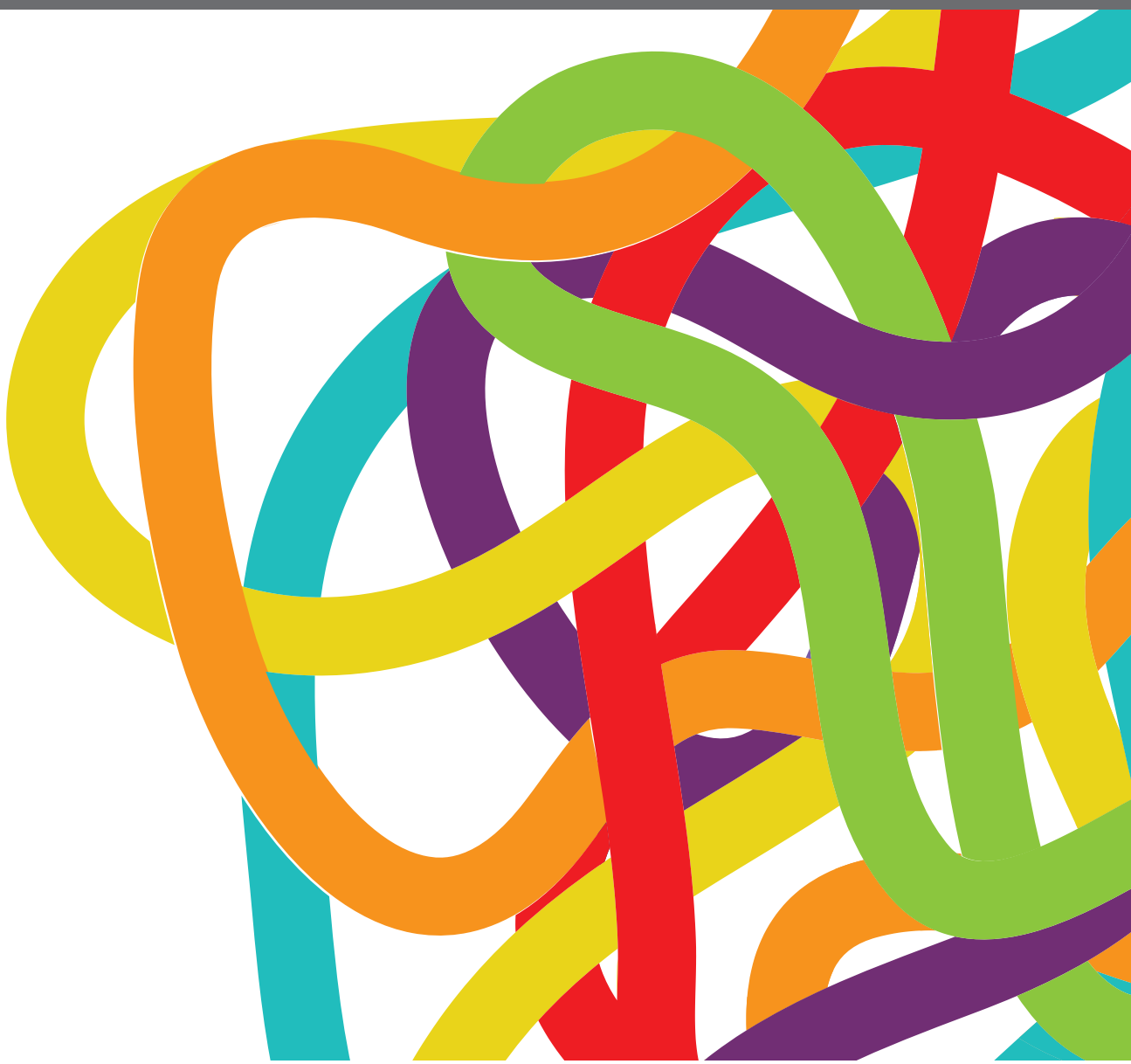
BREAKTHROUGH IN IMAGING-GUIDED PRECISION MEDICINE IN ONCOLOGY

EDITED BY: Laurent Dercle, Samy Ammari, Florent L. Besson, Fatima-Zohra Mokrane, Romain-David Seban and Randy Yeh

PUBLISHED IN: Frontiers in Oncology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-654-5
DOI 10.3389/978-2-88974-654-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

BREAKTHROUGH IN IMAGING-GUIDED PRECISION MEDICINE IN ONCOLOGY

Topic Editors:
Laurent Dercle, Columbia University Irving Medical Center, United States
Samy Ammari, Département d'imagerie médicale, Institut Gustave Roussy, France
Florent L. Besson, Université Paris-Saclay, France
Fatima-Zohra Mokrane, Université Toulouse III Paul Sabatier, France
Romain-David Seban, Institut Curie, France
Randy Yeh, Department of Radiology, Memorial Sloan Kettering Cancer Center, United States

Citation: Dercle, L., Ammari, S., Besson, F. L., Mokrane, F.-Z., Seban, R.-D., Yeh, R., eds. (2022). Breakthrough in Imaging-Guided Precision Medicine in Oncology. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-654-5

Table of Contents

frontiers | Frontiers in Oncology

EDITORIAL
published: 18 May 2022
doi: 10.3389/fonc.2022.908561

Editorial: Breakthrough in Imaging-Guided Precision Medicine in Oncology

Ashley Shaw[1], *Romain-David Seban*[2,3], *Florent L. Besson*[4], *Helena Vila-Reyes*[1], *Samy Ammari*[5], *Fatima-Zohra Mokrane*[6], *Randy Yeh*[7] *and Laurent Dercle*[1]*

[1] Department of Radiology, NewYork-Presbyterian, Columbia University Irving Medical Center, New York, NY, United States, [2] Department of Nuclear Medicine, Institut Curie, Paris, France, [3] Laboratory of Translational Imaging in Oncology, Paris Sciences et Lettres University (PSL) Research University, Institut Curie, Orsay, France, [4] Department of Nuclear Medicine, Université Paris-Saclay, Le Kremlin-Bicêtre, France, [5] Department of Medical Imaging, Institut Gustave Roussy, Villejuif, France, [6] Department of Radiology, Faculté de Médecine Rangueil, Université Toulouse III Paul Sabatier, Toulouse, France, [7] Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Keywords: radiomic, artificial intelligence, deep learning, machine learning, oncology, computed tomography, positron - emission tomography, magnetic resonance imaging

Editorial on the Research Topic

Breakthrough in Imaging-Guided Precision Medicine in Oncology

OPEN ACCESS

Edited by:
Zaver Bhujwalla,
Johns Hopkins Medicine,
United States

***Correspondence:**
Laurent Dercle
laurent.dercle@gmail.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *30 March 2022*
Accepted: *20 April 2022*
Published: *18 May 2022*

Citation:
Shaw A, Seban R-D, Besson FL, Vila-Reyes H, Ammari S, Mokrane F-Z, Yeh R and Dercle L (2022) Editorial: Breakthrough in Imaging-Guided Precision Medicine in Oncology. Front. Oncol. 12:908561. doi: 10.3389/fonc.2022.908561

INTRODUCTION

In the era of precision medicine in oncology, medical imaging is pivotal for a broad spectrum of indications, ranging from early detection of malignant lesions to response assessment in advanced metastatic disease. The Research Topic "*Breakthrough in imaging-guided precision medicine in oncology*" aimed to share disruptive technologies in the field of imaging-guided precision medicine in oncology. The goal was to discuss new concepts and discoveries in the field of imaging biomarkers derived from a quantitative analysis of data contained in medical images. This editorial aims to provide an at-a-glance overview of the 34 articles from 282 authors published in this guest editorial, which collectively support the concept that imaging biomarkers can be used as clinical decision tools, benefiting the outcomes of oncologic patients. Moreover, it provides an overview of research trends at the crossroads between radiology, nuclear medicine, computer science, biochemistry, pathology, and oncology.

NEW IMAGING TECHNOLOGIES FOR PRECISION MEDICINE

Software: Radiomics, Machine-Learning, Deep-Learning, Artificial Intelligence

Capturing the general complexity of tumor biology has become a key challenge in precision medicine. Boosted by the recent evolution of computer science and artificial intelligence (AI) in the field of medical imaging, new powerful image-based analyses have emerged. Radiomics, a large-scale image-based approach derived from OMIC, requires sophisticated workflows consisting of image

acquisition, lesion segmentation, feature extraction, and machine learning that has benefited -and could further benefit- from controlling its variation (Zhao et al.).

In this guest editorial, several authors have unraveled potential indications of radiomics for precision medicine approaches. To face the growing complexity of multidimensional quantitative data, AI appears to be a promising way to assist practitioners in the future. AI can extract and quantify key image information such as its morphological, textural, and molecular features, which can convert subjective qualitative tasks to objective quantitative analysis (Li et al.). For instance, radiomics-based imaging biomarkers, such as 18F-fluciclovine, could be utilized in the detection and management of biochemical-recurrent prostate cancer (Shaikh et al.). Radiomics-based nomogram on 18F-Fluorodeoxyglucose (FDG) positron emission tomography/computed tomography scan (PET/CT) rad-score, combined with clinicopathological factors, was proposed as a new method to personalize management of patients with non-small cell lung cancer. This approach may effectively strengthen the limitations of TNM staging methodology when evaluating lung cancer prognosis (Yang et al.). These radiomics models combining multi-omics analyses could provide a more holistic overview of tumor behavior and predictive capabilities. As an example, combining AI with ultrasound and pathological tests was utilized to tailor personalized therapies for patients with thyroid cancer (Li et al.). Additionally, radiomics-based prediction of histological subtypes was identified through the use of activation maps, which identified specific regions responsible for signature activation and was used to guide treatment decisions in non-small cell lung cancer (Vuong et al.).

There is an ongoing quest toward obtaining reproducibility and generalizable results to characterize image phenotypes (Zhao et al.). Therefore, understanding and controlling the sources of variation is necessary. This requires collaborative multidisciplinary approaches with consistent workflows (Zhao et al.). One approach to mitigate variability is by performing standardized image processing. As illustrated in (Ammari et al.), brain magnetic resonance imaging (MRI) pre-processing techniques are essential to ensure reliable and reproducible radiomics-based models. For instance, the field strength in MRI (1.5T vs 3T) had a significant influence on a wide range of radiomics feature values, and images acquired with a distinct field strength in MRI should not be used interchangeably to build radiomics models (Ammari et al.).

Hardware: New Devices, New Acquisition Protocols

Whatever imaging modality is considered, the improvement of patient management related to hardware evolutions perfectly illustrates the great contribution of technology in precision medicine. For example, preoperative Spectral CT imaging, which utilizes the energy spectrum curve and slope parameters, was beneficial in providing an objective approach to the preoperative staging of thymic epithelial tumors (Zhou et al.). Deep inspiration breath-hold reduced respiratory motion during radiotherapy allowing for more precise and accurate radiotherapy (Naumann et al.). An acquisition protocol incorporating deep inspiration breath-hold with optical surface-guided radiotherapy and image-guided radiotherapy was utilized to ensure reproducibility and accurate tumor localization in surface body radiotherapy (Naumann et al.). Finally, image-guided tissue biopsies using large gauge coaxial needles enabled multiple tissue biopsies in a single pass (Khan et al.). These high-yield samples could be beneficial in better understanding the molecular and genomic characteristics of tumors, ultimately assisting in the development of biomarkers of clinical and translational relevance (Khan et al.).

Multimodal Hybrid Imaging

Multiparametric imaging offers unique opportunities to evaluate tumor characteristics at an advanced multidimensional imaging level. A literature review supported the utility of hybrid 18F-FDG PET/CT imaging modalities in the management of patients with muscle-invasive bladder carcinoma, suggesting that it may be used to guide precision medicine (Girard et al.). Additionally, PET, CT, and MRI used synergistically could provide complementary information with potentially beneficial clinical applications (Decazes et al.).

Metabolic and Molecular Imaging

The pathogenesis of neoplastic tumors and their metabolic processes can be targeted through molecular imaging to better understand and characterize the tumor. PET imaging detection is based on the radionuclide labeling of molecular probes, providing almost unlimited opportunities to map numerous physiological or pathophysiological targeted processes. Numerous relevant molecular probes constitute a powerful arsenal to characterize several tumor biological processes *in vivo*. For example, isocitrate dehydrogenase enzyme (IDH) mutations is a known occurrence in gliomagenesis (Zhou et al.). A metabolic PET/CT-based nomogram model utilized easy-accessible imaging metrics and clinical features and was shown to provide predictive information for IDH mutational status in patients with gliomas (Zhou et al.). Another example is the use of fulvestrant, which is an estrogen receptor antagonist drug approved for postmenopausal women with HR-positive and HER2-negative metastatic breast cancer (Liu et al.). PET with dual tracers 16α-[18F]fluoro-17β-estradiol and 18F-FDG, was utilized as prognostic imaging biomarkers to predict the efficacy of fulvestrant therapy in patients with ER-positive metastatic breast cancer (Liu et al.).

GUIDING INITIAL TREATMENT DECISION

Precision Diagnosis: Grade, Stage, Genomics

The role of imaging in initial staging and treatment planning has been extensively demonstrated. Research combining initial imaging and artificial intelligence, including Radiomics, is a growing field that will help to better address some routine clinical challenges that every physician has to face.

As a radiologist, noninvasive image-based diagnosis can be challenging. As an example, differentiating solitary brain metastasis from glioblastoma multiforme can be difficult, and completely change patient's management. To this end, a radiomics based classifier that evaluates the morphological differences on post-contrast 3DT1 weighted MRI was developed to distinguish between solitary brain metastasis and glioblastoma multiforme (de Causans et al.). Another example is the non-invasive detection of intrahepatic cholangiocarcinoma. Radiomics-based models may improve the diagnostic accuracy of intrahepatic cholangiocarcinoma (Xue et al.). In addition, Radiomics-based biomarkers can offer key insights into precision diagnostics, disease classification, prognostication, and therapeutic management in neuro-oncology (Shaikh et al.).

Lung cancer remains a leading cause of cancer-related deaths (El Ayachy et al.) despite recent therapeutic advances in the field such as immunotherapy. Radiomics was used to identify candidate biomarkers that could enhance our understanding of the microbiology of lung cancer (El Ayachy et al.). In addition, 18F-FDG PET/CT radiomics nomogram combined with clinicopathologic factors could be used to predict survival outcomes in patients with lung adenocarcinoma with an EGFR mutation (Yang et al.). This approach could be utilized to provide a more precise diagnosis and offer personalized treatment guidance for patients with EGFR mutations (Yang et al.).

Another example is the impact of imaging improvements in common cancers such as prostate cancer. Prostate lymph node dissection clinical nomograms have the potential to predict non-regional lymph node metastasis in prostate cancer patients (Jiao et al.). The use of clinical nomograms in conjunction with Ga-PSMA PET/CT could enhance the detection of distant lymph node metastasis and guide clinical decision-making (Jiao et al.).

Baseline Prognostic and Predictive Imaging Biomarkers

Baseline imaging plays a key role in clinical care patient management. Indeed, it allows for assessing tumor burden, potential extension, and can also determine prognosis. The development of radiological or hybrid tools that combine several features will help to better target treatment in order to personalize therapies. For instance, a nomogram and risk stratification system in early-stage cervical cancer patients (stage IB1 and IB2) could have clinical utility in predicting progression-free survival (Xu et al.). SUV peak and HPV-16 were shown to independently impact disease progression (Xu et al.) Artificial Intelligence could also play a key role as prognosis tool. In (Wang et al.), CT radiomics offered an objective approach to predicting the risk of malignant gastrointestinal stromal tumors by using various machine-learning algorithms. Of note, machine-learning models require multicenter testing prior to their use in clinical decision-making (Wang et al.). Metabolic imaging plays a major role as a baseline investigation in clinical routine. It has also a dominant role in predicting response. In (Aide et al.), 18F-FDG PET metabolic heterogeneity of estrogen receptors was used to assist in noninvasively identifying patients with the worst event-free survival in breast cancer.

Baseline imaging is also used to guide treatment strategies in focal therapies such as radiation therapies. CT radiomics could be utilized to localize radioresistant sub-volumes in tumors, serving as a predictive biomarker in radiotherapy (Bogowicz et al.). This methodology could be employed to identify potential targets for dose intensification (Bogowicz et al.).

The use and application of new technologies in the medical field require significant scientific rigor and exact knowledge of limits and bias. This is all the more applicable in this era of artificial intelligence. For instance, it was demonstrated that prognostic radiomics signatures must control for confounding variables to avoid inaccurate prediction of survival outcomes in patients with clear cell renal cell carcinoma (Lu et al.).

ASSESSING TUMOR SENSITIVITY TO TREATMENTS

Immunotherapies have become part of the standard-of-care of a wide range of cancers. These therapies have improved clinical outcomes compared to other treatments, such as chemotherapy or targeted therapies. Unfortunately, only a fraction of patients experiences such therapeutic success. Therefore, the early identification of biomarkers in patients that are unlikely to benefit from immunotherapies is a crucial step in selecting appropriate candidates.

Chimeric antigen receptor (CAR) targeting CD19 antigen, a form of immunotherapy used to treat non-Hodgkin's lymphoma, is frequently complicated by relapses (Vercellino et al.). This strongly suggests further evaluation of the use of medical imaging such as PET/CT for the monitoring of patient response and the optimization of medical care, including risk stratification (Vercellino et al.).

Another clinical drawback with immunotherapies relates to their safety. The incidence of severe immune-related adverse events is not negligible and can even lead to the death of the patient. In (Ederhy et al.) the focus was on cardiovascular toxicities, particularly myocarditis, under immune checkpoints inhibitors. This adverse effect should be recognized promptly due to the high fatality rate (30-50%) (Ederhy et al.). Medical imaging could play an important role in optimizing the management of immune-related myocarditis, including diagnosis, prognostication, treatment decision, and follow-up (Ederhy et al.).

Given the fact that immunotherapy is based on restoring tumor elimination by the immune system, novel patterns of response and progression have emerged, such as pseudoprogression, hyperprogression, abscopal effect, and durable response after treatment. Dissociated response, is an additional pattern defined as heterogeneous-responding lesions in the same patient. While dissociated response has historically been labeled as an unfavorable prognostic pattern, immunotherapy with immune checkpoint inhibitors has revealed that it could in fact be a favorable prognostic pattern (Humbert and Chardin).

New response criteria standardization has been developed, aiming to include these novel patterns of response and progression during the course of immunotherapy. In (Castello et al.), the development of the immune metabolic response criteria by 18FDG PET/CT, called "imPERCIST", was shown to optimize the prediction of clinical benefit in immunotherapy regimens (Castello et al.). These criteria may be beneficial in evaluating the response to immunotherapy, and subsequently guiding clinical decision-making. However, it is important to note that a wide range of novel criteria have been designed for 18F-FDG PET/CT (i.e., PECRIT, PERCIMT, iPERCIST), and have not been prospectively validated. Hence, a large-scale multicenter validation should be performed prior to implementation in the daily routine.

A similar question regarding the response criteria is raised for patients with pancreatic neuroendocrine tumors. Specific criteria should be further developed since the current literature reveals a lack of standardization and comparable methodological approaches in the evaluation of pancreatic neuroendocrine tumors (Partouche et al.).

DIAGNOSIS AND MANAGEMENT OF COVID-19 IN CANCER PATIENTS

The pandemic of COVID-19 has impacted medical, economic, social, and environmental practices worldwide. During its peak, clinicians had to rethink the management of patients with a diagnosis of cancer. On the one hand, cancer patients are at increased risk of serious complications due to their comorbidities, therapy, and immune dysregulation. Therefore, one strategy was to delay medical imaging to avoid exposure and subsequent risk of infection. On the other hand, the role of imaging is critical for cancer management, hence delaying management can be detrimental to the patient's prognosis. In this collection, several papers deciphered strategies to optimize the management of cancer patients while in treatment.

Patients requiring systemic therapies during the pandemic could be screened with reverse transcription-polymerase chain reaction (RT-PCR) prior to initiating therapy and rescreened at 15 days if positive. Those exhibiting respiratory symptoms or signs of acute inflammatory syndrome, could undergo a CT scan or CT angiogram, respectively, to guide clinical decision-making (Viansone et al.). Strategies were proposed so that radiation oncology departments could provide strategies to identify COVID-19 infection and ensure the optimization of patient care by the use of specific workflows, which include RT-PCR, CT scans, and social distancing (Sun et al.). Finally, strategies for screening and early detection of COVID-19 were proposed in patients with hematologic malignancies due to their immunosuppressive state, to mitigate the spread in this patient population (Assi et al.).

CONCLUSION

In conclusion, this guest editorial shared proof of concepts results using disruptive technologies or concepts in the field of imaging-guided precision medicine in oncology. It further demonstrates that medical imaging is pivotal for a broad spectrum of indications. The next step forward will be the prospective validation of these findings in large multicenter prospective studies.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

PERSPECTIVE
published: 03 July 2020
doi: 10.3389/fonc.2020.01267

Screening Strategies for COVID-19 in Patients With Hematologic Malignancies

Tarek Assi[1], *Bachar Samra*[2], *Laurent Dercle*[3], *Elie Rassy*[1], *Joseph Kattan*[1], *Marwan Ghosn*[1], *Roch Houot*[4,5] *and Samy Ammari*[6,7]*

[1] Department of Hematology and Medical Oncology, Faculty of Medicine, Saint-Joseph University, Beirut, Lebanon, [2] Department of Leukemia, The University of Texas MD Anderson Cancer Center, Houston, TX, United States, [3] Radiology Department, Columbia University Medical Center, New York Presbyterian Hospital, New York, NY, United States, [4] Department of Hematology, CHU de Rennes, Université de Rennes, Rennes, France, [5] Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, United States, [6] Radiology Department, Gustave Roussy Cancer Campus, Villejuif, France, [7] BIOMAPS, UMR1281, INSERM.CEA.CNRS, Université Paris-Saclay, Paris, France

COVID-19 has been declared a pandemic by the world health organization. Patients with cancer, and particularly hematologic malignancies may be at higher risk for severe complications due to their malignancy, immune dysregulation, therapy, and associated comorbidities. The oncology community has been proactive in issuing practice guidelines to help optimize management, and limit infection risk and complications from SARS-COV-2. Although hematologic malignancies account for only 10% of all cancers, their management is particularly complex, especially in the time of COVID-19. Screening or early detection of COVID-19 are central for preventative/mitigation strategy, which is the best current strategy in our battle against COVID-19. Herein, we provide an overview of COVID-19 screening strategies and highlight the unique aspects of treating patients with hematologic malignancies.

Keywords: hematologic malignancies, hematology, COVID-19, coronavirus, screening, polymerase chain reaction, CT scan

OPEN ACCESS

Edited by:
Samata Kakkad, Johns Hopkins School of Medicine, United States

Reviewed by:
Dennis Lawrence Cooper, Rutgers Cancer Institute of New Jersey, United States
Francesca Gorini, National Research Council (CNR), Italy

***Correspondence:**
Samy Ammari
samy.ammari@gustaveroussy.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *06 May 2020*
Accepted: *18 June 2020*
Published: *03 July 2020*

Citation:
Assi T, Samra B, Dercle L, Rassy E, Kattan J, Ghosn M, Houot R and Ammari S (2020) Screening Strategies for COVID-19 in Patients With Hematologic Malignancies. Front. Oncol. 10:1267. doi: 10.3389/fonc.2020.01267

INTRODUCTION

In December 2019, infection with a novel betacoronavirus, subsequently named SARS-CoV-2, has been reported in a cluster of pneumonia cases in the Wuhan region of China (1). The pathogen showed a rapid spread that led to a global pandemic and a public health emergency of international concern due to the burden on the healthcare system and lack of specific treatments. By April 26th 2020, the COVID-19 outbreak has affected more than 2.8 million people and claimed the lives of more than 190,000 patients around the globe (2). Patients with cancer are regarded as a more vulnerable population because of their immunosuppressive state induced by the malignancy, therapy, and associated comorbidities (3, 4). As a result, despite the limited available data, the oncology community has been pro-actively engaged in issuing guidelines to help clinical practice with the goal of decreasing exposure and complication risks from COVID-19 among patients with cancer. For instance, recommendations have emerged to consider on a case-by-case basis, whenever possible, options of delaying/dose reducing treatment, using lower-intensity therapy, postponing unnecessary radiological evaluations, and prioritizing telemedicine (5–9).

Hematologic malignancies encompass a wide range of neoplasms, including leukemia, lymphoma, multiple myeloma, myelodysplastic syndrome and myeloproliferative neoplasms, which all account for 10% of all malignancies (10). Although some patients present with indolent

diseases that may not require immediate therapy, many patients harbor aggressive diseases, often life-threatening, which necessitate rapid intervention with intensive chemotherapy, high dose radiation and/or hematopoietic stem cell transplant (HSCT). Prior studies have shown that patients with hematologic malignancies are at higher risk for severe lower respiratory tract infections, notably due to lymphopenia, neutropenia, hypogammaglobinemia, steroid administration, and graft-versus-host disease, it is unclear how this may be applicable to COVID-19 (11–13).

The triage of patients according to their COVID-19 status (confirmed, high- or low level of suspicion) has been the key measure in limiting in-hospital contaminations (14). The absence of a rapid screening system for assessing COVID-19 status in patients is problematic, and the current recommendations are only based on anecdotal and theoretical evidence extrapolated from the management of other infectious diseases. Herein, we provide an overview of COVID-19 screening strategies and highlight the complex aspects of caring for patients with hematologic malignancies.

RISK FACTORS FOR COVID-19 IN PATIENTS WITH HEMATOLOGIC MALIGNANCIES

Literature on COVID-19 in hematologic malignancies is sparse and mainly derived from case series that are focused on patients with solid tumors. Two cases series from China showed a higher incidence (1 vs. 0.3%), and a higher case-fatality rate (5 vs. 1%) in patients with cancer compared with the overall population (4, 15). However, the relative contribution of cancer or cancer therapy to their infectious risk of COVID-19 is uncertain given that the majority of patients either had a remote history of cancer, were not on active anticancer therapy, and/or had multiple comorbidities. Notably, none of the patients had a hematologic malignancy (4, 15).

Patients with hematologic malignancies might be at higher risk of contracting and experiencing complications from COVID-19, such as hospitalization, intensive care unit (ICU)/invasive ventilation, sepsis, cytokine dysregulation, multiorgan failure and/or death (16). Reasons for this higher risk are multifactorial (**Table 1**). First, hematologic malignancies are directly tied to an immunocompromised status due to humoral and cellular immune dysfunction (17–20). Second, a large number of therapies employed for hematologic malignancies are highly immunosuppressive, even more than standard chemotherapies used for solid tumors. Examples include myelosuppressive chemotherapy for acute leukemia, conditioning regimens for hematopoietic stem cell transplant (HSCT), prolonged use of steroids, lymphodepleting agents used for chimeric antigen receptor (CAR) T-cell therapy, and radiation therapy. Third, a high proportion of patients with active hematologic malignancies, especially when on active therapy, are constantly exposed to medical facilities and healthcare staff, which puts them at higher exposure risk. This includes frequent travels and visits for outpatient infusional therapies, laboratory checks or transfusions, longer infusions of certain therapies (anti-CD20 antibodies in lymphoma, daratumumab in multiple myeloma), and the need for hospitalization for other therapies (e.g., induction chemotherapy for acute leukemia). Fourth, many hematologic malignancies occur in older patients (median ages in multiple myeloma, CLL, acute myeloid leukemia, follicular lymphoma, and diffuse large B-cell lymphoma are 72, 70, 68, 65, and 64 years, respectively) with multiple coexisting comorbidities (cardiovascular and lung diseases), which are among the highest risk factors for COVID-19-related morbidity and mortality in large Chinese cohorts (4, 15, 21). Fifth, clinical care for patients with hematologic malignancies requires more medical resources (equipment and staff) than for other patients due to the high risk nature of their diseases such as the need for frequent and close monitoring, transfusion burden, high rate of elective and urgent hospital admissions, and crucial need for clinical trial enrollment. Therefore, as clusters of outbreaks can overwhelm the healthcare system capacity, the detrimental effect may be even more pronounced on patients with hematologic malignancies. Sixth and last, many patients with hematologic malignancies may be directly harmed by travel restrictions affecting delivery of crucial therapies such as stem cells and CAR products from unrelated donors (22, 23). On the other hand, a large proportion of patients with hematologic malignancies, especially acute leukemia or transplant candidates/recipients have been adopting social distancing and preventative precautious measures regardless of the COVID-19 crisis, which may have reduced their exposure risk of COVID-19. Moreover, some targeted therapies commonly used in hematologic malignancies, particularly JAK-STAT pathway inhibitors and bruton tyrosine kinase inhibitors (e.g., ibrutinib), have been postulated to have a protective effect in decreasing virus infectivity (24) and abrogating cytokine-mediated lung injury, respectively (25). Clinical trials are ongoing to further investigate the role of such therapies in patients with COVID-19 (NCT04375397, NCT04320277).

COVID-19 IN PATIENTS WITH HEMATOLOGIC MALIGNANCIES

To our knowledge, at the time of writing this review, only 10 cases of COVID-19 have been published to date in the English literature in patients with hematologic malignancies. Details are summarized in **Table 2** (23, 25–27). Data on screening for COVID-19 in hematologic malignancies are limited. A Chinese cross-sectional survey aimed to evaluate the incidence and outcome of COVID-19 among patients with chronic myeloid leukemia (CML) conducted over 1 week in February 2020 (23). Among 392 patients who took the survey, 12 patients were suspected to have COVID-19 infection based on their clinical presentation, but only 2 cases were confirmed using PCR and CT for an incidence of 0.6%. When the authors classified patients according to CML response milestones (according to the 2020 European Leukemia Net guidelines), the incidence of COVID-19 was 0.3% (1 of 299) among patients with optimal response as opposed to 2% (1 of 50) among patients

TABLE 1 | Unique considerations for COVID-19 in patients with hematologic malignancies.

Patient-related factors
Older age patients
Multiple coexisting comorbidities (cardiovascular and lung diseases)
Disease-related factors
Immunocompromised status due to humoral and cellular immune dysfunction
Treatment-related factors
More frequent use of immunosuppressive therapies
Higher exposure risk to medical facilities and staff (frequent travels/medical visits, transfusions, and the need for hospitalization for certain therapies)
System-related factors
Overwhelmed healthcare system and complexity of care
Potential impact of travel restrictions affecting delivery of crucial therapies such as stem cells and CAR T-cells from unrelated donors
Decreased availability or access to crucial and immediate clinical trials (travel restrictions, trials placed on hold, research team understaffing or re-assignments)

CAR, chimeric antigen receptor.

with poor response raising the possibility that good disease control may be related to a lower incidence of COVID-19 (23). These findings must be interpreted with caution given that the relative contribution of CML diagnosis and therapy (imatinib) as a risk-factor may be questionable. The patient with "suboptimal response" to CML therapy was 89 years old with cardiac history, in hematologic remission and only residual molecular disease on Imatinib and succumbed from COVID-19 due to multiorgan failure including myocardial damage. In addition, tyrosine kinase inhibitors (e.g., imatinib) are not myelosuppressive and not known to increase infectious risks especially in the setting of hematologic and cytogenetic remission such as this case.

The American Society of Hematology Research Collaborative has recently launched a COVID-19 international registry for patients with hematologic malignancies (28). By April 23rd, 2020, there were 64 patients reported to have COVID-19. Underlying malignancies were as follows: acute leukemia (31%), non-Hodgkin lymphoma (22%), chronic lymphocytic leukemias (16%), myeloproliferative neoplasms (16), plasma cell neoplasms (11%), and Hodgkin's lymphoma (6%). Sixty percent of these patients were undergoing active anticancer therapy at the time of COVID-19; among whom 60% were undergoing induction therapy, and 20% were on maintenance/consolidation therapy.

An age-matched comparison of patients with cancer ($n = 105$) and without cancer ($n = 536$) was presented at the American Association for Cancer Research 2020 virtual meeting (29). Multivariable analysis showed that cancer diagnosis was associated with higher rates of death and ICU (intensive care unit) admissions from COVID-19. Patients with hematologic malignancies had the highest severity of symptoms (6 out 9 patients [66.67%]) high risk of ICU admission or use of invasive mechanical ventilation [4 and 2 out 9 patients, respectively [44.4 and 22.2%] and death rate (3 out 9 patients [33%]). Although authors reported that 4 out of the 9 patients with hematologic malignancies had severe immunosuppression, no information was provided on disease status or treatment details. Therefore, the small sample number and lack of detailed clinical data on patients with hematologic malignancies limit generalizability of this observation.

SCREENING STRATEGIES

The COVID-19 pandemic has added another major burden to cancer care of all specialty's units, especially hematology and HSCT units. Due to limited testing and screening, little information is known about the incidence of asymptomatic carriers, who maybe as contagious as symptomatic patients (30–32). Efforts have been made to triage COVID-19 patients to separate units in order to avoid inpatient cross-contamination to other patients, thus, highlighting the importance of screening strategies. Additional testing strategies have been rapidly developed to better identify patients with COVID-19. Since SARS-CoV-2 is an RNA-virus, reverse transcriptase-polymerase chain reaction (RT-PCR) is the obvious diagnostic test to confirm virus shedding (33). However, the sub-optimal sensitivity has limited its reliability in clinical practice. The test can be obtained from nasopharyngeal swabs or sputum samples, and results are generally reported within 4–48 h (34). Nasopharyngeal swabs may lack sensitivity after the first week of COVID-19 infection due to the natural history of the disease. The virus was shown to transiently infect the upper respiratory system until it spreads more toward the lungs. This is extrapolated from data showing a higher rate of RT-PCR positivity $> 90\%$ on days 1–3 of illness, vs. $< 80\%$ at day 6, and vs. $< 50\%$ after day 14 (35). False negatives can also be attributed to the poor quality of the specimen and type of specimen obtained (36). Another appealing screening strategy is serologic testing of SARS-CoV-2 antibodies to identify persons that are potentially "immune" and thus at minimal risk of contagion or complications, both at individual and public health levels (37). This also has important therapeutic implications by identifying potential candidates to donate blood for the preparation of convalescent plasma, an investigational product in the management of severely ill COVID-19 patients (38).

Due to limitations associated with PCR testing, including limited testing capacity, high false-negative rates, and delays in having the results, chest imaging has emerged as a potent diagnostic tool for COVID-19. Chest radiographs have a low sensitivity in early or mild disease and therefore, are not ideal screening tools. In one study, twenty percent of patients had normal chest radiographs at any point during the course of their illness (39). In contrast, chest computed tomography (CT) scan is more sensitive, especially in the presence of typical findings, such as bilateral, peripheral patchy ground-glass opacities, predominantly in the lower lobes (33, 40, 41). Although the American College of Radiology recommends that chest CT scan be reserved for symptomatic patients with suspected COVID-19-related complications and discourages its use for screening purposes, these guidelines may not be applicable for patients with cancer, especially hematologic malignancies undergoing immunosuppressive therapies (42). Since many radiological findings may be seen with other etiologies such as opportunistic infections or drug-induced pneumonitis, the

TABLE 2 | Published cases of COVID-19 in patients with hematologic malignancies.

Disease	Demographics/Therapy	Presentation	Infectious course	Treatment	Outcome	Reference
MM	60/M/thalidomide maintenance	Hypoxia Chest pain + PCR CT: GGO	Nasal O_2 support High IL-6	Antibiotics, steroids, and tocilizumab	Recovery	(26)
CLL	39/M/chlorambucil	Fever/respiratory symptoms + PCR CT: GGO	Non-invasive ventilation	IVIG, steroids, nebulized alpha-interferon. Also resumed Reduced dose chlorambucil	Recovery	(27)
WM	***Pt 1:*** 65/M/ibrutinib ***Pt 2:*** 61/M/ibrutinib ***Pt 3:*** 72/F/ibrutinib ***Pt 4:*** 67/F/ibrutinib ***Pt 5:*** 71/M/ibrutinib ***Pt 6:*** 58/M/ibrutinib	***Pt 1–5:*** Fever/respiratory symptoms ***Pt 6:*** Fever/respiratory symptoms CT: GGO	***Pt 1:*** ICU ***Pt 2:*** Not hospitalized ***Pt 3:*** Not hospitalized ***Pt 4:*** Not hospitalized ***Pt 5:*** Not hospitalized ***Pt 6:*** Mechanical ventilation and intensive care unit	***Pt 1:*** HCQ+AZ ***Pt 2–5:*** None ***Pt 6:*** HCQ+AZ+ IVIG, and tocilizumab All pts continued Ibrutinib	***Pt 1:*** Improved ***Pt 2:*** Recovery ***Pt 3:*** Recovery ***Pt 4:*** Recovery ***Pt 5:*** Recovery ***Pt 6:*** Improved	(25)
CML	***Pt 1:*** 89/F/Imatinib ***Pt 2:*** 47/M/HQP1351	***Pt 1:*** Respiratory symptoms + CT: GGO ***Pt 2:*** Fever/respiratory symptoms + PCR CT: negative	***Pt 1:*** Respiratory and renal failure. Myocardial damage ***Pt 2:*** Hospitalized. Mild course	NA	***Pt 1:*** Death 3 days after hospitalization ***Pt 2:*** Recovery	(23)

MM, multiple myeloma; CLL, chronic lymphocytic leukemia; WM, Waldenstrom macroglobulinemia; CML, chronic myeloid leukemia; M, male; F, female; PCR, polymerase-chain reaction; CT, computed tomography; GGO, ground-glass opacities; IVIG, intravenous immunoglobulin; O_2, oxygen; HCQ, hydroxychloroquine; AZ, azythromycine.

TABLE 3 | Screening strategies for COVID-19 patients.

Screening Test	Advantages	Limitations
RT-PCR	Ease of testing Standardized	High false negative rates Limited accessibility Delay in test result
Chest-X Ray	Low cost Easy access Rapid result	Low sensitivity Radiation exposure
Chest CT scan	High sensitivity Great tool for AI techniques Rapid result	Low specificity Radiation exposure
Antibody serology	identification of immune response Potential therapeutic implication (plasma donors)	Not standardized Limited access

role of radiologists is primordial in confirming diagnosis of COVID-19 (43). Vascular thickening, and peripheral distribution seem to be more characteristic of COVID-19 than other viral infections (44). The concordance between PCR and CT scan has been addressed in a Chinese cohort from the Wuhan region of 1,014 patients with suspected COVID-19. Ninety-seven percent of patients had positive CT scans. However, important observations were notable. Initially, CT scan was more sensitive in detecting early infections (88%, compared with 59% with PCR). Moreover, CT scan was better in assessing recovery; 42% had signs of radiological improvement before RT-PCR turned negative (33). Therefore, CT scan may have a higher yield than PCR for diagnosis and monitoring. Similar findings were shown by Fang et al. (40) (chest CT sensitivity of 98 vs. 71% with PCR; $p < 0.001$). One meta-analysis showed a positive predictive value reaching 90.4% for chest CT scan. Taken together, these findings support the superiority of CT over PCR for detection and screening of COVID-19 especially in endemic areas and for higher-risk patients such as those with hematologic malignancies. Nonetheless, many challenges remain before implementation to screen for asymptomatic carriers (**Table 3**) (45). Artificial intelligence (AI) techniques may increase the sensitivity and specificity of imaging tools to screen for COVID-19 infections (46). Deep learning techniques have been successfully used to differentiate bacterial and viral infections with specific lung injuries, suggesting that it could be a promising tool for the diagnosis and screening of COVID-19 (47, 48). AI can improve image acquisition through automation of the scanning procedure and reshaping of the workflow, which would limit the human-to-human contact and avoid virus transmission. These techniques can also increase the accuracy of the COVID-19 diagnosis through the delineation

of suspected infection on chest CT scan and radiography with different segmentation methods (49). Deep learning tools for the detection of COVID-19 infections are currently being developed with promising results such as the COVID-19 detection neural network (COVNet) (50).

PUBLISHED GUIDELINES FOR COVID-19 SCREENING

Despite the paucity of current available data, several medical societies have provided clinical practice guidelines for management of patients with hematologic malignancies in order to optimize therapy and limit infection risk and complications (9, 51). Consensus recommendations have been based on extrapolated information from other coronavirus epidemics and expert opinions based on educated assumptions. Social distancing, quarantine measures, clear guidance on the importance of hand hygiene, and avoiding high-risk exposures are recommended to halt the virus transmission (52). Preventative strategies also include screening for COVID-19 symptoms or fever at the hospital entrance for all patients, caregiver/visitors, and healthcare workers (53). Serial RT-PCR screening of healthcare workers, once every week or every 2 weeks can be another proposed measure for infection control (54). Initial evaluation of patients with hematologic malignancies or undergoing hematopoietic cell transplantation should at least include clinical history and examination, and laboratory and radiological evaluation to document active respiratory viral infections (55, 56). In patients with acute myeloid leukemia, up to 50–75% of patients present with fever and thus are at risk of delayed or missed diagnosis (57, 58). The European Hematology Association recommends testing for COVID-19 with RT-PCR in all newly diagnosed patients with acute and chronic leukemia undergoing intensive chemotherapies and before every cycle of therapy. No recommendations regarding the screening role of chest CT scan have been issued (5). According to the American Society of Hematology guidelines, screening of patients with chronic lymphocytic leukemias presenting with mild symptoms depends on the availability of testing and the need to isolate those who test positive (59). Regarding multiple myeloma, it is generally recommended to screen all newly diagnosed patients for COVID-19 in the inpatient and outpatient settings (60, 61). Screening is widely used and strongly recommended before autologous and allogeneic HSCT (6, 62).

According to the European Society for Blood and Marrow Transplantation guidelines, COVID-19 should be ruled-out in all patients, including asymptomatic ones, before undergoing conditioning chemotherapy (62). Additionally, patients planned to receive CAR-T cell therapy or HSCT should abide by home isolation for at least 2 weeks before lymphodepleting or conditioning chemotherapy. In case of close exposure to a patient with confirmed COVID-19, any elective therapy should be deferred for at least 2–3 weeks before undergoing a new RT-PCR screening to confirm negativity (62, 63). As for donors, COVID-19 positive patients should be excluded while those in contact should be isolated for 4 weeks unless the patient is asymptomatic and the procedure is urgent, then it can be considered after negative testing.

Screening measures in the pre-CAR T cell phase should include screening for symptoms for all patients before apheresis and CAR T cell infusion. RT-PCR testing is considered within 48–72 h before apheresis and recommended within 48–72 h before lymphodepleting chemotherapy and 7 days within CAR T cell infusion. A repeat evaluation of RT-PCR within 72 h of CAR T cell infusion to detect interim COVID-19 infection and perform serologic testing for COVID-19 seroconversion once available (22, 64). That being said, stem cell and CAR-T cell recipient patients residing in high-endemic areas or who have been in contact with suspected COVID-19 patients should undergo RT-PCR testing while those who tested positive should undergo a CT scan and oxygen need evaluation (62). The American Society for Transplantation and Cellular Therapy recommends that asymptomatic patients whose cellular therapy cannot be delayed should be tested within 72 h before their admission or initiation of the conditioning regimen. Also, asymptomatic allogenic donors should be screened 72 h before collection while autologous donors with chemo-mobilization should be screened 72 h before chemotherapy and cell collection and those on GCSF and plerixafor on the day of GCSF start, respectively (6).

Last, blood transfusions should be carefully monitored during this outbreak. Blood donors should be screened for any suspicious signs or close exposure to COVID-19 present in at least the past 2 weeks (65). The WHO has issued specific guidelines on maintenance and safety of blood products supply during the COVID-19 pandemic (66).

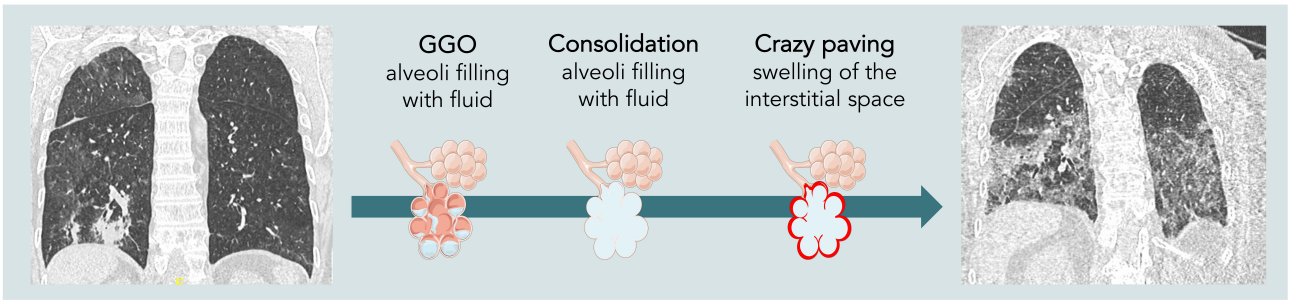
CONCLUSION

Patients with hematologic malignancies are thought to be particularly vulnerable for severe illness from COVID-19. The outbreak has added a large burden to the complexity of cancer care in hematologic malignancies on many aspects. Balancing the management of COVID-19 and its complications and management of cancer is particularly challenging in patients with hematologic malignancies, especially when treating with curative intent, which is often the case in many blood cancers. In the context of limited available scientific data in a such young and rapidly evolving field, physicians are left for now with their clinical judgement, educated guesses and consensual expert opinions (**Figure 1**). Until the advent of effective anti-viral therapy or vaccines, preventative behavioral strategy remains our best model to help protect our patients. Screening for asymptomatic carriers and immune persons has a central role in optimizing mitigation strategy to maintain safety at individual and public health levels. Despite the availability of RT-PCR, antibody serology, and CT scan as complementary diagnostic tools, several limitations exist in implementing a cost-effective

A Proposed screening strategy for COVID-19 in patients with hematological malignancies

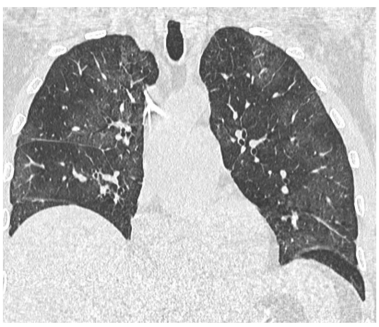
General considerations for all patients	- Clinical evaluation for signs and symptoms (Fever, dyspnea, cough, sputum production) - Preventative measures (social distancing, hand washing, face masks) - Discuss cancer therapy after thorough assessment (Evaluation of risk/benefit of observation vs treatment, delaying therapy, evaluation of disease bulk/risk vs. COVID-19 risk) - Favor out-patient > in-patient treatments - Favor less immunosuppressive therapy > dose-dense immunosuppressive therapy when appropriate - Delay elective procedures or treatment until COVID-19 is ruled out (stem cell collection, leukapheresis)			
Risk level based on management	High Risk - HSCT - CAR-T therapy - Intensive and/or highly immunosuppressive therapies		Intermediate Risk - Standard chemotherapy - Monoclonal antibodies (including anti-CD20) - High dose steroids - Moderately immunosuppressive therapies	Lower-risk - Observation - Supportive care - Radiotherapy - Mildly/non immunosuppressive therapies (TKI,…)
Recommendations before initiation	PCR testing Baseline Chest CT		PCR testing	No screening if asymptomatic

B CT-scan (high sensitivity, low specificity) ± pulmonary angiogram (pulmonary embolism)



C COVID-19 pneumonia

D Pneumocystosis



E Aspergillosis

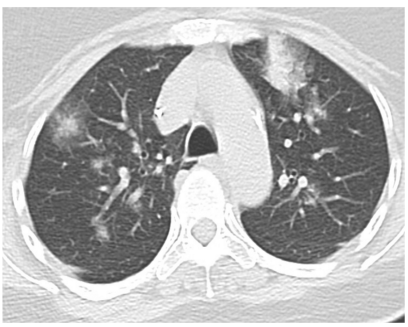


FIGURE 1 | (A) Proposed screening strategy for COVID-19 in patients with hematological malignancies. **(B)** CT-scan is a sensitive tool, but its specificity for the diagnosis of COVID-19 is around 37%, according to recent meta-analysis. It has not been evaluated in patients with hematological malignancies but could be even lower due to a higher rate of infection caused by immune suppression. **(C)** COVID-19 pneumonia in a 58-year-old patient with diffuse B lymphoma treated with R

(Continued)

FIGURE 1 | ACVBP. Symptoms were cough and fever. SARS-CoV-2 rt-PCR was positive. On CT-scan, the CT score CORADS was 5. The axial CT image showed ordinary COVID-19 pneumonia with multiple regions of subpleural GGO (Ground Glass opacity) with superimposed inter and intralobular septal thickening. **(D)** Pneumocystis in a 58-year-old patient followed for acute leukemia. The clinical exam revealed a cough and fever. The patient was lymphopenic. CT-scan showed diffused ground-glass opacities. SARS-CoV-2 rt-PCR was negative. The final diagnosis was pneumocystis. **(E)** Diffuse pulmonary condensation in a 55-year-old patient with AML in febrile aplasia. SARS-CoV-2 rt-PCR was negative. A positive aspergillus antigenemia was confirmed by culture of bronchoalveolar lavage fluid.

screening strategy that can be applicable on a large scale. A tiered approach targeting patients at pre-defined higher risk (pre-conditioning or other intensive chemotherapy regimens, before prolonged use of steroids or anti-CD20 antibodies etc.) may be a better way and warrants more exploration. Despite the numerous difficulties during the pandemic, the oncology community has been very adaptive in adjusting to the new challenges. In addition, scientific and technological progress is evolving rapidly, which gives hope to the oncology, and international communities.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors contributed equally to the manuscript writing and review.

REFERENCES

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. (2020) 382:727–33. doi: 10.1056/NEJMoa2001017
2. WHO. *Coronavirus Disease 2019 (COVID-19) Situation Report – 97*. (2019). Available online at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200426-sitrep-97-covid-19.pdf?sfvrsn=d1c3e800_2 (accessed May 05, 2020).
3. Garnica M, Maiolino A. COVID and Hematology: special considerations regarding patient safety, gold standard therapies and safety for health care professionals. *Hematol Transfus Cell Ther*. (2020) 42:111–2. doi: 10.1016/j.htct.2020.04.001
4. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol*. (2020) 21:335–7. doi: 10.1016/S1470-2045(20)30096-6
5. EHA guidelines. *Recommendations for Specific Hematologic Malignancies*. (2020). Available online at: https://ehaweb.org/covid-19/covid-19-recommendations/recommendations-for-specific-hematologic-malignancies/ (accessed May 05, 2020).
6. ASTCT (American Society for Transplantation and Cellular Therapy). *Guidance for BMT Center Functioning During COVID-19 Pandemic*. (2020). Available online at: https://www.astct.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=5d66683f-c808-4eaf-faa1-782a5bfd687b&forceDialog=0 (accessed May 05, 2020).
7. Lambertini M, Toss A, Passaro A, Criscitiello C, Cremolini C, Cardone C, et al. Cancer care during the spread of coronavirus disease 2019 (COVID-19) in Italy: young oncologists' perspective. *EMSO Open*. (2020) 5:e000759. doi: 10.1136/esmoopen-2020-000759
8. Portnoy J, Waller M, Elliott T. Telemedicine in the Era of COVID-19. *J Allergy Clin Immunol Pract*. (2020) 8:1489–91. doi: 10.1016/j.jaip.2020.03.008
9. Perini GF, Fischer T, Gaiolla RD, Rocha TB, Bellesso M, Teixeira LLC, et al. How to manage lymphoid malignancies during novel 2019 coronavirus (CoVid-19) outbreak: a Brazilian task force recommendation. *Hematol Transfus Cell Ther*. (2020) 42:103–10. doi: 10.1016/j.htct.2020.04.002
10. Leukemia and lymphoma society. *Facts and Statistics*. (2020). Available online at: https://www.lls.org/facts-and-statistics/facts-and-statistics-overview/facts-and-statistics (accessed May 05, 2020).
11. Hakki M, Rattray RM, Press RD. The clinical impact of coronavirus infection in patients with hematologic malignancies and hematopoietic stem cell transplant recipients. *J Clin Virol*. (2015) 68:1–5. doi: 10.1016/j.jcv.2015.04.012
12. Weinkove R, McQuilten Z, Adler J, Agar M, Blyth E, Cheng A, et al. Managing haematology and oncology patients during the COVID-19 pandemic: interim consensus guidance. *Med J Aust*. (2020) 212:481–9. doi: 10.5694/mja2.50607
13. NHS. *Clinical Guide for the Management of Cancer Patients During the Coronavirus Pandemic* (2020).
14. Day M. Covid-19: identifying and isolating asymptomatic people helped eliminate virus in Italian village. *BMJ*. (2020) 368:m1165. doi: 10.1136/bmj.m1165
15. Zhang L, Zhu F, Xie L, Wang C, Wang J, Chen R, et al. Clinical characteristics of COVID-19-infected cancer patients: a retrospective case study in three hospitals within Wuhan, China. *Ann Oncol*. (2020) 31:894–901. doi: 10.1016/j.annonc.2020.03.296
16. Willan J, King AJ, Hayes S, Collins GP, Peniket A. Care of haematology patients in a COVID-19 epidemic. *Br J Haematol*. (2020) 189:241–3. doi: 10.1111/bjh.16620
17. Griggio V, Mandili G, Vitale C, Capello M, Macor P, Serra S, et al. Humoral immune responses toward tumor-derived antigens in previously untreated patients with chronic lymphocytic leukemia. *Oncotarget*. (2017) 8:3274–88. doi: 10.18632/oncotarget.13712
18. Advani SH, Dinshaw KA, Nair CN, Gopal R, Talwalkar GV, Iyyer YS, et al. Immune dysfunction in non-Hodgkin's lymphoma. *Cancer*. (1980) 45:2843–8. doi: 10.1002/1097-0142(19800601)45:11<2843::AID-CNCR2820451121>3.0.CO;2-C
19. Kennedy-Nasser AA, Hanley P, Bollard CM. Hodgkin disease and the role of the immune system. *Pediatr Hematol Oncol*. (2011) 28:176–86. doi: 10.3109/08880018.2011.557261
20. Noonan K, Borrello I. The immune microenvironment of myeloma. *Cancer Microenviron Off J Int Cancer Microenviron Soc*. (2011) 4:313–23. doi: 10.1007/s12307-011-0086-3
21. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. (2020) 70:7–30. doi: 10.3322/caac.21590
22. Bachanova V, Bishop MR, Dahi P, Dholaria B, Grupp SA, Hayes-Lattin B, et al. Chimeric antigen receptor T cell therapy during the COVID-19 pandemic. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant*. (2020) 26:1239–46. doi: 10.1016/j.bbmt.2020.04.008
23. *The First Report of the Prevalence of COVID-19 in Chronic Myelogenous Leukemia Patients in the Core Epidemic Area of China:Multicentre, Cross-Sectional Survey | medRxiv*. (2020). Available online at: https://www.medrxiv.org/content/10.1101/2020.03.12.20034876v2 (accessed May 05, 2020).
24. Richardson P, Griffin I, Tucker C, Smith D, Oechsle O, Phelan A, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *Lancet Lond Engl*. (2020) 395:e30. doi: 10.1016/S0140-6736(20)30304-4
25. Treon SP, Castillo J, Skarbnik AP, Soumerai JD, Ghobrial IM, Guerrera ML, et al. The BTK-inhibitor ibrutinib may protect against pulmonary injury in COVID-19 infected patients. *Blood*. (2020) 135:1912–5. doi: 10.1182/blood.2020006288

26. Zhang X, Song K, Tong F, Fei M, Guo H, Lu Z, et al. First case of COVID-19 in a patient with multiple myeloma successfully treated with tocilizumab. *Blood Adv*. (2020) 4:1307–10. doi: 10.1182/bloodadvances.2020001907
27. Jin XH, Zheng KI, Pan KH, Xie YP, Zheng MH. COVID-19 in a patient with chronic lymphocytic leukaemia. *Lancet Haematol*. (2020) 7:e351–2. doi: 10.1016/S2352-3026(20)30074-0
28. *COVID-19 Registry*. (2020). Available online at: https://www.ashresearchcollaborative.org/covid-19-registry (accessed May 05, 2020).
29. Dai M, Liu D, Liu M, Zhou F, Li G, Chen Z, et al. Patients with cancer appear more vulnerable to SARS-COV-2: a multi-center study during the COVID-19 outbreak. *Cancer Discov*. (2020) 10:783–91. doi: 10.1158/2159-8290.CD-20-0422
30. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *N Engl J Med*. (2020) 382:970–1. doi: 10.1056/NEJMc2001468
31. Bai Y, Yao L, Wei T, Tian F, Jin DY, Chen L, et al. Presumed asymptomatic carrier transmission of COVID-19. *JAMA*. (2020) 323:1406–7. doi: 10.1001/jama.2020.2565
32. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic transmission, the Achilles' heel of current strategies to control COVID-19. *Mass Medical Soc*. (2020) 382:2158–60. doi: 10.1056/NEJMe2009758
33. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*. (2020). doi: 10.1148/radiol.2020200642. [Epub ahead of print].
34. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3
35. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). *Clin Infect Dis Off Publ Infect Dis Soc Am*. (2020). doi: 10.1093/cid/ciaa310. [Epub ahead of print].
36. Liang T. *Handbook of COVID-19 Prevention and Treatment*.Zhejiang University School of Medicine (2020).
37. FDA. *Coronavirus (COVID-19) Update: Serological Test Validation and Education Efforts*. (2020). Available online at: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-serological-test-validation-and-education-efforts (accessed May 05, 2020).
38. Chen L, Xiong J, Bao L, Shi Y. Convalescent plasma as a potential therapy for COVID-19. *Lancet Infect Dis*. (2020) 20:398–400. doi: 10.1016/S1473-3099(20)30141-9
39. Wong HYF, Lam HYS, Fong AHT, Leung ST, Chin TWY, Lo CSY, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiology*. (2020) 2:201160. doi: 10.1148/radiol.2020201160
40. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*. (2020) 2:200432. doi: 10.1148/radiol.2020200432
41. Raptis CA, Hammer MM, Short RG, Shah A, Bhalla S, Bierhals AJ, et al. Chest CT and Coronavirus disease (COVID-19): a critical review of the literature to date. *Am J Roentgenol*. (2020) 1–4. doi: 10.2214/AJR.20.23202. [Epub ahead of print].
42. American College of Radiology (ACR). *ACR Recommendations for the Use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection*. (2020) Available online at: www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CTfor-Suspected-COVID19-Infection (accessed May 05, 2020).
43. Kligerman SJ, Franks TJ, Galvin JR. From the radiologic pathology archives: organization and fibrosis as a response to lung injury in diffuse alveolar damage, organizing pneumonia, and acute fibrinous and organizing pneumonia. *Radiographics*. (2013) 33:1951–75. doi: 10.1148/rg.337130057
44. Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TML, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology*. (2020). doi: 10.1148/radiol.2020200823. [Epub ahead of print].
45. Bao C, Liu X, Zhang H, Li Y, Liu J. COVID-19 Computed tomography findings: a systematic review and meta-analysis. *J Am Coll Radiol*. (2020) 17:701–9. doi: 10.1016/j.jacr.2020.03.006
46. Bullock J, Pham KH, Lam CSN, Luengo-Oroz M. Mapping the landscape of artificial intelligence applications against COVID-19. *ArXiv Prepr*. (2020).
47. Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. (2018) 172:1122–31. doi: 10.1016/j.cell.2018.02.010
48. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans Med Imaging*. (2016) 35:1207–16. doi: 10.1109/TMI.2016.2535865
49. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, et al. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. *ArXiv Prepr*. (2020). doi: 10.1109/RBME.2020.2987975. [Epub ahead of print].
50. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct. *Radiology*. (2020) 284:574–82. doi: 10.1148/radiol.2020200905
51. Zic JA, Ai W, Akilov OE, Carter JB, Duvic M, Foss F, et al. United States Cutaneous lymphoma consortium recommendations for treatment of cutaneous lymphomas during the COVID-19 pandemic. *J Am Acad Dermatol*. (2020). doi: 10.1016/j.jaad.2020.04.049. [Epub ahead of print].
52. Al-Shamsi HO, Alhazzani W, Alhuraiji A, Coomes EA, Chemaly RF, Almuhanna M, et al. A practical approach to the management of cancer patients during the novel coronavirus disease 2019 (COVID-19) pandemic: an international collaborative group. *Oncologist*. (2020) 25:e936–45. doi: 10.1634/theoncologist.2020-0213
53. van de Haar J, Hoes LR, Coles CE, Seamon K, Fröhling S, Jäger D, et al. Caring for patients with cancer in the COVID-19 era. *Nat Med*. (2020) 26:665–71. doi: 10.1038/s41591-020-0874-8
54. ecancer. *What to Expect: Oncology's Response to Coronavirus in Italy - Ecancer*. (2020). Available online at: http://ecancer.org/en/news/17495-what-to-expect-oncologys-response-to-coronavirus-in-italy (accessed May 05, 2020).
55. Hirsch HH, Martino R, Ward KN, Boeckh M, Einsele H, Ljungman P. Fourth European Conference on Infections in Leukaemia (ECIL-4): guidelines for diagnosis and treatment of human respiratory syncytial virus, parainfluenza virus, metapneumovirus, rhinovirus, and coronavirus. *Clin Infect Dis*. (2013) 56:258–66. doi: 10.1093/cid/cis844
56. Dignan FL, Clark A, Aitken C, Gilleece M, Jayakar V, Krishnamurthy P, et al. BCSH/BSBMT/UK clinical virology network guideline: diagnosis and management of common respiratory viral infections in patients undergoing treatment for haematological malignancies or stem cell transplantation. *Br J Haematol*. (2016) 173:380–93. doi: 10.1111/bjh.14027
57. Gavillet M, Carr Klappert J, Spertini O, Blum S. Acute leukemia in the time of COVID-19. *Leuk Res*. (2020) 92:106353. doi: 10.1016/j.leukres.2020.106353
58. Burke PJ, Braine HG, Rathbun HK, Owens JA. The clinical significance and management of fever in acute myelocytic leukemia. *Johns Hopkins Med J*. (1976) 139:1–12.
59. ASH Guidelines. *COVID-19 and CLL: Frequently Asked Questions*. (2020). Available online at: https://www.hematology.org/covid-19/covid-19-and-cll (accessed May 05, 2020).
60. Al Saleh AS, Sher T, Gertz MA. Multiple myeloma in the time of COVID-19. *Acta Haematol*. (2020). doi: 10.1159/000507690. [Epub ahead of print].
61. Mian H, Grant SJ, Engelhardt M, Pawlyn C, Bringhen S, Zweegman S, et al. Caring for older adults with multiple myeloma during the COVID-19 Pandemic: Perspective from the International Forum for Optimizing Care of Older Adults with Myeloma. *J Geriatr Oncol*. (2020) 11:764–8. doi: 10.1016/j.jgo.2020.04.008
62. Styczynski JMM, Ljungman P. EBMT recommendation on: CORONAVIRUS DISEASE COVID-19: EBMT (2020). Available online at: https://www.ebmt.org/sites/default/files/2020-03/EBMT%20COVID19%20guidelines%20v.3.2%20%282020-03-16%29.pdf
63. Dholaria B, Savani BN. How do we plan hematopoietic cell transplant and cellular therapy with the looming COVID-19 threat? *Br J Haematol*. (2020) 189:239–40. doi: 10.1111/bjh.16597
64. Frey NV, Shaw PA, Hexner EO, Pequignot E, Gill S, Luger SM, et al. Optimizing chimeric antigen receptor T-cell therapy for

adults with acute lymphoblastic leukemia. *J Clin Oncol.* (2020) 38:415–22. doi: 10.1200/JCO.19.01892

65. Mascaretti L, De Angelis V, Berti P. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic and Transfusion Medicine: reflections from Italy. *Blood Transfus.* (2020) 18:77–8. doi: 10.2450/2020.0071-20
66. WHO. *Maintaining a Safe and Adequate Blood Supply During the Pandemic Outbreak of Coronavirus Disease (COVID-19).* (2020). Available online at: https://www.who.int/publications-detail/maintaining-a-safe-and-adequate-blood-supply-during-the-pandemic-outbreak-of-coronavirus-disease-(covid-19) (accessed May 05, 2020).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 17 July 2020
doi: 10.3389/fonc.2020.01042

Development and Validation of a Radiomics Nomogram Based on ^{18}F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography and Clinicopathological Factors to Predict the Survival Outcomes of Patients With Non-Small Cell Lung Cancer

Bin Yang[1†], Jian Zhong[1†], Jing Zhong[1†], Lu Ma[1†], Ang Li[1], Hengshan Ji[2], Changsheng Zhou[1], Shaofeng Duan[3], Qinggen Wang[4], Chaohui Zhu[5], Jiahe Tian[6], Longjiang Zhang[1], Feng Wang[7], Hong Zhu[2] and Guangming Lu[1]**

[1] Department of Medical Imaging, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing, China, [2] Department of Nuclear Medicine, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing, China, [3] GE Healthcare China, Shanghai, China, [4] Department of Medical Imaging, Jinling Hospital, The First School of Clinical Medicine, Southern Medical University, Nanjing, China, [5] Department of Nuclear Medicine, Peking Union Medical College Hospital, Beijing, China, [6] Department of Nuclear Medicine, The Chinese PLA General Hospital, Beijing, China, [7] Department of Nuclear Medicine, First People's Hospital of Nanjing, Nanjing, China

Purpose: In this study, we developed and validated a radiomics nomogram by combining the radiomic features extracted from ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) images and clinicopathological factors to evaluate the overall survival (OS) of patients with non-small cell lung cancer (NSCLC).

Patients and Methods: A total of 315 consecutive patients with NSCLC (221 in the training cohort and 94 in the validation cohort) were enrolled in this study. A total of 840 radiomic features were extracted from the CT and PET images. Three radiomic scores (rad-scores) were calculated using the least absolute shrinkage and selection operator (LASSO) Cox regression based on subsets of CT, PET, and PET/CT radiomic features. A multivariate Cox regression analysis was performed for each rad-score combined with clinicopathological factors to determine the independent risk factors. The OS nomogram was constructed based on the PET/CT rad-score and independent clinicopathological factors. Validation and calibration were conducted to evaluate the performance of the model in the training and validation cohorts, respectively.

Results: A total of 144 (45.71%) women and 171 (54.29%) men with NSCLC were enrolled in this study. The PET/CT rad-score combined with the clinical model had the best C-index (0.776 and 0.789 for the training and validation cohorts, respectively). Distant metastasis, stage, carcinoembryonic antigen (CEA), and targeted therapy were independent risk factors for patients with NSCLC. The validation curve

OPEN ACCESS

Edited by:
Florent L. Besson, Université Paris-Saclay, France

Reviewed by:
Mathieu Hatt, Institut National de la Santé et de la Recherche Médicale (INSERM), France
Clément Bailly, CHU de Nantes, France

***Correspondence:**
Hong Zhu
zh_zy@163.com
Guangming Lu
cjr.luguangming@vip.163.com

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *06 March 2020*
Accepted: *26 May 2020*
Published: *17 July 2020*

Citation:
Yang B, Zhong J, Zhong J, Ma L, Li A, Ji H, Zhou C, Duan S, Wang Q, Zhu C, Tian J, Zhang L, Wang F, Zhu H and Lu G (2020) Development and Validation of a Radiomics Nomogram Based on ^{18}F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography and Clinicopathological Factors to Predict the Survival Outcomes of Patients With Non-Small Cell Lung Cancer. Front. Oncol. 10:1042. doi: 10.3389/fonc.2020.01042

showed that the OS nomogram had a strong predictive power in patients' survival. The calibration curve showed that the predicted survival time was significantly close to the observed one.

Conclusion: A radiomic nomogram based on ^{18}F-FDG PET/CT rad-score and clinicopathological factors had good predictive performance for the survival outcome, offering feasible, and practical guidance for individualized management of patients with NSCLC.

Keywords: non-small cell lung cancer, PET/CT, radiomics, survival outcome, risk stratification

INTRODUCTION

Lung cancer is a malignant tumor with the highest morbidity and mortality worldwide (1). Non-small cell lung cancer (NSCLC) is the most common pathological type of lung cancer, accounting for ∼85% of all patients with lung cancer (1, 2). Considering that early signs and symptoms of NSCLC do not manifest in some patients, ∼70% of patients have developed metastasis at the time of diagnosis and thus have lost the opportunity for surgical treatment (3, 4). The tumor-node-metastasis (TNM) staging system is currently the most commonly used tumor staging system worldwide and is considered to be the most valuable method for assessing the prognosis of malignant tumors (5–7). However, the TNM staging system still has several limitations when used to evaluate lung cancer prognosis in clinical practice. There are notable differences in the prognosis of tumors in the same stage, indicating that the TNM staging system cannot be used alone to fully evaluate the prognosis of patients with NSCLC. Thus, a comprehensive analysis of the TNM staging system in combination with other tumor biological characteristics that affect the prognosis of patients with NSCLC should be performed (8, 9). Therefore, determining additional effective prognostic indicators other than the TNM staging system, evaluating patients' responses to treatment at an early stage, and predicting the overall survival (OS) of patients are considered important to achieve individualized medical treatments.

With the development of genomic biology and technology, survival-related genomic characteristics have been included in the prognostic evaluation of several diseases, thereby improving the accuracy of the prognostic evaluation of several patients. However, the main limitation of these invasive technologies is that they cannot capture comprehensive information on the spatiotemporal heterogeneity of tumors (10–13). Therefore, an effective method is urgently required to comprehensively quantify the spatiotemporal heterogeneity of tumors and to evaluate the prognosis of several diseases. ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) is an important imaging method widely used for functional metabolic and anatomical/morphological imaging of various types of malignant tumors and metastatic lesions. ^{18}F-FDG PET/CT provides not only intuitive imaging differences through image comparisons, but also several metabolic parameters to distinguish metabolically active or inactive tumor tissues. In particular, PET/CT has been widely used in clinical practice for the establishment of diagnosis, staging, efficacy monitoring, and prognostic evaluation of NSCLC (14–16). Several studies have confirmed that the FDG uptake of primary tumors is an independent risk factor for patients with early NSCLC (17, 18), but its application value in the prognostic evaluation of NSCLC is still controversial (19, 20). As an emerging and promising image analysis tool, radiomics is a non-invasive quantitative research method that can be used to convert medical images into mineable data for the identification of tumor heterogeneity. The integration of genetic pathology and imaging multimodality could improve the non-invasive quantitative analysis of tumor spatiotemporal heterogeneity and microenvironment (21, 22). Studies have shown that radiomics may have good predictive prognostic performance and decision support in oncology (23, 24). In previous studies, the texture characteristics or radiomics based on ^{18}F-FDG PET/CT have been used to predict the EGFR and KRAS mutation status in patients with NSCLC, to evaluate NSCLC radiation tumor response, to predict the prognosis of patients with NSCLC after stereotactic body radiotherapy, and to stratify the risk of patients with poor prognosis. The results of these studies showed that the PET/CT-based texture characteristics or radiomics had good classification or predictive prognostic performance. Radiomics based on PET/CT may provide complementary information for predicting survival in patients with lung cancer (25–29). A nomogram is based on multivariate regression analysis and includes important influencing factors related to tumor prognosis. By constructing an intuitive graph using a statistical predictive model, the nomogram provides the numerical probability of a clinical event. The nomogram has become the focus of interest in cancer research in recent years and is considered a useful tool for quantifying risk (30–32).

Therefore, this study primarily aimed to construct a predictive model of the OS nomogram based on the radiomic features of PET/CT combined with the clinicopathological factors to predict prognosis and risk stratification as well as to determine the role of radiomic features in predicting the prognosis of

Abbreviations: C-index, Harrell's concordance index; CEA, carcinoembryonic antigen; HR, hazard ratio; MTV, metabolic tumor volume; NSCLC, non-small cell lung cancer; NOS, not otherwise specified; PET/CT, positron emission tomography/computed tomography; rad-score, radiomic score; SUVmax, maximal standard uptake value; SUVmean, mean standard uptake value; TLG, total lesion glycolysis; TNM, tumor-node-metastasis.

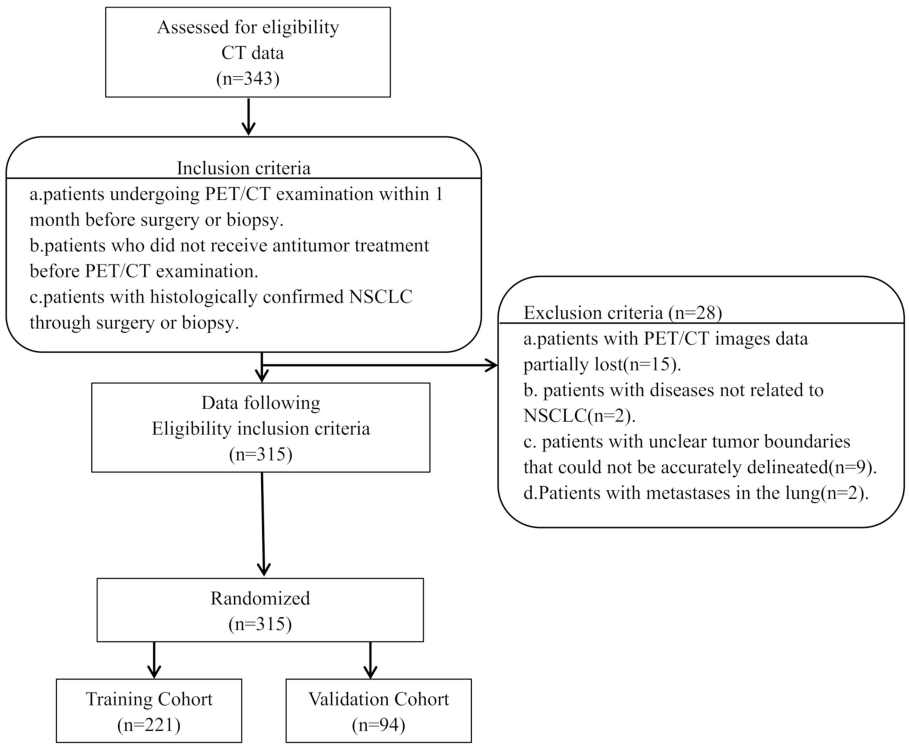


FIGURE 1 | Flowchart of patient enrollment, eligibility, and exclusion criteria of the dataset.

NSCLC. To improve the prognostic assessment of patients with NSCLC, advancements in the areas of individualized treatment and precision medicine are necessary.

PATIENTS AND METHODS

Patients and Clinicopathological Data

The institutional review board of Jinling Hospital, Medical School of Nanjing University approved this retrospective study and waived the need to obtain informed consent from the patients. This was a retrospective study, and the medical records of patients between October 2007 to August 2016 were reviewed. The medical records were searched consecutively, and 343 patients who had a lung tumor as assessed by histopathological analysis were identified. Patients with the following characteristics were included in the study: (a) patients undergoing PET/CT examination within 1 month before surgery or biopsy, (b) patients who did not receive antitumor treatment before PET/CT examination, and (c) patients with histologically confirmed NSCLC through surgery or biopsy. However, patients with the following characteristics were excluded: (a) patients with partial loss of PET or CT images ($n = 15$); (b) patients with diseases not related to NSCLC ($n = 2$), (c) patients with unclear tumor boundaries that could not be accurately delineated ($n = 9$), and (d) patients with metastases in the lung ($n = 2$). The final cohort included 315 patients (**Figure 1**). We randomly divided the patients into the training cohort ($n = 221$) and the validation cohort ($n = 94$) with a 7:3 ratio. Clinicopathological data were obtained from the patients' medical records, which included age, sex, family history, smoking history, histological grade, lymph node metastasis, distant metastasis, and TNM stage (defined according to the eighth edition of the TNM classification and staging system by the American Joint Committee on Cancer), histologic type (adenocarcinoma, squamous cell carcinoma, or not otherwise specified [nos]), treatment methods (surgery, chemotherapy, targeted therapy, and radiotherapy), thyroid transcription factor-1 (TTF-1) level, carcinoembryonic antigen (CEA) level, tumor location, and PET/CT metabolic parameters were obtained (**Table 1**). The survival information of these patients was obtained through telephone calls. Follow-up data were collected from October 2007 to January 2019. The mean and median follow-up periods were 37.99 (95% confidence interval [CI], 35.464–40.522) and 36.00 (range, 20.00–52.00) months, respectively. The endpoint of this study was OS, which was defined as the period from the date of ^{18}F-FDG PET/CT examination to the date of telephone follow-up or the date of the patient's death.

PET/CT Image Acquisition and Analysis

Patients underwent PET/CT imaging (Biography 16, Siemens, Erlangen, Germany) using ^{18}F-FDG synthesized by the Canadian EBCO TR19 medical cyclotron and chemical synthesis system. All PET/CT acquisitions were carried out in free breathing mode, and no steps were taken to correct for motion. The

TABLE 1 | Clinicopathological factors of patients in the training and validation cohorts.

Characteristic	Training cohort *N* = 221	Validation cohort *N* = 94	*P*-value
Gender - no. (%)			1.000
Female	101 (46.0)	43 (46.0)	
Male	120 (54.0)	51 (54.0)	
Age-yr	62 (54.00-69.00)	64 (55.25-70.00)	0.672
Family history- no. (%)			0.487
No	211 (95.0)	92 (98.0)	
Yes	10 (5.0)	2 (2.0)	
T stage–no. (%)			0.879
T1	74 (33.5)	28 (29.8)	
T2	74 (33.5)	34 (36.2)	
T3	24 (10.9)	9 (9.6)	
T4	49 (22.2)	23 (24.5)	
N stage- no. (%)			0.630
N0	85 (38.5)	37 (39.4)	
N1	33 (14.9)	9 (9.6)	
N2	55 (24.9)	25 (26.6)	
N3	48 (21.7)	23 (24.5)	
M stage- no. (%)			0.439
M0	118 (53.4)	45 (47.9)	
M1	103 (46.6)	49 (52.1)	
Histologic type- no. (%)			0.445
Adenocarcinoma	210 (95.0)	92 (97.9)	
Squamous cell carcinoma	9 (4.1)	2 (2.1)	
NOS	2 (0.9)	0 (0.0)	
Surgery- no. (%)			0.054
No	133 (60.2)	68 (72.3)	
Yes	88 (39.8)	26 (27.7)	
Chemotherapy- no. (%)			1.000
No	114 (51.6)	45 (47.9)	
Yes	107 (48.4)	92 (97.9)	
Targeted therapy- no. (%)			0.584
No	168 (76.0)	68 (72.3)	
Yes	53 (24.0)	26 (27.7)	
Radiotherapy- no. (%)			0.352
No	202 (91.4)	82 (87.2)	
Yes	19 (8.6)	12 (12.8)	
Smoking status- no. (%)			0.955
No	141 (64.0)	61 (65.0)	
Yes	80 (36.0)	33 (35.0)	
Histologic grade- no. (%)			0.982
Poorly differentiated	81 (37.0)	34 (36.0)	
Moderately differentiated	102 (46.0)	43 (46.0)	
Highly differentiated	38 (17.0)	17 (18.0)	
Lymph node metastasis- no. (%)			0.663
No	85 (38.0)	33 (35.0)	
Yes	136 (62.0)	61 (65.0)	
Distant metastasis- no. (%)			0.380
No	110 (50.0)	41 (44.0)	
Yes	111 (50.0)	53 (56.0)	
Stage- no. (%)			0.847
I	52 (23.5)	22 (23.4)	
II	18 (8.1)	7 (7.4)	
III	36 (16.3)	12 (12.8)	
IV	115 (52.0)	53 (56.4)	
TTF-1- no. (%)			0.239
Negative	107 (48.0)	38 (40.0)	
Positive	114 (52.0)	56 (60.0)	
CEA	4.55 (2.30-17.50)	7.55 (3.33-37.55)	0.023
SUVmax	7.32 (4.85-10.04)	6.80 (4.07-9.67)	0.358
SUVmean	4.42 (2.96-6.48)	4.07 (2.51-6.12)	0.337
TLG(g)	31.12 (15.67-83.55)	32.05 (16.64-67.81)	0.812
MTV(cm^3)	8.17 (4.93-16.33)	9.01 (5.14-18.76)	0.733

CEA, carcinoembryonic antigen; MTV, metabolic tumor volume; NOS, not otherwise specified; SUVmax, maximal standard uptake value; SUVmean, mean standard uptake value; TTF-1, thyroid transcription factor-1; TLG, total lesion glycolysis.

radiochemical purity was >95%. All acquisitions were carried out in a free-breathing mode. The patients fasted for 6–8 h before undergoing the scan. Patients were intravenously injected with ^{18}F-FDG (5.55 MBq/kg) and underwent a whole-body PET/CT scan of the skull base to the upper part of the thigh, and the data included CT and PET scans. The CT scanning parameters were as follows: power, 120 kV; current, 140 mAs; slice thickness and spacing, 5 mm; matrix, 512×512; and tube rotation speed, 0.8 s/r. The PET acquisition parameters were as follows: three-dimensional at 3 min/bed; iterative algorithm; iterations, 4; subset, 8; resolution, 4.1 mm lateral, 4.6 mm axial; matrix, 128 × 128; voxel size, 5.3 × 5.3 × 5.3 mm^3. These settings were the same for all included patients. Images were reconstructed using an iterative reconstruction method resulting in CT, PET, and PET/CT fusion images that were transferred to a post-processing workstation. We used Microsoft Viewer software (version VB10, Siemens) to calculate the metabolic parameters on the PET images. PET images were first converted to SUV images in the software without other processing methods. Then, the 3-dimensional region of interest (ROI) was manually delineated by a radiologist (W.Q.G.) to calculate the maximum standard uptake value (SUVmax, with a threshold set to 40%), mean standard uptake value (SUVmean), and metabolic tumor volume (MTV). Subsequently, the total lesion glycolysis (TLG) (TLG=SUVmean×MTV) was calculated.

Tumor Segmentation

Our study followed and adhered to the Image Biomarker Standardization Initiative (IBSI) guidelines (33), and the software (Radiomics, Frontier, Siemens) used was IBSI-compliant. A volume of interest(VOI)was drawn semiautomatically around the tumor by a chest radiologist (Y.B., 9 years of experience) in the lung diagnosis using the radiomics prototype (Radiomics, Frontier, Siemens) and confirmed by another chest radiologist

(W.Q.G., 5 years of experience). Both radiologists were blinded to the patients' clinical information. Firstly, we import CT images into radiomics prototype software (Radiomics, Frontier, Siemens). In the segmentation module, a few segmentation tools were available for semiautomatic delineation of the tumor in three dimensions. The segmentation was semiautomatically produced by drawing a line across the boundary of the tumor, then, the tool automatically find the neighboring voxels in 3D space with the same gray level through an automatic algorithm, and this is a Random Walker-based lesion segmentation for solid and subsolid lung lesions (34). The first step is to obtain a superset of voxels that may be part of the lesion. This can be implemented efficiently as a 3 D region growing starting from the center of the ROI. Then the thresholds can be fixed for lesions or determined adaptively from an analysis of the density distribution in the ROI. The region growing results in complete lesion and additionally parts of the attached vasculature. A morphological opening operation is applied to remove the vessels finally (35). If the segmentation wasn't right, the operators could correct it manually in the 3D domain using the radiomics prototype. The algorithm aimed at K-way image segmentation with given seeds indicating regions of the image belonging to the K objects(the objects to be segmented). Each seed specifies a location with a user-defined label. The algorithm labels an unseeded pixel by resolving the question: Given a random walker starting at this location, what is the probability that it first reaches each of the K seed points? It will be shown that this calculation may be performed exactly without the simulation of a random walk. By performing this calculation, we assign a K-tuple vector to each pixel that specifies the probability that a random walker starting from each un-seeded pixel will first reach each of the K seed points. A final segmentation may be derived from these K-tuples by selecting for each pixel the most probable seed destination for a random walker. By biasing the random walker to avoid crossing sharp intensity gradients, a quality segmentation is obtained that respects object boundaries (including weak boundaries) (36). And then in the radiomics module to click the computer features tool to calculate the CT radiomic features, and export the CT Masks+STL at the same time. Then we import the PET image into the software. If the tumor on the PET image is not at the same slice as the CT, we manually adjust the slice of PET image. Then, the CT Masks+STL will be imported into the software to cover the tumor on the PET image. If the CT Masks+STL does not cover the tumor, two radiologists(Y.B; W.QG) manually adjusted the CT Masks+STL through edit tools and reached a consensus to ensure that the CT Masks+STL completely covered the tumor lesions on the PET image as much as possible, and then use the same method to extract PET radiomic features. So, the 3D ROI (VOI) was delineated on CT image, and could be used by the PET image when the PET image were transformed to the CT image space using the transformation matrix obtained in PET-CT fusion.

Radiomic Feature Extraction

The Radiomic features from volumes of interest were then computed with both CT and PET images on a prototype that interfaces with the PyRadiomics library in manner similar to the 3D slicer's Radiomics plugin (34). The PyRadiomics library provides a variety of options to customize image pre-processing before feature extraction. Laplacians of Gaussian filtering, wavelet filtering, and non-linear intensity transforms were selected for image pre-processing. The feature classes contain 162 first-order features, 12 shape features, and 666 texture features. We also extracted numerous features (e.g., wavelets) that have not yet been standardized or validated by the IBSI. As a result, a total of 840 radiomic features were extracted from the CT and PET images using the software (**Figure 2**). The IBSI guidelines for reporting all necessary details are provided in the **Supplemental Material**.

Feature Selection and Radiomics Signature Construction

Considering the redundancy of the features and to reduce model overfitting, feature engineering was performed using two methods, Spearman correlation test and the least absolute shrinkage and selection operator (LASSO) Cox regression analysis. The Spearman correlation test was initially used to reduce feature redundancy, and a cutoff value of 0.9 was adopted. A ten-fold cross-validation LASSO Cox regression method, which is suitable for the regression of high-dimensional data in survival analysis, was conducted to select the most useful predictive features from the training cohort. The specified step of the LASSO Cox analysis included determining the optimized hyperparameter λ, which ensured that the model had the least deviance. Features with non-zero coefficients were preserved. The rad-score was calculated via a linear combination of selected features weighted by their respective coefficients. Three rad-scores including the CT, PET, and PET/CT rad-scores for each patient were calculated using PET, CT, and PET/CT features, respectively (**Figure 2**).

Clinicopathological Factor Analysis

Clinicopathological factors including PET/CT metabolic parameters were analyzed using a univariate Cox proportional hazards regression analysis. Factors with $p < 0.05$ were analyzed using the Kaplan–Meier curve and log-rank test. These significant factors were combined into a multivariate Cox proportional hazards regression analysis to identify independent risk factors.

Construction and Validation of OS Nomogram

Before constructing the OS nomogram, the performance of each rad-score was evaluated using the concordance index (C-index). The largest rad-score integrated with the independent factors was used to construct the nomogram. The prognostic ability of the nomogram was evaluated in the training cohort and validated in the validation cohort. The discrimination performance of the nomogram was assessed using Harrell's C-index. The C-index ranges between 0.5 and 1.0, with 0.5 indicating a random distribution of data and 1.0 indicating the outcome of the model perfectly predicting the observed survival information. The calibration curves of the nomogram were subsequently drawn for the patients' 5-year OS. The calibration curves were used

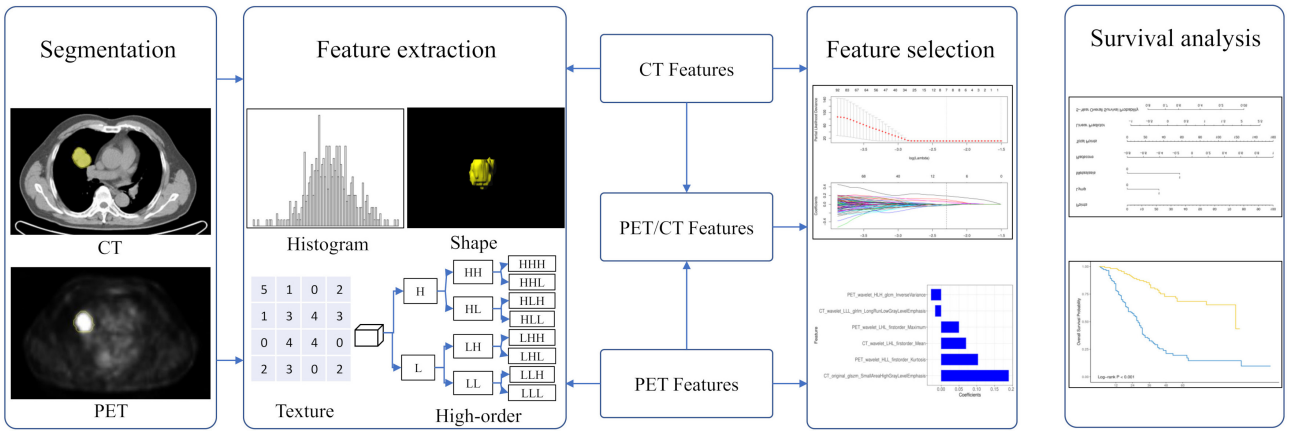


FIGURE 2 | The workflow in developing a radiomic overall survival nomogram. Computed tomography (CT) and positron emission tomography (PET) images were segmented semiautomatically using the Siemens radiomics prototype. Features including histogram features, shape features, texture features, and wavelet features were extracted from CT and PET images using the software. Three rad-scores were calculated using the least absolute shrinkage and selection operator Cox regression based on subsets of CT, PET, and PET/CT radiomic features. The predictive ability of CT, PET, and PET/CT rad-scores on overall patient survival was evaluated. The overall survival nomogram was constructed based on the PET/CT rad-score and clinicopathological factors.

to determine the independent risk factors and also illustrated both survival probabilities predicted by the nomogram and the observed probabilities.

Statistical Analysis

R software (version 3.5.0, www.Rproject.org) was used for statistical analysis in this study. LASSO was conducted using the "glmnet" package, while the "hdnom" package was used for the survival analysis. All statistical tests were two-sided, with a significance level of 0.05. Finally, a decision curve analysis was conducted using the "rmda" package to determine the clinical usefulness of radiomics nomogram by quantifying the net benefits at different threshold probabilities (37).

RESULTS

Clinical Characteristics of Patients

The study patients were divided into two groups: the training cohort with 221 patients (120 men and 101 women) and a validation cohort with 94 patients (51 men and 43 women). There were no significant differences in sex, family history, smoking status, histological grade, lymph node metastasis, distant metastasis, TNM stage, and TTF-1 level ($p = 0.054$–1.000) between the training and validation cohorts. CEA levels were significantly different between the training and validation cohorts ($p = 0.023$). Other clinicopathological characteristics are shown in **Table 1**.

Establishment of Multivariate Cox Proportional Hazards Model

Before constructing the final model, we used a multivariate Cox regression analysis to test the hazard ratio (HR) of each parameter and to determine its significance in the probability of death. The results were as follows: distant metastasis (HR, 1.94 [95% CI, 1.17–3.21]), (HR, 1.71 [95% CI, 0.81–3.61]); stage (HR, 3.24 [95% CI, 1.74–6.02]), (HR, 8.34 [95% CI, 2.28–30.56]); CEA (HR, 1.12 [95% CI, 1.03–1.21]), (HR, 1.18 [95% CI, 1.01–1.37]) and targeted therapy (HR, 0.35 [95% CI, 0.22–0.56]), (HR, 0.41 [95% CI, 0.19–0.89]) were the independent risk factors in the training and validation cohorts, respectively (**Table 2**).

TABLE 2 | HR analysis for the different independent clinicopathological factors for clinical model.

	Training cohort				Validation cohort			
	HR	*p* value	95% CI for HR		HR	*p* value	95% CI for HR	
			Lower	Upper			Lower	Upper
Distant metastasis	1.94	0.010	1.17	3.21	1.71	0.162	0.81	3.61
Stage	3.24	<0.001	1.74	6.02	8.34	0.001	2.28	30.56
CEA	1.12	0.007	1.03	1.21	1.18	0.035	1.01	1.37
Targeted therapy	0.35	<0.001	0.22	0.56	0.41	0.023	0.19	0.89

CEA, carcinoembryonic antigen; HR, hazard ratio.

Important Radiomic Features Selection and Calculation of the Rad-Score: Model Construction and Comparison

We performed a selection using the LASSO regression model on the PET/CT features, as shown in **Figures 3A,B**. To calculate the rad-score, the following six important features were selected from the 840 radiomic features, as shown in **Figure 3C**: PET_wavelet_HLH_glcm_Inverse Variance, CT_wavelet_LLL_glrlm_Long Run Low Gray Level Emphasis,PET_wavelet_LHL_firstorder_Maximum,CT_wavelet_LHL_firstorder_Mean, PET_wavelet_HLL_firstorder_Kurtosis, and CT_original_glszm_Small Area High Gray Level Emphasis. Subsequently, the rad-scores were calculated. The PET/CT

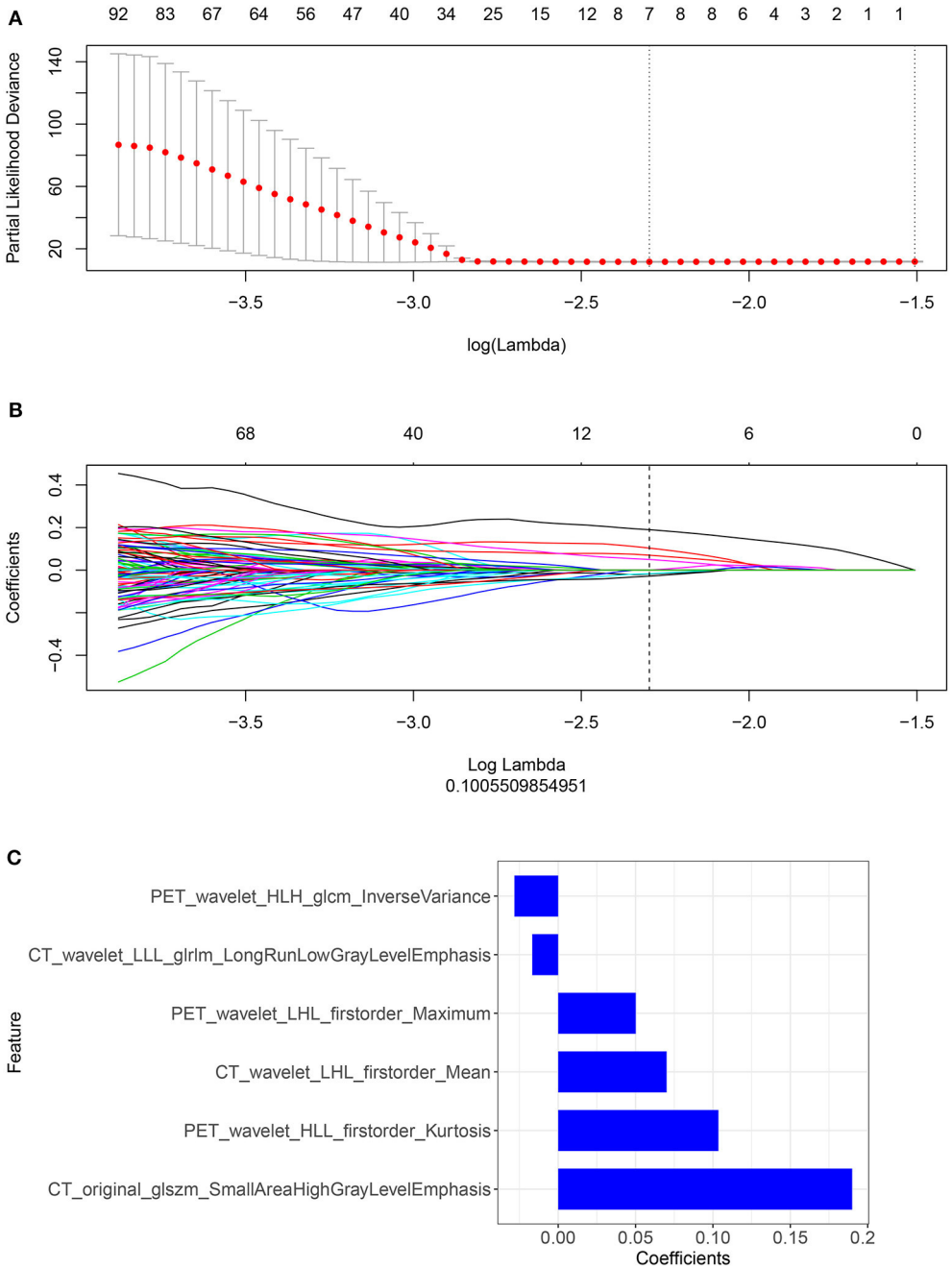


FIGURE 3 | (A,B) Radiomic features were selected using the ten-fold cross-validation least absolute shrinkage and selection operator Cox regression model in the training cohort (number of patients: 221). The following two steps were included: determining the hyperparameter/lambda with a partial likelihood deviance as the criterion (top row) and using the optimized/lambda (the vertical dashed line) to select features with nonzero coefficients (bottom row). **(C)** A total of six important radiomic features were selected.

rad-score was determined using the following formula: Rad-score=0.19×CT_original_glszm_SmallAreaHighGrayLevel Emphasis+0.07×CT_wavelet_LHL_firstorder_Mean+0.017× CT_wavelet_LLL_glrlm_LongRunLowGrayLevelEmphasis+ 0.028×PET_wavelet_HLH_glcm_InverseVariance+0.104×PET_ wavelet_HLL_firstorder_Kurtosis+0.05×PET_wavelet_LHL_first order_Maximum−0.019.

We constructed three rad-scores based on CT features, PET features, and PET/CT combined features. The C-index of the rad-scores is shown in **Table 3**. Among these three rad-scores, the CT rad-scores were 0.685 and 0.658 in the training and validation cohorts, respectively. The PET rad-score had a lower C-index (0.662 and 0.611 for the training and validation cohorts, respectively) than the CT rad-score. The PET/CT rad-score had

TABLE 3 | Harrell's concordance index of different modalities.

Modality	Training cohort	Validation cohort
	C-index 95% CI	C-index 95% CI
CT	0.685 (0.654–0.716)	0.658 (0.593–0.723)
PET	0.662 (0.609–0.715)	0.611 (0.540–0.682)
PET/CT	0.706 (0.663–0.749)	0.661 (0.540–0.682)
Clinical model	0.730 (0.691–0.769)	0.774 (0.707–0.841)
TNM stage	0.552 (0.504–0.072)	0.531 (0.435–0.144)
Tumor volume	0.607 (0.554–0.080)	0.644 (0.560–0.127)
TNM stage and tumor volume	0.618 (0.568–0.076)	0.635 (0.548–0.132)
Radiomics nomogram (PET/CT combined with Clinical model)	0.776 (0.741–0.811)	0.789 (0.724–0.854)

PET/CT, positron emission tomography/computed tomography.

the best C-index (0.706 and 0.661 for the training and validation cohorts, respectively). The C-index values of the clinical model with clinicopathological factors were 0.730 and 0.774 in the training and validation cohorts, respectively (**Table 3**). The C-index values of the TNM stage and tumor volume (0.618 and 0.635 for the training and validation cohorts, respectively) were significantly higher than that of the TNM stage (0.552 and 0.531, respectively) or tumor volume (0.607 and 0.644, respectively) alone (**Table 3**). A rad-score was combined with the clinicopathological factors to construct a nomogram based on LASSO, as shown in **Figure 4A**. The C-index (0.776 and 0.789 for the training and validation cohorts, respectively) of the PET/CT rad-score combined with the clinical model was higher than that of the clinical model without the rad-score (**Table 3**). The validation of the nomogram showed that it had good predictive performance, as shown in **Figures 4B,C**. The calibration curve showed that the predicted probability was significantly close to the actual survival time of patients, as shown in **Figures 4D,E**. We also analyzed the association of PET/CT rad-score, OS nomogram, tumor volume, stage, and clinical model with the survival time of patients with NSCLC using a Kaplan-Meier analysis. **Figures 5A–E** shows the survival probability of the patients in the high-risk or low-risk cohorts. The results of the log-rank test indicate significant discrimination between the two groups.

To determine the clinical usefulness of the radiomics nomogram model, a decision curve analysis was performed. The decision curve analysis showed that the radiomics nomogram had a higher overall net benefit than 4 other clinical models (tumor volume, TNM stage, TNM stage, and tumor volume, and clinical model) across the majority of the range of reasonable threshold probabilities as shown in **Figure 6**.

DISCUSSION

In this study, we used ^{18}F-FDG PET/CT radiomics to investigate the prognosis of patients with NSCLC. We extracted radiomic features from CT and PET, constructed a radiomics signature, and calculated the rad-score. Subsequently, we compared the predictive performance of CT, PET, and PET/CT rad-scores to determine the prognosis of patients with NSCLC. Considering that PET/CT has the best predictive performance among the three modalities, we further combined the PET/CT rad-score with the clinicopathological factors to predict the prognosis of patients with NSCLC. In addition, we performed Cox proportional hazards regression analysis on the clinicopathological risk factors and selected the independent risk factors related to the patient's prognosis. Finally, we constructed two prediction models based on LASSO: clinical models with and without the rad-score. In addition, 3 other clinical models were established (TNM stage, tumor volume, and tumor volume) to predict the prognosis of patients with NSCLC. Our results showed that the OS nomogram had good predictive performance for prognosis and could successfully stratify patients into high-risk and low-risk groups.

We extracted 840 radiomic features from CT, PET, and PET/CT images. To avoid redundancy and overfitting caused by the small sample size and additional radiomic features, we used the LASSO method to select important radiomic features. LASSO can be used to select biomarkers from high-dimensional radiomic features to overcome the problem of a small sample size and to select features that are most relevant to survival time (38). In addition, the LASSO method with cross-validation, as presented in this study, can be used to elegantly address issues of overfitting, collinearity, and multiple-hypothesis testing in feature selection. The LASSO method was also used to select radiomic features related to prognosis that were consistent with previous reports (39, 40). Furthermore, our prediction performance after using the LASSO method was better than the prediction performance of previous studies (38, 41, 42). Finally, we selected a total of six important radiomic features to construct CT, PET, and PET/CT radiomics signatures. The rad-scores were subsequently calculated for the three modalities to compare their predictive performances, revealing that PET/CT had the best predictive performance. Hence, we further studied the PET/CT modality, and combined the PET/CT rad-score with the clinicopathological factors that acquired good predictive performance with the C-index (0.776 and 0.789 for the training and validation cohorts, respectively). Our results showed that TNM staging was inconsistent with prognostic assessment; therefore, the prognosis cannot be predicted well. Radiomics can be used to comprehensively and quantitatively assess the spatiotemporal heterogeneity of tumors, and when combined with clinicopathological factors, the predictive performance of prognosis may be improved. According to Kirienko et al. (43), the Cox proportional hazards regression model was established based on CT, PET, and PET/CT radiomic signatures to predict the disease-free survival of patients with NSCLC. The results showed that the Cox proportional hazards regression models including radiomic features for the CT, PET, and PET/CT images had areas under the curve (AUCs) of 0.75, 0.68, and 0.68, respectively. The addition of clinicopathological risk factors to the Cox proportional hazards regression models resulted in AUCs of 0.61, 0.64, and 0.65 for the CT, PET, and PET/CT images, respectively. Mattonen et al. (44) constructed a Cox proportional hazards model that included stage and an MTV plus penumbra texture

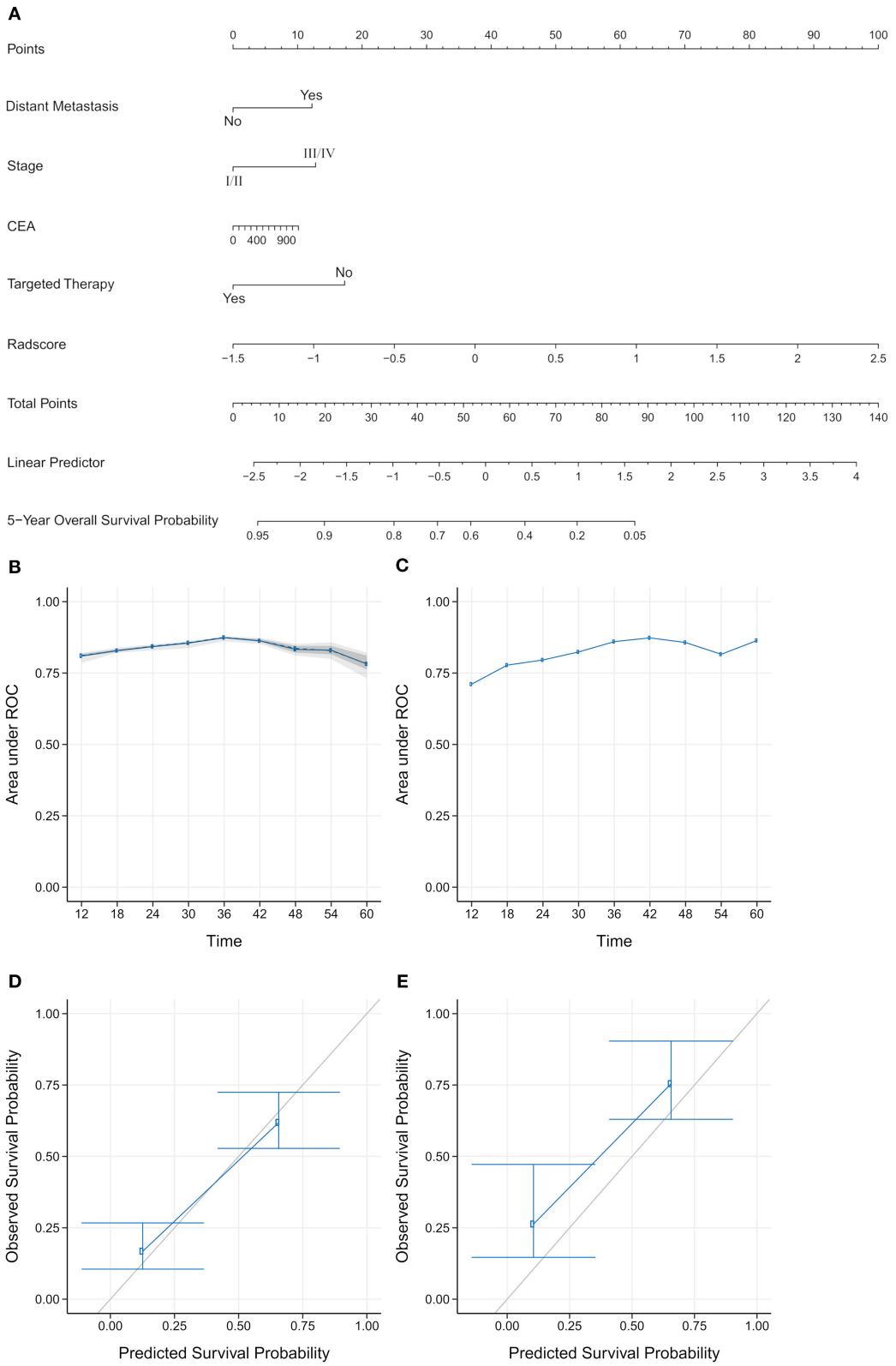


FIGURE 4 | (A) Establishment of a comprehensive nomogram by combining the positron emission tomography/computed tomography (PET/CT) rad-score and clinicopathological factors for predicting the 5-year overall survival of patients with non-small cell lung cancer. **(B,C)** A validation analysis of the nomogram showed

FIGURE 4 | that the area under the curve (AUC) at five time points was obtained on the training and validation cohorts. The AUC for predicting the prognosis from 1–5 years was >0.7. With the extension of follow-up time, the predicted AUC gradually increased, indicating that the nomogram has a good performance in predicting prognosis. **(D,E)** The calibration curve was used to estimate the 5-year overall survival predicted using a nomogram. The diagonal gray line represents an ideal evaluation, and the blue line represents the performance of the nomogram. The calibration curves for the training and validation cohorts showed the calibration of two cohorts in terms of the agreement between the estimated and observed 5-year outcomes.

feature to predict recurrence/progression in NSCLC based on the LASSO method. The results showed that the C-index of the training and validation sets of this multivariate model were both 0.74. Wang et al. (45) used the consensus clustering method to automatically select the stable and prognostic radiomic features and subsequently constructed a multivariate Cox proportional hazards model that incorporated CT radiomic, clinical, and hematological features. These were found to be more predictive with a C-index of 0.792 and retained a C-index of 0.743 in the cross-validation analysis, therefore outperforming the radiomic, clinical, or hematological models. In addition, Dissaux et al. (27) and Oikonomou et al. (28) used PET/CT-based radiomics to predict the prognosis of patients with lung cancer who were treated with stereotactic body radiotherapy (SBRT). The results showed that radiomic features derived from PET/CT were associated with local control in patients with NSCLC undergoing SBRT, and can be used as predictors of OS, disease-specific survival, and regional control. Radiomics on PET/CT provided complementary information for the prediction of control and survival in patients with SBRT-treated lung cancer and could be helpful in clinical decision-making. The above studies showed that radiomic features were related to prognosis and had good prognostic predictive performance. Although some of the studies were multi-center studies, the sample sizes were generally small; therefore, the prediction model may have been overfitted. Our prediction model has better predictive performance and our sample size was much larger compared to previous studies. Our results indicated that the rad-score from the CT, PET, or PET/CT group has a favorable predictive power for survival. Moreover, the PET/CT rad-score had the best performance among the three rad-scores and could improve the predictive performance of the PET/CT models when combined with the clinicopathological factors. Hence, we believe that more tumor details are contained in the PET/CT entity model compared to an individual CT or PET entity model, a finding consistent with the findings of previous studies.

Additionally, we performed univariate Cox regression analysis on the clinicopathological factors to test the HR of each parameter and to determine its significance in the probability of death. The results showed that distant metastasis, stage, CEA, and targeted therapy were independent risk factors; We subsequently constructed a nomogram by combining the PET/CT rad-score and clinicopathological factors. The calibration curve showed that the predicted probability was significantly close to the actual survival time of patients. The validation of the OS nomogram showed that with the extension of follow-up time, the AUC for predicting prognosis gradually increased, and our results indicated that the OS nomogram had good predictive performance. We also evaluated the reliability of the PET/CT rad-score, OS nomogram, tumor volume, stage, and clinical model in predicting patient survival using a Kaplan–Meier analysis. The results of the Kaplan–Meier analysis demonstrated that the OS nomogram can clearly divide the patients into high-risk and low-risk groups, indicating that our nomogram had a strong predictive power in patients with high and low risks. Thus, it is considered significantly robust and reliable, and can be used as evidence for additional treatment and close follow-up in patients with poor prognosis, which is consistent with the research results of Dessroit et al. (29). In addition, the decision curve analysis demonstrated that the radiomics nomogram was superior to 4 other clinical models (tumor volume, TNM stage, TNM stage and tumor volume, and clinical model) across the majority of the range of reasonable threshold probabilities, which indicated that the radiomics nomogram added incremental value to the traditional staging system and other clinicalpathologic factors for individualized estimations. We believe that with the combination of the rad-score and clinicopathological factors to construct an OS nomogram, the predictive performance was largely improved, suggesting that the rad-score played an important role in the predictive accuracy of the OS of patients with NSCLC, a result that was consistent with the results of previous studies (46, 47). It is worth nothing that because of variations in technical parameters or inconsistent imaging parameters, a limited sample size, and heterogeneous patient characteristics, radiomic features may be insignificant in predicting prognosis in certain situations. Therefore, cohorts and validation datasets need to be evaluated, methodologies need to be standardized, and data on studies that evaluate radiomic features need to be harmonized in future studies, especially those with retrospective multi-centric datasets (48, 49).

Our study has some limitations, including the relatively small sample size and single-center cohort, the retrospective nature of the data, and the lack of external validation, which may have introduced selection bias, thereby resulting in poor model generalization and capacity. However, we plan to rapidly expand the sample size, and multi-center cohorts should be recruited for validation in the near future. Secondly, in this study, all texture matrices using 26-connectivity to find the neighboring voxels with distance 1 and 13 angles. Finally, the value of a feature is calculated for each angle separately, after which the mean of these values is used. However, it has been shown this strategy leads to less informative features compared to extracting the feature from as single matrix implementing all 13 directions, so, strategy should be implemented to merge the angle specific features into the texture matrix in future studies. Thirdly, We measured four metabolic parameters in a different way (different volume, different software) than the other radiomic features. It may lead to a bias in the comparison of the potential value of these four metrics with respect to

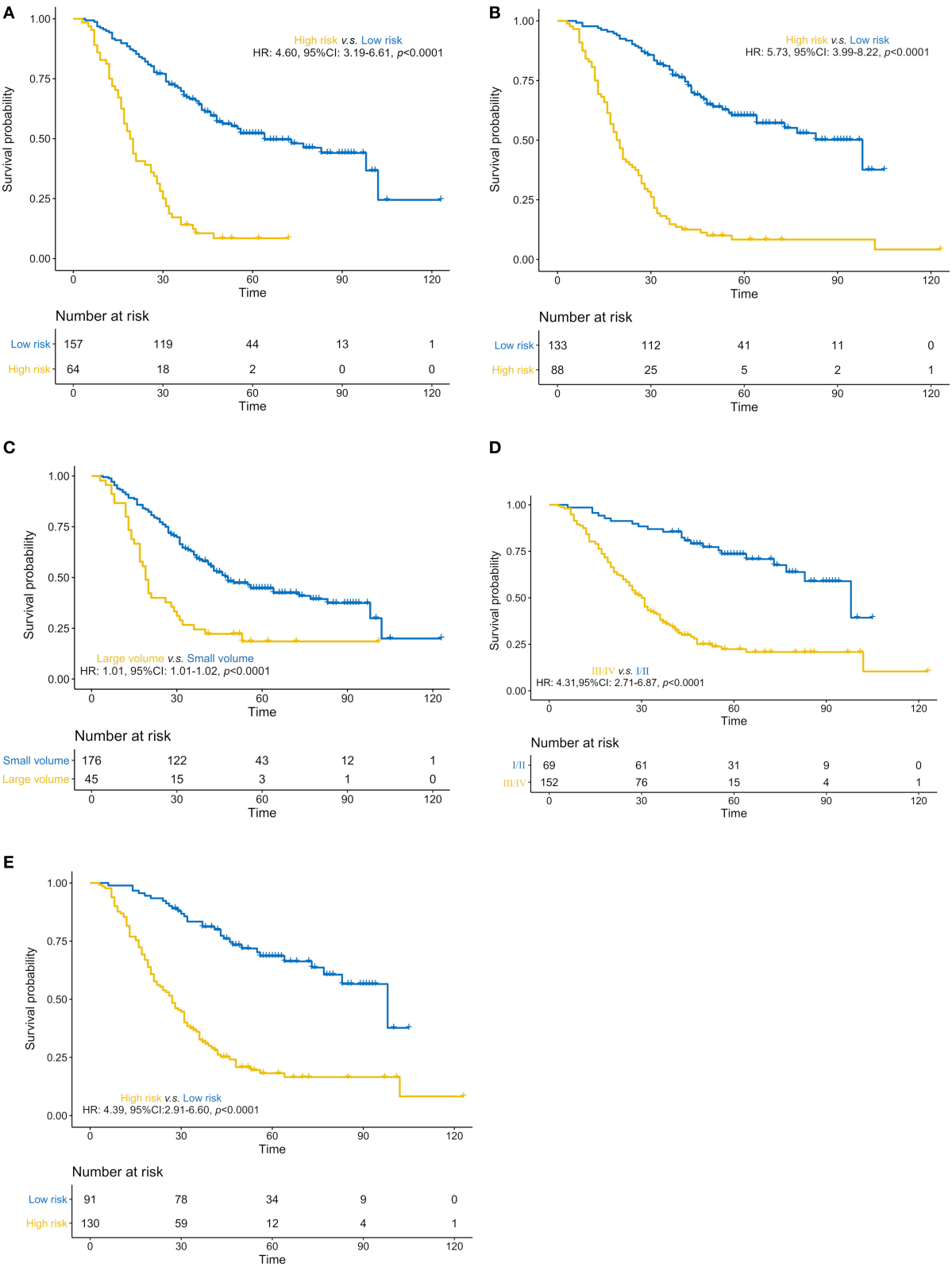


FIGURE 5 | (A–E). Predictive performance of the PET/CT rad-score, overall survival nomogram, tumor volume, stage, and clinical model Kaplan-Meier survival analysis of the patients in the high- and low-risk groups in the training cohort. Kaplan-Meier analysis for the PET/CT rad-score **(A)**, overall survival (OS) nomogram **(B)**, tumor volume **(C)**, stage **(D)**, and clinical model **(E)**. The patients were stratified into high- and low-risk groups based on PET/CT rad-score (**A**, $p < 0.0001$, log-rank test), OS nomogram (**B**, $p < 0.0001$, log-rank test), tumor volume (**C**, $p < 0.0001$, log-rank test), stage (**D**, $p < 0.0001$, log-rank test), and clinical model (**E**, $p < 0.0001$, log-rank test).

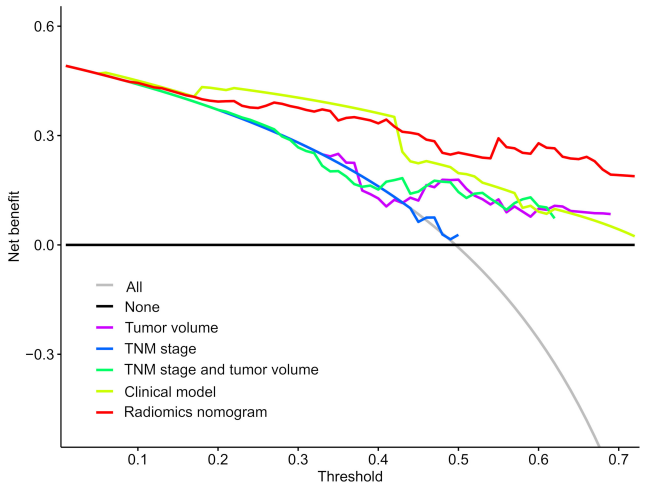


FIGURE 6 | Decision curve analysis for each model. The y-axis denotes the net benefit, which was calculated using true-positive and false-positive results. The radiomics nomogram model has the highest net benefit at the threshold from 0.1 to 0.9 among all positive predictions (line labeled "All"), all negative prediction (line labeled "None"), and another 4 clinical models (line labeled "Tumor volume, tumor-node-metastasis [TNM] stage, TNM stage and tumor volume, and clinical model").

all other features calculated by pyradiomics. And the stability and repeatability of the MTV and TLG values derived from a volume determined through a fixed threshold at 40% of maximum intensity are still controversial. So, we will try to use a consensus of several manual delineation instead of fixed thresholding to calculate tumor metabolic volume in the near future study (50, 51). In addition, all PET/CT acquisitions were carried out in free breathing mode, and no steps were taken to correct for motion may lead to extraction of features might have been suboptimal in the case of small lesions affected by motion, but also in some larger heterogeneous uptakes affected by motion blur.

In conclusion, the identified radiomic signature based on PET/CT can be potentially used as a biomarker for risk stratification of the OS in patients with NSCLC. The OS nomogram combining radiomics and clinicopathological factors for individualized OS estimation may provide more precise guidance for the accurate diagnosis and treatment of NSCLC in clinical practice.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher. Requests to access the datasets should be directed to Hong Zhu, zh_zy@163.com.

ETHICS STATEMENT

The Institutional Review Board of Jinling Hospital, Medical School of Nanjing University approved this retrospective study and waived the need to obtain informed consent from the patients.

AUTHOR CONTRIBUTIONS

BY conceived the idea of the study. BY, JiaZ, JinZ, LM, AL, HJ, CZhou, and QW collected the data. HZ and GL performed image analysis. BY wrote the manuscript. SD performed the statistical analysis. SD, LZ, CZhu, JT, FW, and GL edited and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2017YFC0113400 for LZ) and Natural Science Foundation of Jiangsu Province (BK2011665).

ACKNOWLEDGMENTS

We would like to thank all the personnel from the Department of Nuclear Medicine and Department of Medical Imaging, Jinling Hospital, Medical School of Nanjing University for their collective efforts on the image post-processing, clinical management, and follow-up. Additionally, we would like to thank Dr. Yingqian Ge and Dr. Chengyan Wang for their guidance in our project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01042/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *Cancer J Clin.* (2018) 68:7–30. doi: 10.3322/caac.21442
2. Addario BJ. Lung cancer is a global epidemic and requires a global effort. *Ann Transl Med.* (2015) 3:26. doi: 10.3978/j.issn.2305-5839.2015.01.20
3. Curran WJ, Paulus R, Langer CJ, Komaki R, Lee JS, Hauser S, et al. Sequential vs. concurrent chemoradiation for stage III non-small cell lung cancer: randomized phase III trial RTOG 9410. *J Natl Cancer Inst.* (2011) 103:1452–60. doi: 10.1093/jnci/djr325
4. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc.* (2008) 83:584–94. doi: 10.4065/83.5.584
5. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (Eighth) Edition of the TNM classification for lung cancer. *J Thorac Oncol.* (2016) 11:39–51. doi: 10.1016/j.jtho.2015.09.009
6. Chansky K, Sculier JP, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P, et al. The International Association for the Study of Lung Cancer Staging Project: prognostic factors and pathologic TNM stage in surgically

managed non-small cell lung cancer. *J Thorac Oncol.* (2009) 4:792–801. doi: 10.1097/JTO.0b013e3181a7716e
7. Kawaguchi T, Takada M, Kubo A, Matsumura A, Fukai S, Tamura A, et al. Performance status and smoking status are independent favorable prognostic factors for survival in non-small cell lung cancer: a comprehensive analysis of 26, 957 patients with NSCLC. *J Thorac Oncol.* (2010) 5:620–30. doi: 10.1097/JTO.0b013e3181d2dcd9
8. Higashi K, Ueda Y, Arisaka Y, Sakuma T, Nambu Y, Oguchi M, et al. 18F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resected non-small cell lung cancer. *J Nucl Med.* (2002) 43:39–45. doi: 10.1587/transcom.E96.B.802
9. William WN, Lin HY, Lee JJ, Lippman SM, Roth JA, Kim ES. Revisiting stage IIIB and IV non-small cell lung cancer: analysis of the surveillance, epidemiology, and end results data. *Chest.* (2009) 136:701–9. doi: 10.1378/chest.08-2968
10. Halabi S, Lin CY, Kelly WK, Fizazi KS, Moul JW, Kaplan EB, et al. Updated prognostic model for predicting overall survival in first-line chemotherapy for patients with metastatic castration-resistant prostate cancer. *J Clin Oncol.* (2014) 32:671–7. doi: 10.1200/JCO.2013.52.3696
11. Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol.* (2013) 14:1295–306. doi: 10.1016/S1470-2045(13)70491-1
12. Tran B, Dancey JE, Kamel-Reid S, McPherson, JD, Bedard PL, et al. Cancer genomics: technology, discovery, and translation. *J Clin Oncol.* (2012) 30:647–60. doi: 10.1200/JCO.2011.39.2316
13. Hofman V, Ilie M, Long E, Lassalle S, Butori C, Bence C, et al. Immunohistochemistry and personalised medicine in lung oncology: advantages and limitations. *Bull Cancer.* (2014) 101:958–65. doi: 10.1684/bdc.2014.2041
14. Grootjans W, de Geus-Oei LF, Troost EG, Visser EP, Oyen WJ, Bussink J. PET in the management of locally advanced and metastatic NSCLC. *Nat Rev Clin Oncol.* (2015) 12:395–407. doi: 10.1038/nrclinonc.2015.75
15. Kirienko M, Gallivanone F, Sollini M, Veronesi G, Voulaz E, Antunovic L, et al. FDG PET/CT as theranostic imaging in diagnosis of non-small cell lung cancer. *Front Biosci.* (2017) 22:1713–23. doi: 10.2741/4567
16. Hatt M, Majdoub M, Vallières M, Tixier F, Le Rest CC, Groheux D, et al. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med.* (2015) 56:38-44. doi: 10.2967/jnumed.114.144055
17. Goodgame B, Pillot GA, Yang Z, Shriki J, Meyers BF, Zoole J, et al. Prognostic value of preoperative positron emission tomography in resected stage I non-small cell lung cancer. *J Thorac Oncol.* (2008) 3:130–4. doi: 10.1097/JTO.0b013e318160c122
18. Kim HR, Kim DJ, Lee WW, Jheon S, Sung SW. The significance of maximum standardized uptake values in patients with stage I pulmonary adenocarcinoma. *Eur J Cardiothorac Surg.* (2009) 35:712–6. doi: 10.1016/j.ejcts.2008.12.030
19. Hoang JK, Hoagland LF, Coleman RE, Coan AD, Herndon JE, Patz EF Jr. Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography imaging in patients with advanced-stage non-small-cell lung carcinoma. *J Clin Oncol.* (2008) 26:1459–64. doi: 10.1200/JCO.2007.14.3628
20. Inal A, Kucukoner M, Kaplan MA, Urakci Z, Karakus A, Komek H, et al. Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography in patients with advanced non-small cell lung cancer: single center experience. *J BUON.* (2012) 17:724–8. doi: 10.1158/1078-0432.CCR-12-1202
21. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
22. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* (2014) 5:4006. doi: 10.1038/ncomms5644
23. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. *Clin Cancer Res.* (2017) 23:4259–69. doi: 10.1158/1078-0432.CCR-16-2910
24. Gillies RJ, Kinahan PE, Hricack H. Radiomics: images are more than pictures, they are data. *Radiology.* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
25. Shiri I, Maleki H, Hajianfar G, Abdollahi H, Ashrafinia S, Hatt M, et al. Next-generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using multimodal imaging and machine learning algorithms. *Mol Imaging Biol.* (2020). doi: 10.1007/s11307-020-01487-8
26. Vaidya M, Creach KM, Frye J, Dehdashti F, Bradley JD, El Naqa I. Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol.* (2012) 102:239–45. doi: 10.1016/j.radonc.2011.10.014
27. Dissaux G, Visvikis D, Da-Ano R, Pradier O, Chajon E, Barillot I, et al. Pre-treatment F-FDG PET/CT Radiomics predict local recurrence in patients treated with stereotactic radiotherapy for early-stage non-small cell lung cancer: a multicentric study. *J Nucl Med.* (2019) 61:814–20. doi: 10.2967/jnumed.119.228106
28. Oikonomou A, Khalvati F, Tyrrell PN, Haider MA, Tarique U, Jimenez-Juan L, et al. Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. *Sci Rep.* (2018) 8:4003. doi: 10.1038/s41598-018-22357-y
29. Desseroit MC, Visvikis D, Tixier F, Majdoub M, Perdrisot R, Guillevin R, et al. Development of a nomogram combining clinical staging with (18)F-FDG PET/CT image features in non-small-cell lung cancer stage I-III. *Eur J Nucl Med Mol Imaging.* (2016) 43:1477–85. doi: 10.1007/s00259-016-3325-5
30. Han DS, Suh YS, Kong SH, Lee HJ, Choi Y, Aikou S, et al. Nomogram predicting long-term survival after d2 gastrectomy for gastric cancer. *J Clin Oncol.* (2012) 30:3834–40. doi: 10.1200/JCO.2012.41.8343
31. Tang XR, Li YQ, Liang SB, Jiang W, Liu F, Ge WX, et al. Development and validation of a gene expression-based signature to predict distant metastasis in locoregionally advanced nasopharyngeal carcinoma: a retrospective, multicentre, cohort study. *Lancet Oncol.* (2018) 19:382–93. doi: 10.1016/S1470-2045(18)30080-9
32. Liang W, Zhang L, Jiang G, Wang Q, Liu L, Liu D, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol.* (2015) 33:861–9. doi: 10.1200/JCO.2014.56.6661
33. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* (2020) 295:328–38. doi: 10.1148/radiol.2020191145
34. Wels MG, Félix Lades, Muehlberg A, Suehling M. General purpose radiomics for multi-modal clinical research. *Computer-Aided Diagnosis.* (2019). doi: 10.1117/12.2511856
35. Moltz J H, Bornemann L, Kuhnigk JM, Dicken V, Peitgen E, Meier S, et al. Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in ct scans. *IEEE J Select Top Sig Process.* (2009) 3:122–34. doi: 10.1109/JSTSP.2008.2011107
36. Grady L. Random walks for image segmentation. *IEEE Trans Pattern Anal Mach Intell.* (2006) 28:1768–83. doi: 10.1109/TPAMI.2006.233
37. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* (2006) 26:565–74. doi: 10.1177/0272989X06295361
38. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep.* (2015) 5:13087. doi: 10.3389/fonc.2015.00272
39. Ohri N, Duan F, Snyder BS, Wei B, Machtay M, Alavi A, et al. Pretreatment 18F-FDG PET textural features in locally advanced non-small cell lung cancer: secondary analysis of ACRIN 6668/RTOG 0235. *J Nucl Med.* (2016) 57:842–8. doi: 10.2967/jnumed.115.166934
40. Wu J, Aguilera T, Shultz D, Gudur M, Rubin DL, Loo BW Jr, et al. Early-stage non-small cell lung cancer: quantitative imaging characteristics of (18)f fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology.* (2016) 281:270–8. doi: 10.1148/radiol.2016151829
41. Leger S, Zwanenburg A, Pilz K, Lohaus F, Linge A, Zöphel K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep.* (2017) 7:13206. doi: 10.1038/s41598-017-13448-3

42. Deist TM, Dankers FJWM, Valdes G, Wijsman R, Hsu IC, Oberije C, et al. Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers. *Med Phys.* (2018) 45:3449–59. doi: 10.1002/mp.12967
43. Kirienko M, Cozzi L, Antunovic L, Lozza L, Fogliata A, Voulaz E, et al. Prediction of disease-free survival by the PET/CT radiomic signature in non-small cell lung cancer patients undergoing surgery. *Eur J Nucl Med Mol Imaging.* (2018) 45:207–17. doi: 10.1007/s00259-017-3837-7
44. Mattonen SA, Davidzon GA, Bakr S, Echegaray S, Leung ANC, Vasanawala M, et al. [18F] FDG positron emission tomography (PET) tumor and penumbra imaging features predict recurrence in non-small cell lung cancer. *Tomography.* (2019) 5:145–53. doi: 10.18383/j.tom.2018.00026
45. Wang L, Dong T, Xin B, Xu C, Guo M, Zhang H, et al. Integrative nomogram of CT imaging, clinical, and hematological features for survival prediction of patients with locally advanced non-small cell lung cancer. *Eur Radiol.* (2019) 29:2958–67. doi: 10.1007/s00330-018-5949-2
46. Yang L, Yang J, Zhou X, Huang L, Zhao W, Wang T, et al. Development of a radiomics nomogram based on the 2D and 3D CT features to predict the survival of non-small cell lung cancer patients. *Eur Radiol.* (2019) 29:2196–206. doi: 10.1007/s00330-018-5770-y
47. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology.* (2016) 281:947–57. doi: 10.1148/radiol.2016152234
48. van Timmeren JE, Carvalho S, Leijenaar RTH, Troost EGC, van Elmpt W, de Ruysscher D, et al. Challenges and caveats of a multi-center retrospective radiomics study: an example of early treatment response assessment for NSCLC patients using FDG-PET/CT radiomics. *PLoS ONE.* (2019) 14:e0217536. doi: 10.1371/journal.pone.0217536
49. Krarup MMK, Nygård L, Vogelius IR, Andersen FL, Cook G, Goh V, et al. Heterogeneity in tumours: Validating the use of radiomic features on F-FDG PET/CT scans of lung cancer patients as a prognostic tool. *Radiother Oncol.* (2020) 144:72–8. doi: 10.1016/j.radonc.2019.10.012
50. Hatt M, Laurent B, Ouahabi A, Fayad H, Tan S, Li L, et al. The first MICCAI challenge on PET tumor segmentation. *Med Image Anal.* (2018) 44:177–95. doi: 10.1016/j.media.2017.12.007
51. Berthon B, Spezi E, Galavis P, Shepherd T, Apte A, Hatt M, et al. Toward a standard for the evaluation of PET-Auto-Segmentation methods following the recommendations of AAPM task group No.211: requirements and implementation. *Med Phys.* (2017) 44:4098–111. doi: 10.1002/mp.12312

Conflict of Interest: SD was employed by GE Healthcare China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 24 July 2020
doi: 10.3389/fonc.2020.01200

A Nomogram Modeling ^{11}C-MET PET/CT and Clinical Features in Glioma Helps Predict IDH Mutation

Weiyan Zhou[1†], Zhirui Zhou[2†], Jianbo Wen[3†], Fang Xie[1], Yuhua Zhu[1], Zhengwei Zhang[1], Jianfei Xiao[1], Yijing Chen[1], Ming Li[1], Yihui Guan[1*] and Tao Hua[1*]*

[1] PET Center, Huashan Hospital, Fudan University, Shanghai, China, [2] Department of Radiotherapy, Huashan Hospital, Fudan University, Shanghai, China, [3] Department of Radiology, Huashan Hospital, Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
Fatima-Zohra Mokrane, Université Toulouse III Paul Sabatier, France

Reviewed by:
Qijun Shen, Hangzhou First People's Hospital, China
Koichiro Abe, Tokyo Medical University, Japan
Lawrence O. Dierickx, Institut universitaire du Cancer de Toulouse Oncopole, France

***Correspondence:**
Ming Li
mingli_cool@163.com
Yihui Guan
guanyihui@fudan.edu.cn
Tao Hua
tao.hua@live.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 30 March 2020
Accepted: 12 June 2020
Published: 24 July 2020

Citation:
Zhou W, Zhou Z, Wen J, Xie F, Zhu Y, Zhang Z, Xiao J, Chen Y, Li M, Guan Y and Hua T (2020) A Nomogram Modeling ^{11}C-MET PET/CT and Clinical Features in Glioma Helps Predict IDH Mutation. Front. Oncol. 10:1200. doi: 10.3389/fonc.2020.01200

Purpose: We developed a ^{11}C-Methionine positron emission tomography/computed tomography (^{11}C-MET PET/CT)-based nomogram model that uses easy-accessible imaging and clinical features to achieve reliable non-invasive isocitrate dehydrogenase (IDH)-mutant prediction with strong clinical translational capability.

Methods: One hundred and ten patients with pathologically proven glioma who underwent pretreatment ^{11}C-MET PET/CT were retrospectively reviewed. IDH genotype was determined by IDH1 R132H immunohistochemistry staining. Maximum, mean and peak tumor-to-normal brain tissue (TNRmax, TNRmean, TNRpeak), metabolic tumor volume (MTV), total lesion methionine uptake (TLMU), and standard deviation of SUV (SUV_{SD}) of the lesions on MET PET images were obtained via a dedicated workstation (Siemens. syngo.via). Univariate and multivariate logistic regression models were used to identify the predictive factors for IDH mutation. Nomogram and calibration plots were further performed.

Results: In the entire population, TNRmean, TNRmax, TNRpeak, and SUV_{SD} of IDH-mutant glioma patients were significantly lower than these values of IDH wildtype. Receiver operating characteristic (ROC) analysis suggested SUV_{SD} had the best performance for IDH-mutant discrimination (AUC = 0.731, cut-off ≤0.29, $p < 0.001$). All pairs of the ^{11}C-MET PET metrics showed linear associations by Pearson correlation coefficients between 0.228 and 0.986. Multivariate analyses demonstrated that SUV_{SD} (>0.29 vs. ≤0.29 OR: 0.053, $p = 0.010$), dichotomized brain midline structure involvement (no vs. yes OR: 26.52, $p = 0.000$) and age (≤45 vs. >45 years OR: 3.23, $p = 0.023$), were associated with a higher incidence of IDH mutation. The nomogram modeling showed good discrimination, with a C-statistics of 0.866 (95% CI: 0.796–0.937) and was well-calibrated.

Conclusions: ^{11}C-Methionine PET/CT imaging features (SUV_{SD} and the involvement of brain midline structure) can be conveniently used to facilitate the pre-operative prediction of IDH genotype. The nomogram model based on ^{11}C-Methionine PET/CT and clinical age features might be clinically useful in non-invasive IDH mutation status prediction for untreated glioma patients.

Keywords: methionine, PET, nomogram, gliomas, isocitrate dehydrogenase mutation

INTRODUCTION

Gliomas are the most prevalent malignant primary tumors of the brain. Over the past years, isocitrate dehydrogenase enzyme (IDH) mutations have been proven to be an inciting event in gliomagenesis, which made a great difference in the molecular and genetic route of oncogenic progression and clinical outcome (1). IDH mutations were identified in low grade glioma (LGG) and secondary glioblastoma multiforme (GBM) with a high percentage but in primary GBM with a much lower percentage (2). Glioma patients with IDH mutation had been prone to significantly better progression-free survival than those IDH wildtype counterparts, irrespective of grade or received treatments (3). Thereafter, some IDH wildtype LGGs can be as aggressive and have prognoses that are quite similar to GBMs (4). The gold standard of IDH mutations detection relies on immunohistochemistry or genetic sequencing of the surgical specimens. Given the inherent risk of surgery or biopsy, substantial research efforts have focused on the pre-operative non-invasive prediction of IDH mutational status in gliomas.

In 2016, the World Health Organization (WHO) updated the classification criteria for central nervous system tumors, in which IDH mutation and 1p/19q codeletion made a significant difference in the latest classification of glioma (5). The amino acid PET imaging has become increasingly important in evaluating the atypical non-enhancing gliomas as well as the differentiating tumor progression from treatment-related changes (6). Response Assessment in Neuro-Oncology (RANO) working group proposed that amino acid positron emission tomography (PET) imaging should be used in all aspects of glioma management combined with magnetic resonance imaging (MRI). L-[methyl-11C]methionine (^{11}C-MET) PET imaging has been widely used in glioma grading, differential diagnosis, tumor scope definition, brain biopsy site determination, radiotherapy planning, prognostication, and treatment monitoring (7–12).

Radiomics analysis from multimodality MRI or FDG PET images have been reported to be sufficient for IDH prediction (13). A recent study (14) by Maldjian et al. evaluated the usefulness of a non-invasive, only T2 weighted MRI based deep-learning method for the determination of IDH status. The results are inspiring since T2-weighted MR imaging is widely available and routinely performed in the assessment of gliomas. Some studies have explored the relationship between amino acid uptake characteristics of gliomas and IDH mutation status (15–19). We aimed to develop a novel and convenient statistical model that combines PET features and clinical factors for an IDH-predictive signature. Nomogram is a prediction tool that creates a simple pictorial representation of a statistical prediction model that generates a probability of a clinical event and aid in clinical decision-making (20, 21). It is "a form of line chart showing scales of the variables involved in a particular formula in a way that corresponding values for each variable lie in a straight line intersecting all the scales."(22) Therefore we tried to establish a MET PET/CT-based nomogram model that uses easy-accessible imaging metrics and clinical features to add reliable predictive information for IDH mutational status in patients with gliomas.

MATERIALS AND METHODS

Study Population

We conducted a retrospective study of patients with histologically proven diffuse glioma who underwent ^{11}C-MET PET/CT between February 2012 and November 2017 at a single center. Inclusion criteria: (1) all patients were confirmed to have glioma histological diagnosis and IDH1 R132H immunohistochemistrical staining results. (2) PET images for every patient were of good quality with no obvious artifacts. Exclusion criteria: (i) patients who received treatment by radiotherapy, chemotherapy, or chemoradiotherapy before PET imaging. (ii) glioma patients with no precise histological grading or IDH1R132H staining results. (iii) poor image quality with artifacts affecting the semi-quantitative analysis. (iv) hypo- or iso-metabolism of ^{11}C-MET compared to the background which is not applicable for threshold-based tumor volume delineation procedures. Moreover, the interval between PET imaging and subsequent tumor resection or biopsy was no more than 100 days for grade II or III gliomas and no more than 30 days for grade IV glioblastomas. A total of 110 cases were eligible for inclusion.

^{11}C-MET PET/CT Imaging Protocol

^{11}C-MET was synthesized by the GE Healthcare-Tracerlab-FXc ^{11}C radiolabelling module semi-automatically. ^{11}C-CO_2 was produced by SIEMENS RDS III cyclotron, and homo-cysteine was used as the precursor. The radiochemical purity of the obtained sterile product was higher than 95%. All patients had fasted for at least 4 h before imaging. At 10–15 min after an intravenous bonus injection of ^{11}C-MET (370–550 MBq), a static PET scan was subsequently collected for 20 min with a Siemens Biograph 64 HD PET/CT (Siemens, Erlangen, Germany) in 3-dimensional (3D) mode. PET images were reconstructed using the filter back projection (FBP) with Gaussian filter (FWHM 3.5 mm) and a 256*256 matrix, providing 64 contiguous transaxial slices of 5 mm-thick spacing. Attenuation correction was performed using a low-dose CT (150 mAs, 120 kV, Acq. 64*0.6 mm) before the emission scan.

^{11}C-MET PET/CT Data Analysis

All PET/CT images were analyzed using a dedicated workstation (Siemens.syngo.via). Semi-quantitative analysis of tumor metabolic activity was obtained using SUV normalized to body weight. All parameters were assessed in 3-dimensional volumes. Mean standardized uptake values (SUVmean) of the normal contralateral frontal cortex were calculated as references. A predefined threshold method at 1.3-times of the corresponding reference SUVmean (23, 24) was applied. The brain MRI of patients were reviewed initially to locate the possible tumor region. A VOI isocontour of the tumor region were applied, semi-quantitative PET imaging analysis were carried out after the lesion delineation procedures. The above-mentioned procedures were carried out by two experienced nuclear medicine physicians separately to double confirm the correct inclusion and reproducible parameters measurements of the glioma lesion. For those multifocal glioma patients in our group, the specific surgical resected or biopsied lesion for pathological

examination were included in our research in order to avoid bias. Each VOI generated a maximum of SUV (SUVmax), a mean SUV (SUVmean), a peak SUV (SUVpeak), a standard deviation of SUVmean (SUV_{SD}), a metabolic tumor volume (MTV) and a total lesion methionine uptake (TLMU). The total lesion methionine uptake (TLMU) was defined as the MTV multiplied by the SUVmean within the tumor boundary. SUVpeak was the highest mean SUV from a fixed 1-cm^3 spherical volume centered over the highest metabolic part of the tumor. The lesion SUV/normal contralateral cortical SUVmean was defined as the tumor-to-normal brain tissue ratio (TNR) of ^{11}C-MET uptake.

Physicians would examine the interested glioma lesions to decide whether brain midline structure were involved or not, mainly taking MET PET images for reference. The brain midline structures included corpus calloum, cingulate gyrus, thalamus, third ventricle and brain stem. As illustrated above, two physicians performed VOI delineation for each included patient to confirm the brain midline structure involvement status and to obtain two sets of MET metric features. In order to build a relatively stable integrated ^{11}C-MET PET/CT metrics-based model, we evaluated the inter-observer agreement indices for those obtained results.

Neuropathologic Analyses

Histological specimens were obtained by surgery or stereotactic brain biopsy. H&E staining and immunohistochemical analysis were performed by an experienced neuropathologist according to the current WHO guidelines. IDH status of the surgical samples was identified with an antibody to the IDH1 (R132H) mutation by immunohistochemical staining.

Establishment of a MET PET-Based Nomogram and Validation of the Model Performance

Participant's age, gender and brain midline structure involvement were used as potential predictors, together with those MET-PET metrics, to perform the univariate logistic regression analysis for developing a prediction model of IDH mutation. Those MET-PET metrics share a deep homology, so we first evaluated their correlations to avoid overfitting in the nomogram model building. After that, the MET PET metrics-based nomogram was then designed based on a multivariable logistic analysis results in the whole group with the aim of providing the clinician with a quantitative tool used in the prediction of IDH mutation status. The nomogram model validation involved the quantitative assessment of the nomogram's accuracy in IDH mutation prediction by use of Harrell's concordance index (C-statistic) and calibration curve. The corrected C-index, which is used to quantify the level of concordance between predicted probabilities and actual chance, was measured to predict the accuracy (discrimination) of the nomogram (20, 25). A relatively corrected C-index could be calculated after bootstrap analyses using 1,000 resamples. The calibration curve was used to estimate how closely the modeled nomogram estimated the risk relative to the actual risk of IDH status (mutant or wildtype), accompanied by the Hosmer-Lemeshow test (26).

Statistical Analysis

All continuous variables are expressed as mean ± standard deviation or median and range. Categorical variables are expressed as percentages. For continuous variables, an independent sample *t*-test was used to compare the two groups, while the chi-square test was applied to calculate *P*-values for categorical variables. Inter-observer agreements on ^{11}C-MET PET metrics and dichotomized location results were assessed with interclass correlation coefficients (ICC) and Cohen's kappa coefficient analysis, respectively, defined as poor (<0.2), fair (0.21–0.4), moderate (0.41–0.6), good (0.61–0.8), and very good (0.8–1.0). All PET activity measuring indices were compared with each other using scatter plots and Pearson correlation coefficients. Receiver operating characteristic (ROC) analysis was performed to calculate the area under the ROC curve (AUC) for each PET semi-quantitative parameters. The Delong test was used in the comparison of ROC curves. The AUC of ROC curves analysis and the Delong test were performed by using MedCalc for windows (version11.3.3.0, MedCalc software, Mariakierke, Belgium). Univariate and multivariate logistic regression models were used to identify the predictive factors for an IDH mutation. A nomogram was formulated based on the results of multivariate logistic regression analysis and by using the rms package of R, version 3.6.1 (http://www.r-project.org/). The predictive performance of the nomogram was measured by concordance index (C-Statistics) and calibration with 1000 bootstrap samples to decrease the overfit bias. All other statistical analysis was performed using the Prism Software version 8.0 (GraphPad, San Diego. CA). In all analyses, $P < 0.05$ was considered to indicate statistical significance.

RESULTS

Patient Demographics

The demographic data of the patients included in this study are listed in **Table 1**. Of the 110 patients who were evaluated retrospectively, 67 (59.32%) were male, and 43 (40.68%) were female, with a mean age of 45.5 years (range 10–71). The majority of patients (80/110) underwent tumor resection. The post-surgical histological examination demonstrated 59 grade II diffuse glioma, 32 grade III anaplastic tumors, and 19 grade IV glioblastomas, among which 61 patients confirmed IDH-wildtype while 49 patients were IDH-mutant. Patients with IDH-wildtype were more likely to present with lesions involving the midline structures, while there was no significant difference in gender distribution between these two groups (detailed in **Table 2**).

Inter-reader Agreement in ^{11}C-MET PET Results

The dichotomized location results of the interested tumor lesion yielded very similar values for both readers, and accordingly, the inter-observer kappa was satisfactory ($\kappa = 1.0$, $p < 0.0001$). The ICC also showed perfect agreement for the five ^{11}C-MET PET volume-based metrics (ICC > 0.95, $p < 0.0001$). Therefore, only the results of reader one were considered for further analysis.

The absolute values for all ^{11}C-MET PET metrics based on IDH-genotype were given in **Table 2**.

Correlations of ^{11}C-MET PET Metrics

All pairs of volume-based ^{11}C-MET PET metrics showed a linear association, which was quantified by Pearson correlation coefficients. There were strong correlations between paired TNRs, i.e., TNRmax, TNRmean, and TNRpeak, and SUV_{SD} with r values ranging from 0.843 to 0.986 ($p < 0.0001$). The volume-related features, including MTV and TLMU, also correlated strongly with each other ($r = 0.927$, $p < 0.0001$). Intratumoral heterogeneity feature SUV_{SD} and TNRs demonstrated fair or moderate associations with MTV ($r = 0.228$–0.370, $p < 0.05$) and TLMU ($r = 0.342$–0.430, $p < 0.0001$) (detailed in **Table 3** and **Supplementary Figures 1A,B**).

TABLE 1 | Clinicopathological features of 110 patients.

Characteristic	Numbers (Percentage %)
Age (median, range)	45.5 years old (10–71)
Gender	
Male	67 (59.32%)
Female	43 (40.68%)
Primary tumor location	
Frontal	27 (24.54%)
Parietal	6 (5.45%)
Temporal	20 (18.18%)
Occipital	1 (0.91%)
Cerebellum	5 (4.55%)
Deep brain structure	14 (12.73%)
Multifocal	37 (33.64%)
WHO grade classification	
Grade II	59 (55.09%)
Grade III	32 (28.81%)
Grade IV	19 (16.10%)
IDH status	
Mutant (Grade II/III/IV)	42/7/0 (85.71%/14.29%/0.00%)
Wildtype (Grade II/III/IV)	17/25/19 (27.87%/40.98%/31.15%)
Type of operation (surgery/stereotactic biopsy)	
Grade II	47/12 (79.66%/20.34%)
Grade III	21/11 (65.63%/34.37%)
Grade IV	14/5 (73.68%/26.32%)

Pre-operative ^{11}C-MET PET/CT ROC Analysis for IDH Mutation

As shown in **Table 2**, IDH-wildtype patients had significantly higher TNRmax, TNRmean, and TNRpeak values. Lower SUV_{SD} values were shown in IDH-mutant patients. Lower MTV and TLMU values were also observed in the IDH-mutant group, albeit not significantly so.

In the ROC analysis, the highest AUC of 0.731 (95%CI: 0.638–0.811) was reached by the SUV_{SD} value, with the best cut-off value at 0.29, a specificity of 60.66% and a sensitivity of 77.55%, followed by the TNRmax value with an AUC of 0.678, the best cut-off value at 2.99, a specificity of 59.02% and a sensitivity of 75.51%. Their optimal cutoff, AUC, sensitivity, specificity values, etc., for the abovementioned ^{11}C-MET MET PET metrics were listed in **Table 4**. Their AUC curves were displayed in **Supplementary Figure 2**.

Further pairwise comparisons of ROC curves confirmed that the AUC of SUV_{SD} differed significantly from any other MET PET metrics ($p < 0.05$, details in **Supplementary Table 1**).

Predicting IDH-Mutant Gliomas and Construction of the Nomogram

In univariate analysis, the MET PET metrics including SUV_{SD}, TNRmax, TNRpeak and TNRmean, except for MTV and TLMU, were significantly associated with IDH mutation ($p < 0.05$ for all the variables). Considering their collinearity and the AUC curve comparison results for the MET PET metrics, SUV_{SD} was selected as the only MET PET feature for further multivariate logistic regression analysis. In multivariate logistic regression analysis, the three factors, i.e., participant's age, the involvement of midline structure, and SUV_{SD}, were found to be significant independent predictors. We demonstrated that SUV_{SD} (>0.29 vs. ≤0.29 OR: 0.053, $p = 0.010$), brain midline structure involvement

TABLE 2 | Clinical features and ^{11}C-MET PET metrics based on IDH-genotype.

PET Metric	All patients ($n = 110$)	IDH-mutant ($n = 49$)	IDH-wildtype ($n = 61$)	*P*-value[a]
Age (mean ± SD)	45.08 ± 13.56	42.63 ± 10.6	47.05 ± 15.35	0.090
Age (≤45/>45 years)	55/55	31/18	24/37	0.013
Gender (M/F)	67/43	27/22	40/21	0.263
Midline Involvement (yes/no)	37/73	2/47	35/26	0.000
TNRmax	1.7719 ± 0.3038	1.6692 ± 0.2474	1.8544 ± 0.3211	0.001
TNRmean	3.2421 ± 1.3193	2.8277 ± 1.1741	3.5749 ± 1.3440	0.002
TNRpeak	2.8114 ± 1.1222	2.4739 ± 0.9704	3.0824 ± 1.1690	0.004
MTV	48.6750 ± 54.9081	44.9551 ± 53.7083	51.6631 ± 56.1161	0.525
TLMU	87.9881 ± 104.0785	68.5628 ± 86.9409	103.5920 ± 114.3505	0.071
SUV_{SD}	0.3783 ± 0.2819	0.2551 ± 0.1781	0.4772 ± 0.3108	0.000

[a]*comparison between IDH-mutant and IDH-wildtype.*

TABLE 3 | Correlation of ^{11}C-MET PET metrics.

PET Metric	TNRmax	TNRmean	TNRpeak	MTV	TLMU	SUV_{SD}
TNRmax	1	0.843	0.986	0.364	0.419	0.861
TNRmean		1	0.872	0.242	0.342	0.855
TNRpeak			1	0.370	0.430	0.876
MTV				1	0.927	0.228
TLMU					1	0.401
SUV_{SD}						1

TABLE 4 | The Performance of ^{11}C-MET PET metrics and age feature for IDH-mutation prediction.

PET metric	Cutoff	AUC (95% CI)	ACC	SEN	SPE	PPV	NPV	Youden-index
SUV_{SD}	≤0.29	0.731 (0.638–0.811)	68.18%	77.55%	60.66%	61.30%	77.10%	0.3821
TNRmax	≤2.9886	0.678 (0.582–0.764)	66.36%	75.51%	59.02%	59.70%	75.00%	0.3453
TNRmean	≤1.7051	0.679 (0.583–0.765)	65.46%	69.39%	62.30%	59.60%	71.70%	0.3168
TNRpeak	≤2.8191	0.660 (0.564-0.748)	65.45%	79.59%	54.10%	58.20%	76.70%	0.3369
Age	≤45	0.630 (0.533–0.720)	62.73%	65.31%	60.66%	57.10%	68.50%	0.2596
Midline involvement	yes	0.766 (0.676–0.842)	74.54%	95.92%	57.38%	64.40%	94.60%	0.5330

CI, Confidence interval; AUC, Area under the receiver-operating characteristic curve; ACC, Accuracy; SEN, Sensitivity; SPE, Specitivity; PPV, Positive predictive value; NPV, Negative predictive value.

TABLE 5 | Univariate and multivariate regression analyses for IDH mutation prediction.

			Multivariate analysis			
	Univariate analysis		Model 1		Model 2	
Variable	**OR (95% CI)**	***P***	**OR (95% CI)**	***P***	**OR (95% CI)**	***P***
Age (≤45 vs. >45)	2.655 (1.223–5.765)	0.014	3.232 (1.180–8.854)	0.023		
Gender (male vs. female)	0.644 (0.298–1.394)	0.264				
Midline structure Involvement (no vs. yes)	31.635 (7.035–142.248)	0.000	26.523 (5.547–126.831)	0.000	24.461 (5.305–112.789)	0.000
MTV	0.998 (0.991–1.005)	0.525				
TLMU	0.996 (0.992–1.001)	0.090				
TNRmean	0.093 (0.020–0.435)	0.003				
TNRmax	0.605 (0.427–0.857)	0.005				
TNRpeak	0.574 (0.385–0.856)	0.007				
SUV_{SD}	0.02 (0.003–0.157)	0.000	0.053 (0.006–0.497)	0.010	0.048 (0.006–0.411)	0.006
C–index			0.866 (0.796–0.937)		0.843 (0.766–0.920)	

(no vs. yes OR: 26.52, $p = 0.000$) and age (≤45 vs. >45 years OR: 3.23, $p = 0.023$), were associated with a higher incidence of IDH mutation (shown **Table 5**).

Age does not correlate with SUV_{SD} ($r = 0.0370$, $p = 0.7010$). Lesions involving brain midline structure (73 cases) showed higher SUV_{SD} (0.3147 ± 0.2501 vs. 0.5038 ± 0.3017, $p = 0.007$) compared to lesions without brain midline involvement (37 cases), but no age predominance (45.67 ± 12.10 vs. 43.92 ± 16.19, $p = 0.5245$) was observed for brain midline involvement (shown in **Supplementary Figure 3**).

After that, a nomogram was constructed on the basis of the multivariate logistic regression (for details, see **Figure 1A**). The nomogram (model 1) showed good discrimination efficacy, with a C-statistics of 0.866 (95%CI: 0.796–0.937). The calibration curve of the nomogram also indicated good agreement between predicted probability and actual occurrence in the whole cohort (**Figure 1B**). The Hosmer-Lemeshow test indicated no significant difference ($p > 0.05$), suggesting that there was no departure from a perfect fit. Meanwhile, we built model 2 only using MET PET information, i.e., brain midline structure involvement and SUV_{SD} derived from lesion VOI, with a C-statistics of 0.843 (95%CI: 0.766–0.920), suggesting that age information dichotomized by 45 years old do enhance the predictive ability for IDH genotype. **Figure 2** illustrates a comparison of two representative grade II glioma cases with similar images by visual analysis. Our ^{11}C-MET PET/CT-based nomogram could effectively distinguish between IDH-mutant and IDH-wildtype gliomas.

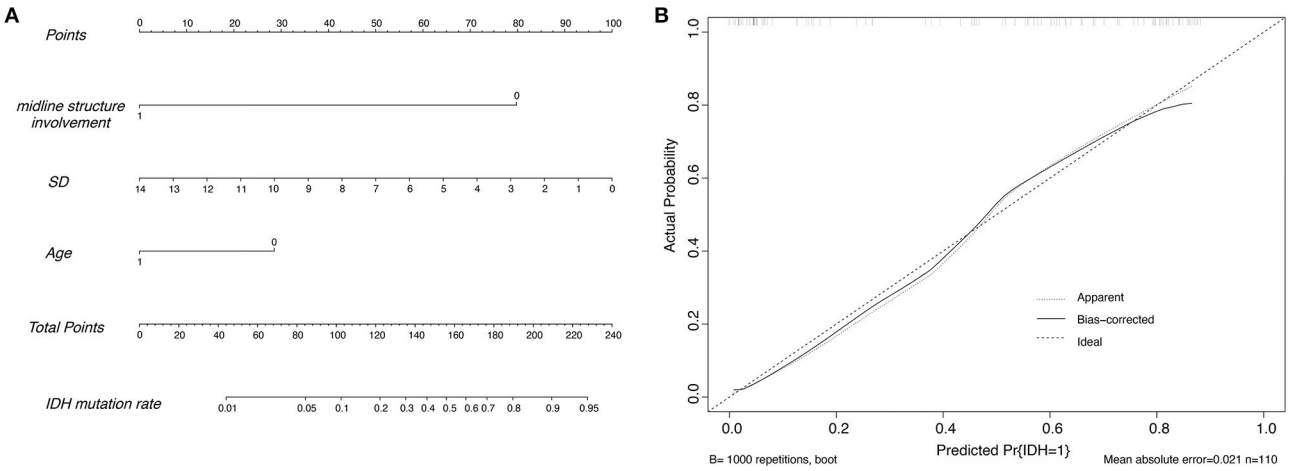


FIGURE 1 | (A) The nomogram developed in the whole cohort using the SUV_{SD} metric, the age, and the brain midline structure involvement of the patients. 0 represents ≤ 45 years for age or without brain midline structure involvement and 1 represents >45 years for age or with brain midline involvement, respectively; SD = 100*SUV_{SD}. **(B)** Calibration plots of the nomogram for predicting IDH mutation. The y-axis represents the actual probability, and the x-axis represents the predicted probability.

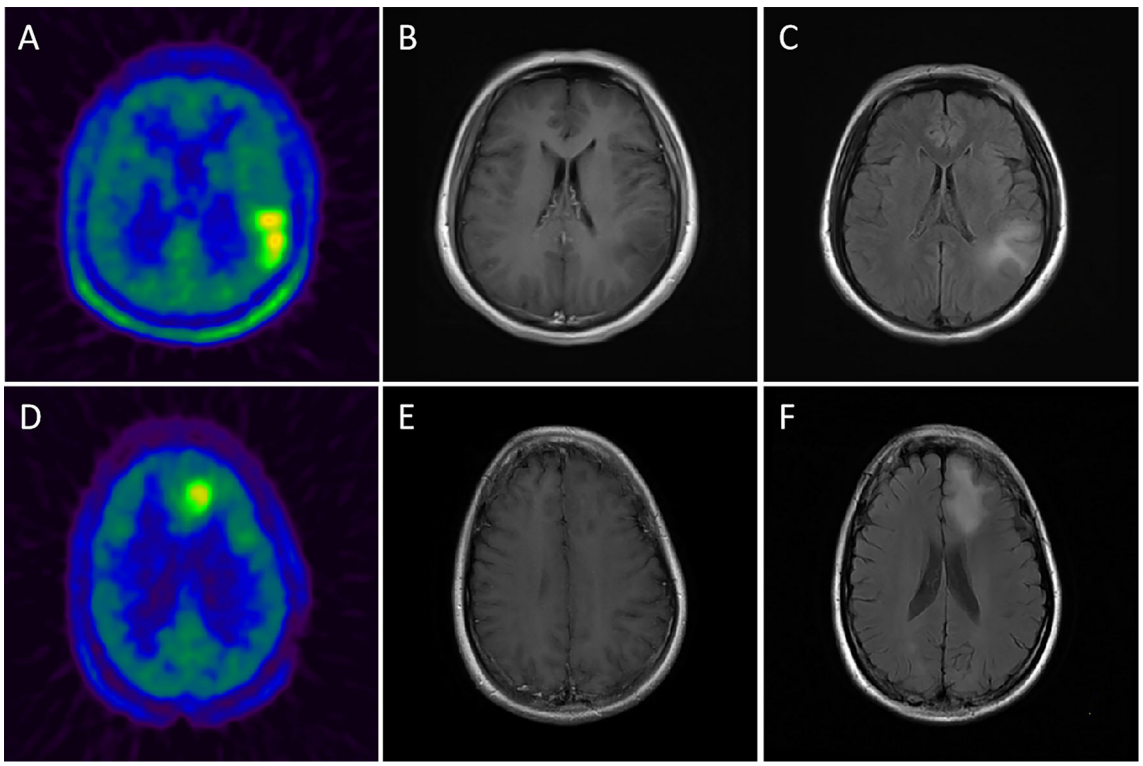


FIGURE 2 | Upper row: example of a patient with a ^{11}C-MET-positive lesion in the left temporo-parietal lobe, moderate TBRmax of 3.21 and moderate MTV of 46.1 ml **(A)** without contrast-enhancement in CE-T1 MRI **(B)** and an obvious flair alteration in T2 MRI **(C)** but with a moderate SUV_{SD} of 0.31 (>0.29); histopathological analysis revealed an IDH-wildtype, diffuse astrocytoma (WHO grade II). Lower row: example of a patient with a ^{11}C-MET-positive lesion in the left frontal lobe, moderate TBRmax of 2.46 and moderate MTV of 31.3 ml **(D)** without contrast-enhancement in CE-T1 MRI **(E)** and a flair alteration in T2 MRI **(F)** but with a small SUV_{SD} of 0.26 (<0.29); histopathological analysis revealed an IDH-mutant, diffuse astrocytoma (WHO grade II).

DISCUSSION

In the present study, we confirmed an association between volume-based ^{11}C-MET PET quantification metrics and IDH mutational status for untreated glioma patients and further constructed a novel and intuitive statistical model to help clinicians and radiologists non-invasively predict glioma IDH mutation. As expected, our data demonstrated that TNRs and SUV_{SD} were significantly lower in the IDH-mutant group compared with those IDH-wildtypes, which are consistent with those of Kim et al. (15) ^{11}C-MET PET derived SUV_{SD} showed the most excellent ability to identify whether glioma had an IDH mutation or not besides other MET PET metrics. Single-parameter SUV_{SD}, which is a sort of tumor imaging heterogeneity feature, had the best prediction efficacy in IDH mutation. It is reasonable to hypothesize that the more heterogeneous the tumor MET PET imaging, the more likely IDH status is to be wildtype.

^{11}C-MET PET played a significant role in evaluating the O_6-methylguanylmethyltransferase methylation (MGMT) status in gliomas (27, 28). PET imaging was suggested to be informative for preoperatively differentiating gliomas according to 2016 WHO classification, particularly for differentiating IDH-wildtype and IDH-mutant tumors (19). A study of hybrid ^{11}C-MET PET/MRI imaging including 39 glioma patients described that ROC analysis of TNRmax had a high AUC of 0.79 for predicting IDH status (16). Another study retrospectively evaluated 109 patients with newly diagnosed glioma also indicated that ^{11}C-MET uptake was negatively correlated with IDH mutational status. The MET uptake of IDH-wildtype glioblastoma was significantly higher than that of IDH-mutant glioma (17). TNRmax derived from ^{11}C-MET PET appears to be superior to MRS in differentiating IDH status with a ROC of 0.67 (18). The investigations on the relationship between amino acid tracer uptake and IDH status were not totally consistent. One O-(2-^{18}F-fluoroethyl)-L-tyrosine(^{18}F-FET) PET research in a mixed group of glioma patients, which included 16 oligodendrogliomas (IDH mutated and 1p/19q co-deleted), 27 astrocytomas (IDH mutated only) and 47 glioblastomas (IDH-wildtype), suggested that gliomas with IDH mutation are typically shown with a lower tumor to brain ratios(TNRmean and TNRmax), prolonged time to peak, and a slow-rise time-activity curve of 20–50 min (29). By contrast, another 3,4-dihydroxy-6-[^{18}F]fluoro-L-phenylalanine (^{18}F-FDOPA) imaging study in a total of 43 newly diagnosed glioma cases described paradoxically higher ^{18}F-FDOPA uptake in diffuse grade II and III gliomas with IDH mutation (30). This inconsistency may be explained by the different amino acids PET probe uptake models in glioma (31). The expression level of L-type amino acid transporter in glioma is positively proportional to the intake value of MET, while the expression level of amino acid transporter is positively correlated with the microvascular density of glioma (32). Literature has shown that local blood flow in IDH wildtype glioma is higher than that of IDH mutants (33).

Our study differs from the abovementioned ones in that it indicated one predictive model for proper pre-operative prediction of IDH status in glioma patients. We aimed to develop a nomogram which is independent of histopathologic features, such as tumor grade or oligodendrocyte component, etc. The model showed good discrimination and was well-calibrated. Pre-operative lesion VOI SUV_{SD} should be important in daily clinical practice, which is a convenient and repetitive PET imaging parameter obtained through glioma VOI delineation. The origin of this PET imaging parameter could reflect the intratumoral heterogeneity to some degree. The SUV_{SD} difference derived from ^{11}C-MET PET images between IDH-mutant and IDH-wildtype gliomas may help understand the possible internal link of intratumor heterogeneity and IDH mutation. Our model showed that the middle line structure involvement is associated with IDH mutational status. Things that need to be clarified is that this kind of brain midline structure involvement was also decided by VOI delineation, which could be more extensive and broader than the enhanced tumor volume in MRI. Moreover, age information has shown reasonable predictive potential and enhanced the predictive ability for IDH genotype. We report for the first time the application of ^{11}C-MET PET/CT metrics and clinical age feature based nomogram in IDH genotyping for untreated glioma patients.

From specific clinical perspective, this nomogram model has some positive features. Firstly, our predictive model takes the advantage of being able to be rapidly acquired by a radiologist without requiring specialized software extracting texture features from high-order matrixes. Secondly, it is based on repetitive MET-PET metrics and some important clinical features, which is easily for understanding and clinically viable. This nomogram model displayed the potential to be used as a standalone diagnostic modality for patients with excessive surgical risk related to patient's comorbidities, advanced age, deep-seated, or brain stem tumors, etc.

There are some limitations to the present study. First, as this was a single-center study, with more cases are recruited, the training and validation group will be set for further external validation or multicenter validation to assess the potential clinical utility of our model further. Furthermore, next-generation sequencing for the IDH genotype was not available for this retrospective study, and some patients with the mutation may have been misidentified. Non-canonical IDH mutations can be found in IDH1 R132H immune-negative LGG (34). These points would be addressed in future work.

CONCLUSIONS

This study proved that SUV_{SD} derived from regular glioma VOI delineation in MET PET imaging is a novel and convenient semiquantitative parameter for the glioma IDH prediction. The nomogram model combining with age, brain midline structure involvement, and SUV_{SD} demonstrates the potential in non-invasive IDH mutation status prediction for untreated glioma patients and showed reasonable convenience in clinical practice.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (Ethics Committee of HuaShan Hospital Fudan University-approval number 2008-82) and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by the National Natural Science Foundation (Grant no. 81701755), Shanghai Municipal Science and Technology Committee Funds (Grant nos. 18411952100 and 17411953500), Shanghai Municipal Science and Technology Major Project (Grant no. 2018SHZDZX01), and ZJLab.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01200/full#supplementary-material

REFERENCES

1. Turkalp Z, Karamchandani J, Das S. IDH mutation in glioma: new insights and promises for the future. *JAMA Neurol.* (2014) 71:1319–25. doi: 10.1001/jamaneurol.2014.1205
2. Sun H, Yin L, Li S, Han S, Song G, Liu N, et al. Prognostic significance of IDH mutation in adult low-grade gliomas: a meta-analysis. *J Neurooncol.* (2013) 113:277–84. doi: 10.1007/s11060-013-1107-5
3. Wick W, Hartmann C, Engel C, Stoffels M, Felsberg J, Stockhammer F, et al. NOA-04 randomized phase III trial of sequential radiochemotherapy of anaplastic glioma with procarbazine, lomustine, and vincristine or temozolomide. *J Clin Oncol.* (2009) 27:5874–80. doi: 10.1200/JCO.2009.23.6497
4. Dunn GP, Andronesi OC, Cahill DP. From genomics to the clinic: biological and translational insights of mutant IDH1/2 in glioma. *Neurosurg Focus.* (2013) 34:E2. doi: 10.3171/2012.12.FOCUS12355
5. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1
6. Lohmann P, Werner JM, Shah NJ, Fink GR, Langen KJ, Galldiks N. Combined amino acid positron emission tomography and advanced magnetic resonance imaging in glioma patients. *Cancers (Basel).* (2019) 11:153. doi: 10.3390/cancers11020153
7. Shinozaki N, Uchino Y, Yoshikawa K, Matsutani T, Hasegawa A, Saeki N, et al. Discrimination between low-grade oligodendrogliomas and diffuse astrocytoma with the aid of 11C-methionine positron emission tomography. *J Neurosurg.* (2011) 114:1640–7. doi: 10.3171/2010.11.JNS10553
8. Albert NL, Weller M, Suchorska B, Galldiks N, Soffietti R, Kim MM, et al. Response assessment in neuro-Oncology working group and European association for neuro-Oncology recommendations for the clinical use of PET imaging in gliomas. *Neuro Oncol.* (2016) 18:1199–208. doi: 10.1093/neuonc/now058
9. Katsanos AH, Alexiou GA, Fotopoulos AD, Jabbour P, Kyritsis AP, Sioka C. Performance of 18F-FDG, 11C-Methionine, and 18F-FET PET for glioma grading: a meta-analysis. *Clin Nucl Med.* (2019) 44:864–9. doi: 10.1097/RLU.0000000000002654
10. Glaudemans AW, Enting RH, Heesters MA, Dierckx RA, van Rheenen RW, Walenkamp AM, et al. Value of 11C-methionine PET in imaging brain tumours and metastases. *Eur J Nucl Med Mol Imaging.* (2013) 40:615–35. doi: 10.1007/s00259-012-2295-5
11. Herholz K, Holzer T, Bauer B, Schroder R, Voges J, Ernestus RI, et al. 11C-methionine PET for differential diagnosis of low-grade gliomas. *Neurology.* (1998) 50:1316–22. doi: 10.1212/WNL.50.5.1316
12. Poetsch N, Woehrer A, Gesperger J, Furtner J, Haug AR, Wilhelm D, et al. Visual and semi-quantitative 11C-methionine PET: an independent prognostic factor for survival of newly diagnosed and treatment-naive gliomas. *Neuro Oncol.* (2018) 20:411–9. doi: 10.1093/neuonc/nox177
13. Li L, Mu W, Wang Y, Liu Z, Liu Z, Wang Y, et al. A non-invasive radiomic method using (18)F-FDG PET predicts isocitrate dehydrogenase genotype and prognosis in patients with glioma. *Front Oncol.* (2019) 9:1183. doi: 10.3389/fonc.2019.01183
14. Yogananda CGB, Shah BR, Vejdani-Jahromi M, Nalawade SS, Murugesan GK, Yu FF, et al. A Novel fully automated Mri-based deep learning method for classification of idh mutation status in brain gliomas. *Neuro Oncol.* (2020) 22:402–11. doi: 10.1093/neuonc/noz199
15. Kim D, Chun JH, Kim SH, Moon JH, Kang SG, Chang JH, et al. Re-evaluation of the diagnostic performance of (11)C-methionine PET/CT according to the 2016 WHO classification of cerebral gliomas. *Eur J Nucl Med Mol Imaging.* (2019) 46:1678–84. doi: 10.1007/s00259-019-04337-0
16. Kebir S, Weber M, Lazaridis L, Deuschl C, Schmidt T, Monninghoff C, et al. Hybrid 11C-MET PET/MRI combined with "machine learning" in glioma diagnosis according to the revised glioma WHO classification 2016. *Clin Nucl Med.* (2019) 44:214–20. doi: 10.1097/RLU.0000000000002398
17. Lopci E, Riva M, Olivari L, Raneri F, Soffietti R, Piccardo A, et al. Prognostic value of molecular and imaging biomarkers in patients with supratentorial glioma. *Eur J Nucl Med Mol Imaging.* (2017) 44:1155–64. doi: 10.1007/s00259-017-3618-3
18. Kebir S, Lazaridis L, Weber M, Deuschl C, Stoppek AK, Schmidt T, et al. Comparison of L-Methyl-11C-Methionine PET with magnetic resonance spectroscopy in detecting newly diagnosed glioma. *Clin Nucl Med.* (2019) 44:e375–81. doi: 10.1097/RLU.0000000000002577
19. Takei H, Shinoda J, Ikuta S, Maruyama T, Muragaki Y, Kawasaki T, et al. Usefulness of positron emission tomography for differentiating gliomas according to the 2016 World Health Organization classification of tumors of the central nervous system. *J Neurosurg.* (2019) 16:1–10. doi: 10.3171/2019.5.JNS19780
20. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* (2008) 26:1364–70. doi: 10.1200/JCO.2007.12.9791
21. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* (2015) 16:e173–80. doi: 10.1016/S1470-2045(14)71116-7
22. Grimes DA. The nomogram epidemic: resurgence of a medical relic. *Ann Intern Med.* (2008) 149:273–5. doi: 10.7326/0003-4819-149-4-200808190-00010
23. Kawai N, Maeda Y, Kudomi N, Miyake K, Okada M, Yamamoto Y, et al. Correlation of biological aggressiveness assessed by 11C-methionine PET and hypoxic burden assessed by 18F-fluoromisonidazole PET in newly diagnosed glioblastoma. *Eur J Nucl Med Mol Imaging.* (2011) 38:441–50. doi: 10.1007/s00259-010-1645-4
24. Galldiks N, Ullrich R, Schroeter M, Fink GR, Jacobs AH, Kracht LW. Volumetry of [(11)C]-methionine PET uptake and MRI contrast

enhancement in patients with recurrent glioblastoma multiforme. *Eur J Nucl Med Mol Imaging.* (2010) 37:84–92. doi: 10.1007/s00259-009-1219-5
25. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal Carcinoma. *Clin Cancer Res.* (2017) 23:4259–69. doi: 10.1158/1078-0432.CCR-16-2910
26. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. *Crit Care Med.* (2007) 35:2052–6. doi: 10.1097/01.CCM.0000275267.64078.B0
27. Yu P, Ning J, Xu B, Liu J, Dang H, Lin M, et al. Histogram analysis of 11C-methionine integrated PET/MRI may facilitate to determine the O6-methylguanylmethyltransferase methylation status in gliomas. *Nucl Med Commun.* (2019) 40:850–6. doi: 10.1097/MNM.0000000000001039
28. Okita Y, Nonaka M, Shofuda T, Kanematsu D, Yoshioka E, Kodama Y, et al. (11)C-methinine uptake correlates with MGMT promoter methylation in nonenhancing gliomas. *Clin Neurol Neurosurg.* (2014) 125:212–6. doi: 10.1016/j.clineuro.2014.08.004
29. Verger A, Stoffels G, Bauer EK, Lohmann P, Blau T, Fink GR, et al. Static and dynamic F-FET PET for the characterization of gliomas defined by IDH and 1p/19q status. *Eur J Nucl Med Mol Imaging.* (2018) 45:443–51. doi: 10.1007/s00259-017-3846-6
30. Verger A, Metellus P, Sala Q, Colin C, Bialecki E, Taieb D, et al. IDH mutation is paradoxically associated with higher (18)F-FDOPA PET uptake in diffuse grade II and grade III gliomas. *Eur J Nucl Med Mol Imaging.* (2017) 44:1306–11. doi: 10.1007/s00259-017-3668-6
31. Verger A, Taieb D, Guedj E. Is the information provided by amino acid PET radiopharmaceuticals clinically equivalent in gliomas? *Eur J Nucl Med Mol Imaging.* (2017) 44:1408–10. doi: 10.1007/s00259-017-3710-8
32. Okubo S, Zhen HN, Kawai N, Nishiyama Y, Haba R, Tamiya T. Correlation of L-methyl-11C-methionine (MET) uptake with L-type amino acid transporter 1 in human gliomas. *J Neurooncol.* (2010) 99:217–25. doi: 10.1007/s11060-010-0117-9
33. Kickingereder P, Sahm F, Radbruch A, Wick W, Heiland S, Deimling A, et al. IDH mutation status is associated with a distinct hypoxia/angiogenesis transcriptome signature which is non-invasively predictable with rCBV imaging in human glioma. *Sci Rep.* (2015) 5:16238. doi: 10.1038/srep16238
34. Barresi V, Eccher A, Simbolo M, Cappellini R, Ricciardi GK, Calabria F, et al. Diffuse gliomas in patients aged 55 years or over: a suggestion for IDH mutation testing. *Neuropathology.* (2020) 40:68–74. doi: 10.1111/neup.12608

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 31 July 2020
doi: 10.3389/fonc.2020.01090

Comparison of Metabolic and Morphological Response Criteria for Early Prediction of Response and Survival in NSCLC Patients Treated With Anti-PD-1/PD-L1

Angelo Castello[1], *Sabrina Rossi*[2], *Luca Toschi*[2] *and Egesta Lopci*[1]*

[1] Department of Nuclear Medicine, Humanitas Clinical and Research Hospital—IRCCS, Milan, Italy, [2] Medical Oncology, Humanitas Clinical and Research Hospital—IRCCS, Milan, Italy

OPEN ACCESS

Edited by:
Romain-David Seban,
Institut Curie, France

Reviewed by:
Olivier Humbert,
Université Côte d'Azur, France
Antoine Girard,
Centre Eugène Marquis, France

***Correspondence:**
Egesta Lopci
egesta.lopci@humanitas.it;
egesta.lopci@gmail.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 07 April 2020
Accepted: 01 June 2020
Published: 31 July 2020

Citation:
Castello A, Rossi S, Toschi L and Lopci E (2020) Comparison of Metabolic and Morphological Response Criteria for Early Prediction of Response and Survival in NSCLC Patients Treated With Anti-PD-1/PD-L1. Front. Oncol. 10:1090. doi: 10.3389/fonc.2020.01090

Introduction/Aim: Immunotherapy with immune checkpoint inhibitors (ICIs) has positively changed the history of several malignant tumors. In parallel, new challenges have emerged in the evaluation of treatment response as a result of their peculiar anticancer effect. In the current study, we aimed to compare different response criteria, both morphological and metabolic, for assessing response and outcome in patients with advanced non-small cell lung cancer (NSCLC) treated with ICI.

Materials and Methods: Overall, 52 patients with advanced NSCLC candidate to ICI were prospectively evaluated. Inclusion criteria comprised whole-body contrast-enhanced CT and ^{18}F-FDG PET/CT at baseline and at the first response evaluation 3 or 4 cycles after ICI. Response assessment on CT was performed according to RECIST 1.1 and imRECIST criteria, whereas metabolic response on PET was computed by EORTC, PERCIST, imPERCIST, and PERCIMT criteria. The concordance between the different tumor response criteria and the performance of each criterion to predict progression-free survival (PFS) and overall survival (OS) were calculated.

Results: Inclusion criteria were fulfilled in 35 out of 52 patients. We observed a low agreement between imRECIST and imPERCIST ($\kappa = 0.143$) with discordant response in 20 patients, particularly regarding stable disease and progressive disease groups. Fair agreement between imRECIST and EORTC ($\kappa = 0.340$), and PERCIST ($\kappa = 0.342$), and moderate for PERCIMT ($\kappa = 0.413$) were detected. All criteria were significantly associated with PFS, while only PERCIMT and imPERCIST were associated with OS. Of note, in patients classified as immune stable disease (iSD), imPERCIST, and PERCIMT well-differentiated those with longer PFS ($p < 0.001$, $p = 0.009$) and OS ($p = 0.001$, $p = 0.002$). In the multivariate analysis, performance status [hazard ratio (HR) = 0.278, $p = 0.015$], imRECIST (HR = 3.799, $p = 0.026$), and imPERCIST (HR = 4.064, $p = 0.014$) were predictive factors for PFS, while only performance status (HR = 0.327, $p = 0.035$) and imPERCIST (HR = 3.247, $p = 0.007$) were predictive for OS.

Conclusions: At the first evaluation during treatment with ICI, imPERCIST criteria correctly evaluated treatment response and appeared able to predict survival. Moreover, in patients with iSD on CT, imPERCIST were able to discriminate those with longer survival. This advantage might allow for earlier therapy modification based on metabolic response.

Keywords: non-small cell lung cancer, checkpoint inhibitors, ^{18}F-FDG PET/CT, RECIST, EORTC, PERCIST, imPERCIST, PERCIMT

INTRODUCTION

Several clinical studies have demonstrated the successful therapeutic approach of immune checkpoint inhibitors (ICIs) in patients affected by different malignancies when compared with chemotherapy. As a matter of fact, these new agents, acting against cytotoxic T-lymphocyte-associated antigen (CTLA)-4, and anti-programmed death (PD)-1 or its ligand (PD-L1), have been approved so far for over 18 types of cancer (1–3). However, owing to the peculiar response patterns observed in these immune-modulating agents, in parallel with the increased use of ICI, also the assessment of tumor response by medical imaging has become more challenging. Indeed, PD-1/L1 and CTLA-4 blockage aims to restore the immune response by recalling neutrophils, macrophages, and activating T cells within the tumor microenvironment. Consequently, because of tumor inflammation, malignant lesions might appear stable, or even larger either in size or in metabolic activity before effective shrinkage occurs, making it difficult to discriminate between true progression from the so-called pseudo-progression (4–7). To overcome these limitations, numerous response criteria have been proposed, starting with the traditional Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 (8). The main peculiarity of the new morphological criteria developed in the ICI era, such as immune-related response criteria (irRC) and immune-modified (im)RECIST, is that the appearance of new lesions is not always synonymous with progression of disease, but requires confirmation at least after 4–8 weeks (9, 10). Likewise, metabolic criteria based on ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) have been modified aiming to improve diagnostic accuracy during immunotherapy. Of note, new lesions are considered a sign of progression according to their number and size or if metabolic activity is greater than a determined cut-off, as proposed by PET Response Evaluation Criteria for Immunotherapy (PERCIMT) and Immunotherapy-modified PET Response Criteria in Solid Tumors (imPERCIST) criteria, respectively (11, 12).

The purpose of the present study was to investigate the concordance between morphological and metabolic criteria for early response evaluation and to correlate findings with survival in patients with advanced non-small cell lung cancer (NSCLC) undergoing treatment with checkpoint inhibitors.

MATERIALS AND METHODS

Study Population

From December 2015 to May 2019, patients with histopathologically proven advanced NSCLC who were scheduled to undergo ICI treatment were enrolled. Prospective data were collected from patients ($n = 42$) adhering to the same diagnostic trial, registered at https://clinicaltrials.gov/ (NCT03563482), and from other clinical trials for ICI ($n = 10$). Eligible patients were required to have both contrast-enhanced CT and ^{18}F-FDG PET/CT scan within 1 month before starting ICI and a second scan at the first restaging after 3 cycles for pembrolizumab or 4 cycles for nivolumab (**Figure 1**). Moreover, all patients repeated CT every 3 or 4 cycles until confirmed progression. Other exclusion criteria were as follows: primary malignancy other than NSCLC; no lesion on ^{18}F-FDG PET/CT above the minimum standardized uptake value (SUV) normalized to lean body mass (SUL) as defined by PERCIST (1.5 × liver SUL + 2 SDs of liver SUL) (13); plasma glucose level was ≥200 mg/dL before ^{18}F-FDG PET/CT. The study has been approved by the local institutional review board and in accordance with Declaration of Helsinki and Good Clinical Practice guidelines. Written informed consent was obtained in all cases.

Imaging Protocol

^{18}F-FDG PET/CT

PET/CT scans were performed as previously described (14). A GE ADW4.6 workstation (GE Healthcare, Waukesha, WI, USA) was used to display images, which were interpreted by two experienced nuclear medicine physicians. For the semi-quantitative analysis, the threshold of the volumes of interest (VOIs) was set at 0.5 by PETVCAR (GE Healthcare). The maximum SUV (SUVmax) was defined as the value of the highest pixel and average SUV (SUVmean) as the mean SUV related to the tumor burden. To determine the peak SUV corrected for lean body mass (SULpeak), the reviewer placed a sphere or cube as the VOI around the hottest lesions (up to five lesions, no more than two per organ). Within this VOI, the software searched for the 1.0-cm^3 sphere that encompassed the voxels with the highest average SUL. For background activity, a 3-cm-diameter spherical VOI was delineated in the right lobe of the liver or in the descending thoracic aorta for patients with liver involvement. Response of SULpeak (%) was defined as (sum of baseline SULpeak—sum of follow-up SULpeak)/(sum of baseline

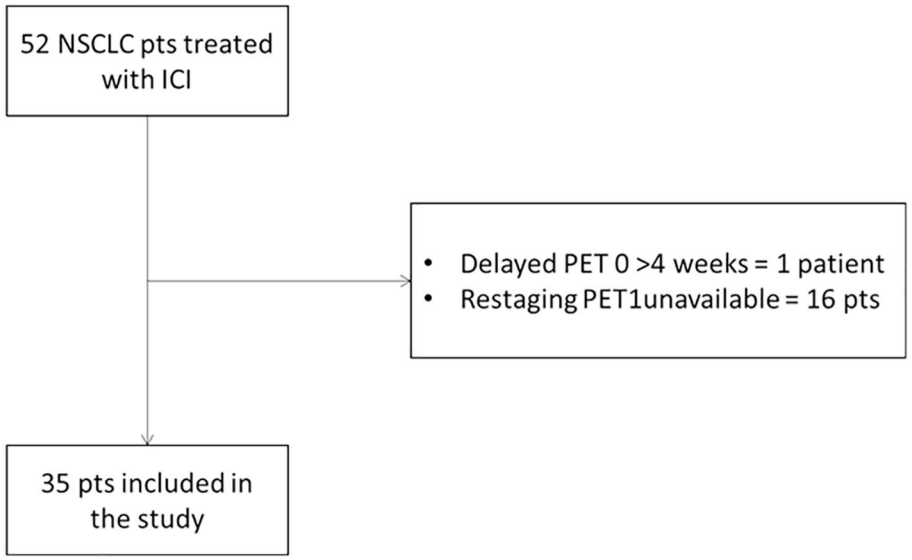


FIGURE 1 | Flowchart of patient accrual.

SULpeak) × 100. Target lesions on follow-up scans were not necessarily the same as target lesions at baseline (13).

Response Assessment

Two physicians (E.L., A.C.), specializing in immunotherapy evaluation, reviewed all consecutive scans to reach a consensus. Morphological evaluation was determined according to RECIST 1.1 and imRECIST (8, 10). Metabolic response on ^{18}F-FDG PET/CT was defined according to the European Organization for Research and Treatment of Cancer (EORTC) criteria, PERCIST, and its variation imPERCIST (12, 13, 15). Response Evaluation Criteria for Immunotherapy (PERCIMT) were also considered in our analysis (11). **Supplementary Table 1** details the response categories. Briefly, we consider four response categories: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD) for morphological criteria. Likewise, complete metabolic response (CMR), partial metabolic response (PMR), stable metabolic disease (SMD), and progressive metabolic disease (PMD) were considered for metabolic criteria.

Statistical Analysis

The concordance among response criteria was assessed using Cohen's κ coefficient. Agreement between the two assessments was categorized as poor (weighted $\kappa < 0.2$), fair (weighted $\kappa = 0.21$–0.40), moderate (weighted $\kappa = 0.41$–0.60), good (weighted $\kappa = 0.61$–0.80), and almost perfect (weighted $\kappa > 0.80$) (16). Progression-free survival (PFS) was calculated as the interval from the date of initiation of ICI to the date of either disease progression or death, whereas overall survival (OS) was calculated as the duration between the date of initiation of immunotherapy and the date of death from any cause (17). PFS and OS were analyzed using the Kaplan–Meier method and log-rank test. Then, forward stepwise multivariate regression analysis was performed to identify factors correlated with PFS and OS based on the calculation of hazard ratios (HRs) and 95% CI (14). Variables included in the final multivariate analysis were selected according to their clinical relevance and statistical significance in a univariate model (cut-off, $p < 0.10$). All statistical analyses were carried out using the Statistical Package for Social Sciences, version 23.0, for Windows (SPSS, Chicago, IL), and $p < 0.05$ were considered to be statistically significant (17).

RESULTS

Patient Characteristics

Out of the 52 patients with metastatic NSCLC enrolled in the clinical trial, 35 patients (23 men and 12 women) were included in the analysis as they had both CT and ^{18}F-FDG PET/CT at baseline and at the first restaging. Patients were treated with a standard schedule of nivolumab ($n = 19$), pembrolizumab ($n = 14$), and nivolumab/ipilimumab ($n = 2$). Twelve patients (34.3%) presented at diagnosis with advanced metastatic NSCLC, whereas the other 23 patients (65.7%) were treated with one or more anticancer therapies. The median number of immunotherapy cycles was 9 (range, 2–47).

The clinical characteristics of patients are summarized in **Table 1**.

Response Comparison for imRECIST and Standard Metabolic Criteria (EORTC, PERCIST)

In our study, all cases of PD at first evaluation according to RECIST 1.1 were all confirmed after at least 4 weeks according to imRECIST. As response rates between RECIST 1.1 and imRECIST were comparable, we used only the latter for our analysis. Classification between imRECIST and EORTC

TABLE 1 | Patient characteristics.

	N (%)
Age median (range)	75 (51–86)
Gender	
Male	23 (65.7)
Female	12 (34.3)
Smoking history	
Former/current	31 (88.6)
Never	4 (11.4)
Performance status	
0	19 (54.3)
≥1	16 (45.7)
Line of treatment	
0	12 (34.3)
1	12 (34.3)
≥2	11 (31.4)
Histology	
Adenocarcinoma	25 (71.4)
Squamous cell carcinoma	6 (17.1)
Other	4 (11.5)
Tumor PD-L1 expression level	
Positive	15 (42.9)
Negative	8 (22.9)
Indeterminate or missing	12 (34.2)

TABLE 2A | Comparison between imRECIST and metabolic criteria (EORTC).

imRECIST	EORTC				
	CMR	**PMR**	**SMD**	**PMD**	**Total**
CR	0	0	0	0	0
PR	1	2	0	1	4
SD	1	5	3	6	15
PD	0	1	0	15	16
Total	2	8	3	22	35

TABLE 2B | Comparison between imRECIST and metabolic criteria (PERCIST).

imRECIST	PERCIST				
	CMR	**PMR**	**SMD**	**PMD**	**Total**
CR	0	0	0	0	0
PR	1	2	0	1	4
SD	1	5	4	5	15
PD	0	1	1	14	16
Total	2	8	5	20	35

CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; CMR, complete metabolic response; PMR, partial metabolic response; SMD, stable metabolic disease; PMD, progressive metabolic disease.

TABLE 3A | Comparison between imRECIST and immuno-related metabolic criteria (imPERCIST).

imRECIST	imPERCIST				
	CMR	**PMR**	**SMD**	**PMD**	**Total**
CR	0	0	0	0	0
PR	1	2	0	1	4
SD	1	5	8	1	15
PD	0	2	9	5	16
Total	2	9	17	7	35

TABLE 3B | Comparison between imRECIST and immuno-related metabolic criteria (PERCIMT).

imRECIST	PERCIMT			
	CMR	**SMD**	**PMD**	**Total**
CR	0	0	0	0
PR	1	3	0	4
SD	1	11	3	15
PD	0	4	12	16
Total	2	18	15	35

CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; CMR, complete metabolic response; PMR, partial metabolic response; SMD, stable metabolic disease; PMD, progressive metabolic disease.

criteria was concordant in 20 patients (57.1%) with a moderate agreement between the two assessments ($\kappa = 0.340$, **Table 2A**). In particular, the change of response category was most frequently seen in patients classified as SD by imRECIST criteria: of 15 patients with SD, 6 (40%) were reclassified to CMR/PMR as the decrease in the sum of the diameters of the target lesions was <30%, while the decrease in the sum of SUVmax was more than 25%, whereas another 6 (40%) were reclassified to PMD by EORTC as new lesions were detected on ^{18}F-FDG PET/CT. Similar levels of agreement were obtained comparing imRECIST and PERCIST criteria ($\kappa = 0.342$, **Table 2B**).

Response Comparison for imRECIST and Immune-Related Metabolic Criteria (imPERCIST, PERCIMT)

imRECIST and imPERCIST were discordant in 20 patients (57.1%) with low agreement in the response classification between the two assessments ($\kappa = 0.143$, **Table 3A**). When adopting imPERCIST criteria, tumor responses were upgraded in 2 (10%) patients and downgraded in 18 (90%) patients. Notably, of 16 patients classified as PD according to imRECIST, 9 were reclassified as SMD according to imPERCIST, as the increase in the sum of the longest diameters of the target lesions was more than 20%, while the increase of overall SULpeak was <30%, and 2 patients as PMR because SULpeak reduction was >30%. Furthermore, of 15 patients classified as SD according to imRECIST, 1 was classified as PMD according to imPERCIST, as new lesions were detected on PET/CT contributing to summed SULpeak for PMD, but not on CT, and 6 patients as PMR, as the decrease in the sum of the longest diameters of the target lesions was <30%, while the decrease in the SULpeak was more than 30%. On the other hand, the level of agreement was higher, although moderate, when comparing imRECIST with PERCIMT ($\kappa = 0.413$), with concordance in 23 patients (65.7%). Overall, 9 patients were downgraded and 3 upgraded (**Table 3B**).

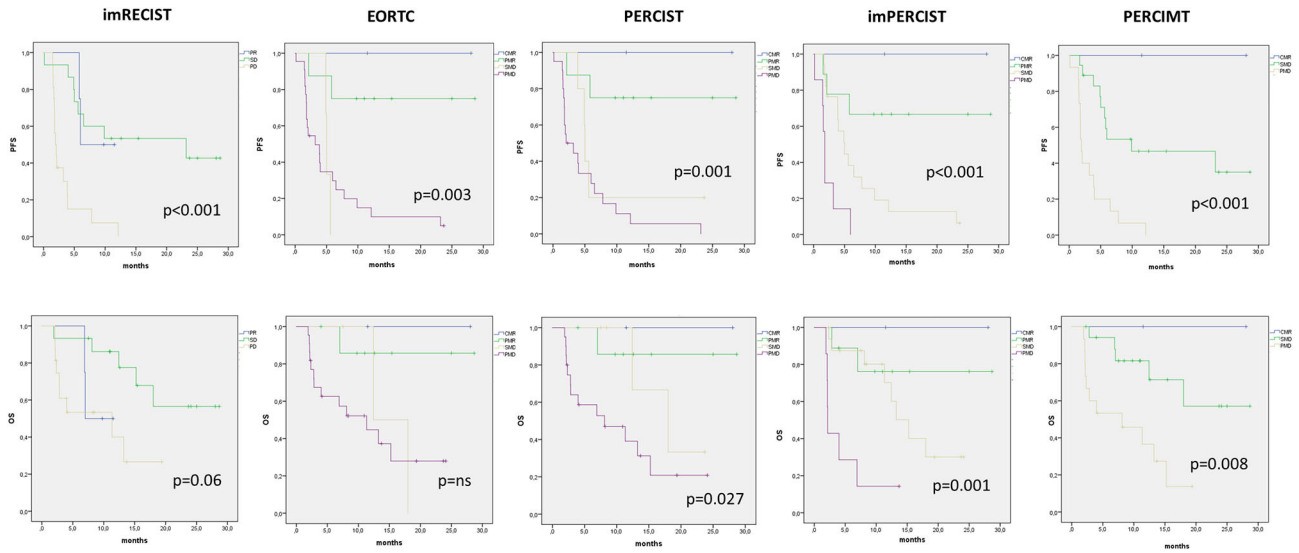


FIGURE 2 | Kaplan–Meier curves with log-rank (Mantel–Cox) test obtained for PFS and OS according to the different morphological and metabolic response criteria.

Clinical Outcome and Prognosis

The median duration of follow-up was 13.7 months (range, 2–28.7 months). Median PFS and OS for all patients was 5.6 months (95% CI, 3.1–8.1 months) and 15.3 months (95% CI, 8.9–21.6 months), respectively. With imRECIST, the median PFS was 6 months for patients with PR, 23 months in those with SD, and 2 months in those with PD. The median PFS in patients with PR was significantly longer than in those with PD ($p = 0.031$), but was not significantly longer than in those with SD (**Figure 2**). Among all metabolic parameters, PMD rate was comparable according EORTC, PERCIST, imPERCIST, and PERCIMT, with a median PFS of 3.2, 2.6, 1.8, and 1.9 months, respectively (**Figure 2**). However, while there was no statistical difference between SMD and PMD according to EORTC and PERCIST criteria, patients with SMD according imPERCIST and PERCIMT had longer PFS than those with PMD ($p = 0.004$ and $p < 0.001$, respectively). At the time of analysis, 16 patients (45.7%) had died. OS curve according to EORTC criteria was not significant (**Figure 3**) and showed only a tendency for imRECIST criteria ($p = 0.06$) (**Figure 2**). On the other hand, PERCIST, imPERCIST, and PERCIMT were significantly associated with OS ($p = 0.027$, $p = 0.001$, and $p = 0.008$, respectively), with similar survival for CMR/PMR group (median not reached) (**Figure 2**). Moreover, OS between SMD and PMD according to imPERCIST and PERCIMT was statistically significant ($p = 0.002$ and $p = 0.006$, respectively), whereas according to PERCIST, it was not.

We then analyzed the value of immune-metabolic criteria, i.e., imPERCIST and PERCIMT, in patients showing SD on CT. In these patients, both imPERCIST and PERCIMT well-differentiate patients with longer survival, expressed by both PFS ($p < 0.001$ and $p = 0.009$, respectively) and OS ($p < 0.001$ and $p = 0.002$, respectively) (**Figure 3**).

Finally, we also performed a multivariate analysis including all clinical variables and response criteria which were significant at univariate Cox proportional-hazards model. According to our results, performance status (HR = 0.278, $p = 0.015$), imRECIST (HR = 3.799, $p = 0.026$), and imPERCIST (HR = 4.064, $p = 0.014$) were predictive factors for PFS, while only performance status (HR = 0.327, $p = 0.035$) and imPERCIST (HR = 3.247, $p = 0.007$) were predictive for OS (**Table 4**).

DISCUSSION

Immunotherapy with ICI has introduced new challenges for medical imaging. This involves anatomical imaging, such as CT or MRI, as well as functional imaging, expressed by nuclear medicine techniques. In fact, with the growing use of ICI, atypical response patterns have been detected and described, such as pseudo-progression, hyper-progression, and dissociated response (14, 18). With this regard, one of the primary goals for the medical community is the early identification of patients who will not respond to ICI to permit a rapid switch of therapeutic line, to reduce the risk of immune-related adverse events, and to decrease the economic impact of these drugs, which remain very expensive. At the same time, it is important to avoid the premature treatment withdrawal for patients with therapeutic benefit (19, 20). For these reasons, many different immune-related scales have been proposed in the last years, but none of them has been routinely adopted in clinical practice, hence the debate is still open (21). Furthermore, only few studies have been published providing a direct comparison between CT-based and PET-based criteria, most with small cohorts and in melanoma setting (22–25). Only Rossi et al. (26) have recently compared anatomic and metabolic criteria in NSCLC patients treated with nivolumab.

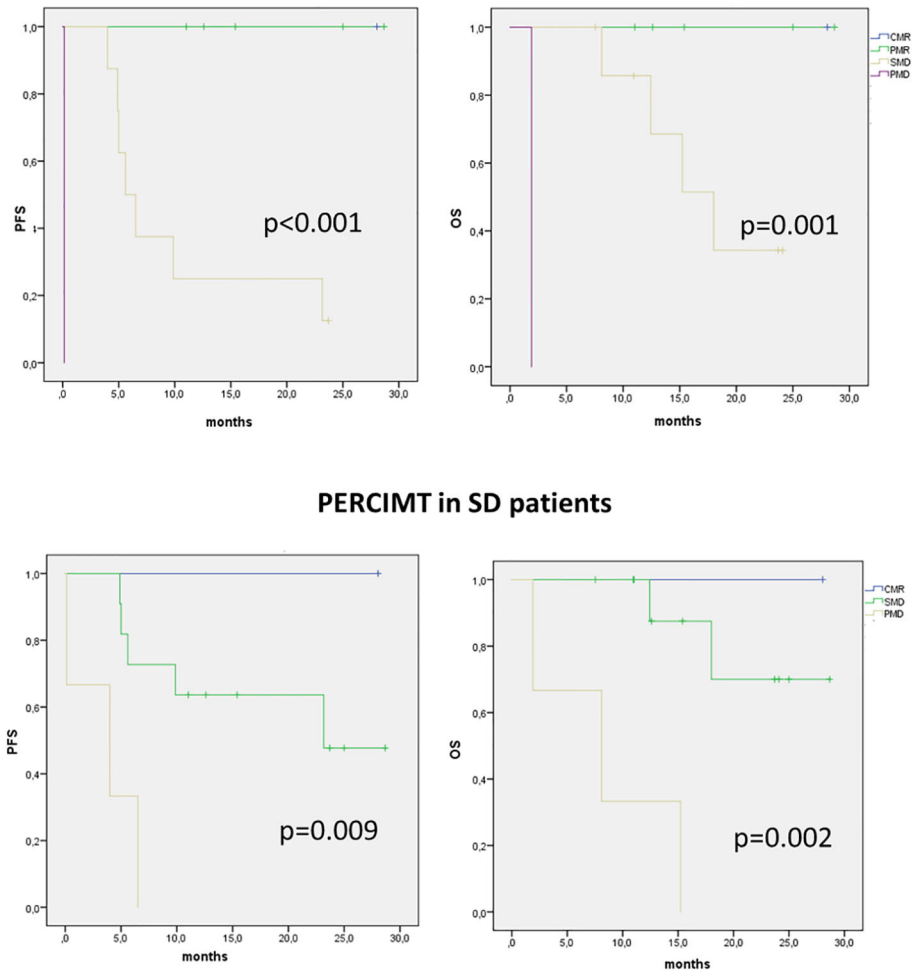


FIGURE 3 | Kaplan–Meier curves with log-rank (Mantel–Cox) test obtained for PFS and OS in patients presenting with SD on imRECIST and classified according to immune-related metabolic criteria (imPERCIST, PERCIMT).

In our study, we aimed to compare different response criteria in patients with advanced NSCLC treated with anti-PD-1/PD-L1 at first evaluation, after approximately 8 weeks. We adopted standard anatomic criteria as RECIST 1.1 and its immune-variation imRECIST, the latter combining cut-off values and unidimensional size of RECIST 1.1 and irRC criteria for interpretation of new lesions. Indeed, the main caveat is that irRC criteria require bidimensional measurements of tumor lesions hardly to apply in routine. Moreover, along with standard EORTC and PERCIST criteria, we also investigated imPERCIST and PERCIMT criteria. The latter can be considered the variation of irRC in which the metabolic dimensions of new lesions are embedded in the overall tumor burden (11).

In our study, we demonstrated a low overall agreement between imRECIST and imPERCIST, particularly for patients in the PD category. In fact, more than half of patients, i.e., 69%, classified as PD were downgraded to either SMD (9/16) or PMR (2/16) according to imPERCIST. Our results suggest that the sum of SULpeak from new lesions appears more reliable than diameter measurement, allowing to detect a therapeutic response as early as 8 weeks since ICI started. Hence, imPERCIST could help to avoid an early interruption of ICI therapy. On the other hand, our findings showed a moderate agreement between imRECIST and PERCIMT, with only 4 out of 16 patients with PD downgraded according to SMD, highlighting once again that metabolic activity expressed by SULpeak is optimal than metabolic measurement for new lesions. This is in line with a recent study in melanoma patients treated with ICI, where PERCIMT criteria were demonstrated suboptimal for the identification of disease progression (22). Furthermore, our results are apparently opposite to those of Rossi et al. (26), who compared different PET- and CT-based response criteria in a similar cohort. In fact, they demonstrated limited prognostic value of the SMD group who had a survival similar to

TABLE 4 | Univariate and multivariate Cox proportional-hazards regression analysis for prediction of PFS and OS.

Parameters	PFS			OS		
	Hazard ratio	95% CI	*P*-value	Hazard ratio	95% CI	*P*-value
Age (median)	0.935	0.423–2.067	ns	1.110	0.413–2.984	ns
Gender	0.519	0.231–1.164	ns	0.332	0.122–0.905	**0.031**
Smoking history	1.407	0.480–4.129	ns	2.668	0.747–9.545	ns
Histology	0.700	0.208–2.356	ns	1.701	0.583–4.962	ns
Performance status	0.479	0.216–0.998	**0.071**	0.289	0.102–0.815	**0.019**
imRECIST	3.962	1.826–8.596	**0.001**	1.893	0.786–4.560	ns
EORTC	2.330	1.360–3.994	**0.002**	2.445	1.116–5.359	**0.026**
PERCIST	2.572	1.483–4.460	**0.001**	3.020	1.289–7.078	**0.011**
imPERCIST	3.388	1.826–8.596	**0.001**	3.904	1.701–8.958	**0.001**
PERCIMT	3.321	1.571–7.019	**0.002**	4.157	1.477–11.706	**0.007**
Multivariate cox proportional-hazards regression analysis						
Gender	–	–	–	0.729	0.301–1.524	ns
Performance status	0.278	0.099–0.791	**0.015**	0.327	0.116–0.922	**0.035**
imRECIST	3.799	1.169–12.340	**0.026**	–	–	–
EORTC	0.730	0.078–6.800	ns	0.005	0.001–2.394	ns
PERCIST	0.890	0.086–9.256	ns	1.235	0.053–2.210	ns
imPERCIST	4.064	1.329–12.426	**0.014**	3.247	1.385–7.611	**0.007**
PERCIMT	1.742	0.418–7.258	ns	3.749	0.635–22.114	ns

ns, not significant; OS, overall survival; PFS, progression-free survival.

PMD patients. However, in their study, no significant difference between PERCIST and imPERCIST was found, whereas in our study among 20 patients classified as PMD according PERCIST, 13 were downgraded to SMD or PMR. In our opinion, this is the main reason for the different results obtained between our study and that from Rossi et al.

As shown in the Kaplan–Meier curves, we observed a positive impact of early PET and CT response on PFS, while only metabolic immune-related response criteria were prognostic for OS, confirming the role of ^{18}F-FDG PET/CT in predicting final clinical response to immunotherapy already underscored in previous studies in melanoma (24, 25). In fact, as visible on imRECIST survival curves, SD patients had a survival profile similar to PR curve. When selecting only patients with SD by imRECIST, both imPERCIST and PERCIMT criteria were able to identify three further survival curves (**Figure 3**). This evidence supports the hypothesis that SD group comprises a heterogeneous cohort with different prognosis, some with clinical benefit and others without. Hence, from this perspective, immune-related response criteria could be useful for monitoring the efficacy of immunotherapy, by identifying responders vs. non-responders as well as by predicting clinical outcomes, as arisen from our multivariate analysis.

Nevertheless, our study presents some limitations. The main one is related to its relatively small cohort. Furthermore, the use of different PET/CT scanners may have caused some variability in metabolic parameter measurement, although all patients were imaged in the same scans throughout the study. Third, in our study we have investigated the most validated ^{18}F-FDG PET-based criteria so far, whereas we did not consider other metabolic parameters or variables, such as metabolic tumor volume, total lesion glycolysis, circulating tumor cells, neutrophil-to-lymphocyte ratio, and their combination, which recently have been demonstrated to predict PFS and OS in patients treated with ICI (17, 27, 28).

In conclusion, our study encourages the use of immune-metabolic response criteria by ^{18}F-FDG PET/CT, in particular imPERCIST, to assess early response and to predict long-term outcomes in patients with NSCLC under ICI therapy. However, our first findings need to be validated in a larger prospective study.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available from the corresponding author, upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comitato Etico Humanitas. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AC: protocol development, data collection and management, data analysis, and manuscript writing.

EL: project development, data collection and management, and manuscript editing. SR and LT: protocol development and data collection and management. All authors: contributed to the article and approved the submitted version.

FUNDING

This study was supported by Fondazione AIRC (Associazione Italiana per la Ricerca sul Cancro) with grant no. 18923.

ACKNOWLEDGMENTS

The Italian Association for Research on Cancer (AIRC—Associazione Italiana per la Ricerca sul Cancro) is acknowledged for supporting this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01090/full#supplementary-material

REFERENCES

1. Egen JG, Kuhns, MS, Allison JP. CTLA-4: New insights into its biological function and use in tumor immunotherapy. *Nat. Immunol.* (2002) 3:611–18. doi: 10.1038/ni0702-611
2. Okazaki T, Chikuma S, Iwai Y, Fagarasan S, Honjo T. A rheostat for immune responses: the unique properties of PD-1 and their advantages for clinical application. *Nat. Immunol.* (2013) 14:1212–18. doi: 10.1038/ni.2762
3. Pons-Tostivint E, Latouche A, Vaflard P, Ricci F, Loirat D, Hescot S, et al. Comparative analysis of durable responses on immune checkpoint inhibitors versus other systemic therapies: a pooled analysis of phase III trials. *JCO Precis Oncol.* (2019) 3:1–10. doi: 10.1200/PO.18.00114
4. Li X, Wenes M, Romero P, Huang SC, Fendt SM, Ho PC. Navigating metabolic pathways to enhance antitumour immunity and immunotherapy. *Nat Rev Clin Oncol.* (2019) 16:425–41. doi: 10.1038/s41571-019-0203-7
5. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell.* (2015) 162:1229–41. doi: 10.1016/j.cell.2015.08.016
6. Castello A, Lopci E. Update on tumor metabolism and patterns of response to immunotherapy. *Q J Nucl Med Mol Imaging.* (2020) 64:175–85. doi: 10.23736/S1824-4785.20.03251-3
7. Decazes P, Bohn P. Immunotherapy by immune checkpoint inhibitors and nuclear medicine imaging: current and future applications. *Cancers.* (2020) 12:371. doi: 10.3390/cancers12020371
8. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* (2009) 45:228–47. doi: 10.1016/j.ejca.2008.10.026
9. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbé C, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune related response criteria. *Clin Cancer Res.* (2009) 15:7412–20. doi: 10.1158/1078-0432.CCR-09-1624
10. Hodi FS, Ballinger M, Lyons B, Soria J-C, Nishino M, Tabernero J, et al. Immune-modified response evaluation criteria in solid tumors (imRECIST): refining guidelines to assess the clinical benefit of cancer immunotherapy. *J Clin Oncol.* (2018) 36:850–8. doi: 10.1200/JCO.2017.75.1644
11. Anwar H, Sachpekidis C, Winkler J, Kopp-Schneider A, Haberkorn U, Hassel JC, et al. Absolute number of new lesions on (18)F-FDG PET/CT is more predictive of clinical response than SUV changes in metastatic melanoma patients receiving ipilimumab. *Eur J Nucl Med Mol Imaging.* (2018) 45:376–83. doi: 10.1007/s00259-017-3870-6
12. Ito K, Teng R, Schöder H, Humm JL, Ni A, Michaud L, et al. 18F-FDG PET/CT for monitoring of ipilimumab therapy in patients with metastatic melanoma. *J Nucl Med.* (2019) 60:335–41. doi: 10.2967/jnumed.118.213652
13. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med.* (2009) 50(Suppl. 1):122s—50s. doi: 10.2967/jnumed.108.057307
14. Castello A, Rossi S, Toschi L, Mazziotti E, Lopci E. Hyper-progressive disease in patients with non-small cell lung cancer treated with checkpoint inhibitors: the role of 18F-FDG PET/CT. *J Nucl Med.* (2019) 61:821–26. doi: 10.2967/jnumed.119.237768
15. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and (1999). EORTC recommendations. European organization for research and treatment of cancer (EORTC) PET study group. *Eur J Cancer.* (1999) 35:1773–82. doi: 10.1016/S0959-8049(99)00229-4
16. Kundel HL, Polansky M. Measurement of observer agreement. *Radiology.* (2003) 228:303–8. doi: 10.1148/radiol.2282011860
17. Castello A, Toschi L, Rossi S, Mazziotti E, Lopci E. The immune-metabolic-prognostic index and clinical outcomes in patients with non-small cell lung carcinoma under checkpoint inhibitors. *J Cancer Res Clin Oncol.* (2020) 146:1235–43. doi: 10.1007/s00432-020-03150-9
18. Borcoman E, Kanjanapan Y, Champiat S, Kato S, Servois V, Kurzrock R, et al. Novel patterns of response under immunotherapy. *Ann Oncol.* (2019) 30:385–96. doi: 10.1093/annonc/mdz003
19. Verma V, Sprave T, Haque W, Simone CB, Chang JY, Welsh JW, et al. A systematic review of the cost and cost-effectiveness studies of immune checkpoint inhibitors. *J Immunother Cancer.* (2018) 6:128. doi: 10.1186/s40425-018-0442-7
20. Anagnostou V, Yarchoan M, Hansen AR, Wang H, Verde F, Sharon E, et al. Immuno-oncology trial endpoints: capturing clinically meaningful activity. *Clin Cancer Res.* (2017) 23:4959–69. doi: 10.1158/1078-0432.CCR-16-3065
21. Laura E, Sepulcri M, Pasello G. PET/CT and the response to immunotherapy in lung cancer. *Curr Radiopharm.* (2019). doi: 10.2174/1874471013666191220105449. [Epub ahead of print].
22. Annovazzi A, Vari S, Giannarelli D, Pasqualoni R, Sciuto R, Carpano S, et al. Comparison of 18F-FDG PET/CT criteria for the prediction of therapy response and clinical outcome in patients with metastatic melanoma treated with ipilimumab and PD-1 inhibitors. *Clin Nucl Med.* (2020) 45:187–94. doi: 10.1097/RLU.0000000000002921
23. Cho SY, Lipson EJ, Im HJ, Rowe SP, Gonzalez EM, Blackford A, et al. Prediction of response to immune checkpoint inhibitor therapy using early-time-point 18F-FDG PET/CT imaging in patients with advanced melanoma. *J Nucl Med.* (2017) 58:1421–8. doi: 10.2967/jnumed.116.188839
24. Sachpekidis C, Larribere L, Pan L, Haberkorn U, Dimitrakopoulou-Strauss A, Hassel JC, et al. Predictive value of early 18F-FDG PET/CT studies for treatment response evaluation to ipilimumab in metastatic melanoma: preliminary results of an ongoing study. *Eur J Nucl Med Mol Imaging.* (2015) 42:386–96. doi: 10.1007/s00259-014-2944-y
25. Sachpekidis C, Anwar H, Winkler J, Kopp-Schneider A, Larribere L, Haberkorn U, et al. The role of interim (18)F-FDG PET/CT in prediction of response to ipilimumab treatment in metastatic melanoma. *Eur J Nucl Med Mol Imaging.* (2018) 45:1289–96. doi: 10.1007/s00259-018-3972-9
26. Rossi G, Bauckneht M, Genova C, Rijavec E, Biello F, Mennella S, et al. Comparison between 18F-FDG-PET- and CT-based criteria in non-small cell

lung cancer (NSCLC) patients treated with nivolumab. *J Nucl Med.* (2019) 61:990–8. doi: 10.2967/jnumed.119.233056

27. Seban RD, Mezquita L, Berenbaum A, Dercle L, Botticella A, Le Pechoux C et al. Baseline metabolic tumor burden on FDG PET/CT scans predicts outcome in advanced NSCLC patients treated with immune checkpoint inhibitors. *Eur J Nucl Med Mol Imaging.* (2019) 47:1147–57. doi: 10.1007/s00259-019-04615-x
28. Castello A, Carbone FG, Rossi S, Monterisi S, Federico D, Toschi L, et al. Circulating tumor cells and metabolic parameters in NSCLC patients treated with checkpoint inhibitors. *Cancers.* (2020) 12:487. doi: 10.3390/cancers12020487

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

OPINION
published: 19 August 2020
doi: 10.3389/fonc.2020.01359

Advanced Imaging of Biochemical Recurrent Prostate Cancer With PET, MRI, and Radiomics

Faiq Shaikh[1], Diana Dupont-Roettger[1], Jamshid Dehmeshki[1,2], Olga Kubassova[1] and Mohammed I. Quraishi[3]*

[1] Image Analysis Group, Philadelphia, PA, United States, [2] Faculty of Science, Engineering and Computing, Kingston University, Kingston-upon-Thames, United Kingdom, [3] Department of Radiology, University of Tennessee Medical Center, Knoxville, TN, United States

OPEN ACCESS

Edited by:
Laurent Dercle, Columbia University Irving Medical Center, United States

Reviewed by:
Yuming Jiang, Stanford University, United States
Michael Albert Thomas, University of California, Los Angeles, United States
Xiaohua Zhu, Huazhong University of Science and Technology, China
Cristina Messa, University of Milano Bicocca, Italy
Qijun Shen, Zhejiang University School of Medicine, China

***Correspondence:**
Faiq Shaikh
faiq.shaikh@ia-grp.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *02 April 2020*
Accepted: *29 June 2020*
Published: *19 August 2020*

Citation:
Shaikh F, Dupont-Roettger D, Dehmeshki J, Kubassova O and Quraishi MI (2020) Advanced Imaging of Biochemical Recurrent Prostate Cancer With PET, MRI, and Radiomics. Front. Oncol. 10:1359. doi: 10.3389/fonc.2020.01359

Keywords: PET, MRI, recurrent, prostate, cancer, biochemical

INTRODUCTION

Prostate cancer is a challenging disease for both physicians and patients. It requires a multidisciplinary team of urologists, medical oncologists, radiation oncologists, radiologists, and pathologists. Current management options include radical prostatectomy (RP), external beam therapy, brachytherapy, high-intensity focused ultrasound, cryotherapy, or watchful waiting (1). Although initial management of prostate cancer is difficult, there is even more uncertainty when patients have biochemical recurrence (BCR) prostate cancer (BCRPCa), which is described as a rise in prostate-specific antigen (PSA) levels in patients with prostate cancer who have undergone surgery or radiation (1). This is because with BCRPCa, the site of recurrence can be elusive. The multidisciplinary team needs the best data possible to ascertain treatment and management options, while the patient deserves answers on the state of his disease.

After radical prostatectomy, up to a third of patients will experience BCRPCa (1). BCRPCa has risen in recent years and now affects, by some estimates, 25,000 men annually in the United States (2). Spratt et al. (2) reason that this rise is largely due to the discouragement of routine PSA screening from the US Preventative Task Force, causing an increase of men presenting with high-risk localized cancer (2, 3). This trend has also been observed in Europe and was the impetus for the European Association of Urology (EAU) latest policy statement to reevaluate PSA screening (4, 5). In addition, there is $<$10% utilization of adjuvant radiation therapy despite support from the American Urological Association (AUA), American Society for Radiation Oncology (ASRO), and American Society of Clinical Oncology (ASCO) (2).

The definition of BCRPCa depends on the initial treatment strategy. Any strategy that does not remove all prostate epithelial tissue will demonstrate a nadir in PSA values instead of the expected undetectable PSA values seen with RP. The AUA as well as the EAU guidelines define BCR after RP as an initial PSA value of $\geq$0.2 ng/ml confirmed by subsequent PSA value of $\geq$0.2 ng/ml (1). To predict the probability of metastasis, BCR must be taken with clinical factors such as initial PSA level, Gleason score, pathological findings after surgery, and post-BCRPCa PSA kinetics.

After confirmation of BCRPCa, imaging is vital to supply the data needed by the multidisciplinary team to direct management. Imaging can change management in up to 70% of patients (1, 6). The determination of local salvage therapy, systemic therapy, surveillance, or the addition of androgen deprivation depends on confident detection (or the lack thereof) of recurrence and distinguishing between local recurrent and metastatic disease (7). It should be noted that a change in management does not necessarily translate to a change in morbidity or mortality.

Current National Comprehensive Cancer Network (NCCN) guidelines allow consideration of a multitude of imaging modalities (8). However, it is our opinion that the recommendations should be streamlined to the most effective imaging modalities available in answering the clinical question with the highest level of confidence available. The imaging studies with the highest positive rate at the lowest PSA can lead to early salvage radiation therapy.

CURRENT LANDSCAPE OF IMAGING IN BIOCHEMICAL RECURRENCE PROSTATE CANCER

Transrectal ultrasound (TRUS) can only evaluate the prostate bed and detects <50% of recurrence when PSA is <0.5 ng/ml (1). Computed tomography (CT) has poor anatomical resolution in the treated prostate bed, and unless recurrence is of substantial size, it is of limited use for local recurrence. CT can be helpful in evaluating for distant metastasis; however, CT has been reported to be positive in only 14% of cases (9). Any lesion seen on Tc-99m methyldiphosphonate (MDP) bone scintigraphy is highly non-specific. In fact, bone scintigraphy with BCRPCa has a positive rate of <5% when PSA is <7.0 ng/ml (10). The other obvious limitation of bone scintigraphy is that it cannot detect soft tissue recurrence.

The benefit of PET/CT is that it combines functional data ascertained by the radiotracer with limited anatomical data from the CT portion. 18F-NaF PET/CT is a bone imaging study that detects areas of increased bone turnover similar to Tc-99m MDP, allowing it to detect osseous metastases (11). Although 18F-NaF PET/CT has been shown by Jadvar et al. (12) to outperform 18-FDG PET/CT in the detection of occult osseous metastases, it has a similar constraint as bone scintigraphy in that it is confined to detecting osseous recurrence where other modalities can detect both osseous and soft tissue recurrence. The true-positive detection rate for occult osseous metastases by 18F-NaF PET/CT is 16.2%, and the median PSA levels for positive vs. negative PET/CT scans is reported as 4.4 and 2.9 ng/ml, respectively (12). 18F-FDG PET/CT, making use of glucose metabolism with a radiolabeled glucose analog, has a low sensitivity for BCRPCa, with only 28% detection of recurrence when PSA is <1.5 ng/ml (1). 11C-choline leverages the function of choline in cell membranes and lipid biosynthesis. 18F- or 11C-choline PET/CT is only of utility when PSA is >2.0 ng/ml (1). It has been observed that when PSA is <0.4 ng/ml, 11C-choline PET shows a dismal positive rate of only 21% (2). 18F-fluciclovine is a leucine amino acid analog and a novel PET radiotracer recently Food and Drug Administration (FDA) approved for use. Prostate cancer upregulates amino acid metabolism, giving 18F-fluciclovine its effectiveness as a radiotracer. At low PSA levels, it has a substantial positive detection rate. At PSA values of <1.0 ng/ml, 1.0–2.0 ng/ml, and ≥2.0 ng/ml, detection rates are reported as 72.0, 83.3, and 100%, respectively (13). Additionally, Lovec et al. (14) reported a positive rate above 50% with men with PSA values below or equal to 0.3 ng/ml. Although the NCCN guidelines report only a marginally better sensitivity and specificity range for 18F-fluciclovine compared to 11C-choline, studies comparing them head-to-head have shown that 18F-fluciclovine is superior (8, 15). Furthermore, Nanni et al. (15) reported the true positives at all PSA levels were generally higher with 18F-fluciclovine than 11C-choline.

Multi-parametric magnetic resonance imaging (mpMRI) is highly sensitive for local recurrence with its superior anatomic and tissue resolution. A positive rate of up to 94% has been reported with median PSA of 0.59 ng/ml (1). With respect to its application in prostate cancer imaging, mpMRI sequences involve various advanced sequences. The two most important sequences include diffusion-weighted imaging (DWI), which measures Brownian motion of water molecules within a voxel of tissue, and dynamic contrast enhancement (DCE) T1 imaging, which highlights vascular perfusion to tissue. DWI signal may be degraded secondary to the blooming artifact caused by surgical metallic clips or retained rectal air (16). Additionally, with short tau inversion recovery (STIR) imaging and DCE T1 imaging, osseous lesions are readily detected. In fact, MRI can detect changes in bone marrow prior to osteoblastic response which is needed for other types of bone-specific imaging (17). Post-therapy scar and fibrosis either does not enhance or demonstrates late enhancement. Malignancy, however, demonstrates early enhancement (18). The added benefit of mpMRI is that it can tease out local disease from focal treatment change that often occurs from focal therapies such as cryoablation and high-intensity focused ultrasound (18). Diagnostic CT or the CT portion of a PET/CT cannot provide the same level of anatomical detail of the treatment-altered prostate bed as mpMRI of the prostate.

In patients with BCRPCa, it is imperative to deliver salvage radiation therapy (RT) as early as possible (ideally PSA <0.5 ng/ml). This means that finding recurrence with the lowest possible PSA is invaluable. Of the imaging modalities available, the ones that detect disease with the lowest PSA value are 18F-fluciclovine PET/CT and mpMRI. 18F-fluciclovine is effective in detecting both local recurrence and distant metastatic disease, while mpMRI has very high utility in detecting local recurrence. In fact, a whole-body MRI would obviate the need for bone-specific imaging modalities given its superiority to both bone scintigraphy and 18F-NaF PET/CT (17). Hence, it is our opinion that there is no need for any other imaging modality except 18F-fluciclovine PET/CT combined with mpMRI, including a whole-body sequence, for BCRPCa, and ideally, 18F-fluciclovine PET/MRI, if available, for the added benefit of superior osseous detection (**Figure 1**). This approach will give the multidisciplinary team the structural and functional information to make early management decisions with high confidence.

FUTURE DIRECTIONS IN BIOCHEMICAL RECURRENCE PROSTATE CANCER IMAGING

Molecular imaging approaches applied in the management of BCRPCa management include prostate-specific membrane

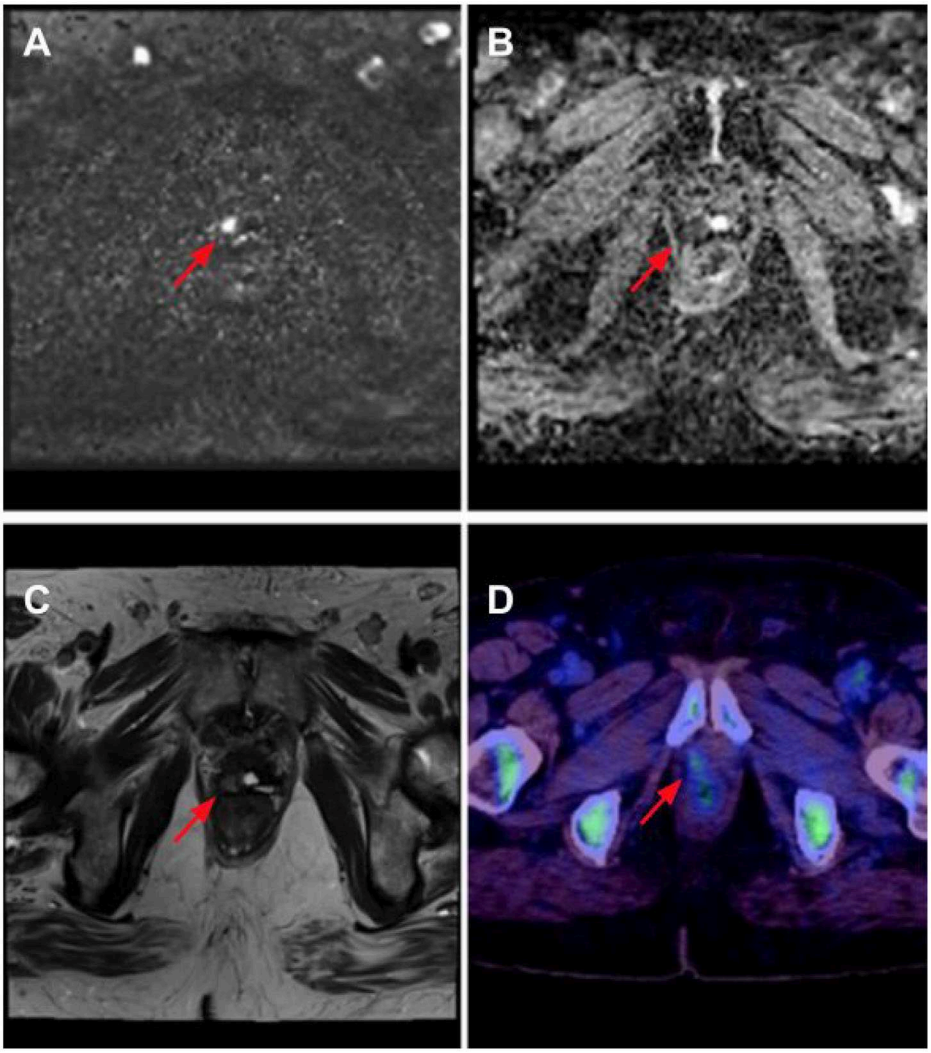


FIGURE 1 | Right anterior prostate bed recurrence as seen on multi-parametric MRI (mpMRI) with 18F-fluciclovine PET/CT. There is diffusion signal on calculated b-1400 diffusion-weighted imaging (DWI) (red arrow) **(A)** with corresponding low apparent diffusion coefficient (ADC) values (red arrow). **(B)** Anatomical correlation is noted on T2 Half-Fourier acquisition single-shot turbo spin echo (HASTE) imaging (red arrow). **(C)** Lesion is confirmed to contain upregulated amino acid transport, seen in prostate cancer, in the 18F-fluciclovine image (red arrow) **(D)**.

antigen (PSMA) radiotracers bound either to gallium (68Ga-PSMA) or to fluoride (18F-DCFPyL). PSMA is a membrane glycoprotein that is overexpressed by prostate cancer cells. Ga-PSMA PET is currently undergoing Phase III trials in the US and appears to outperform 18F-fluciclovine with a positive rate of 73% at a PSA range as low as 0.5 to 1.0 ng/ml and a positive rate of >50% at the remarkably low PSA range of 0.20–0.29 ng/ml (1). It should be noted that 68Ga-PSMA is already clinically available in Europe and outperforms 18F-fluciclovine (19). 18F-DCFPy is a PSMA radiotracer that produces images with higher resolution and is currently in phase II trials (2). It has been shown to successfully identify recurrent disease and lead to a change in management in 60% of patients and in up to 28% of patients who had negative CT or MR findings (20). It has been shown to detect bone metastases as accurately as 18F-NaF PET/CT but is superior to the latter given its ability to detect non-osseous disease at low PSA values, making it a more useful study overall (21).

BCRPCa as well as primary prostate cancer is ripe for quantitative imaging biomarker development using radiomics as a methodology. Radiomics may be defined as a process of extracting quantified data from medical images as single-order (histogram-based) and second-order (texture analysis-based) features, which are then classified into clusters (or signatures) that best align with an underlying pathophysiologic process (**Figure 2**). Radiomic analysis performed on pretreatment mpMRI has been shown to predict BCRPCa, which has implications for predicting response to adjuvant therapy (22, 23). In addition, radiomic texture analysis has been shown to

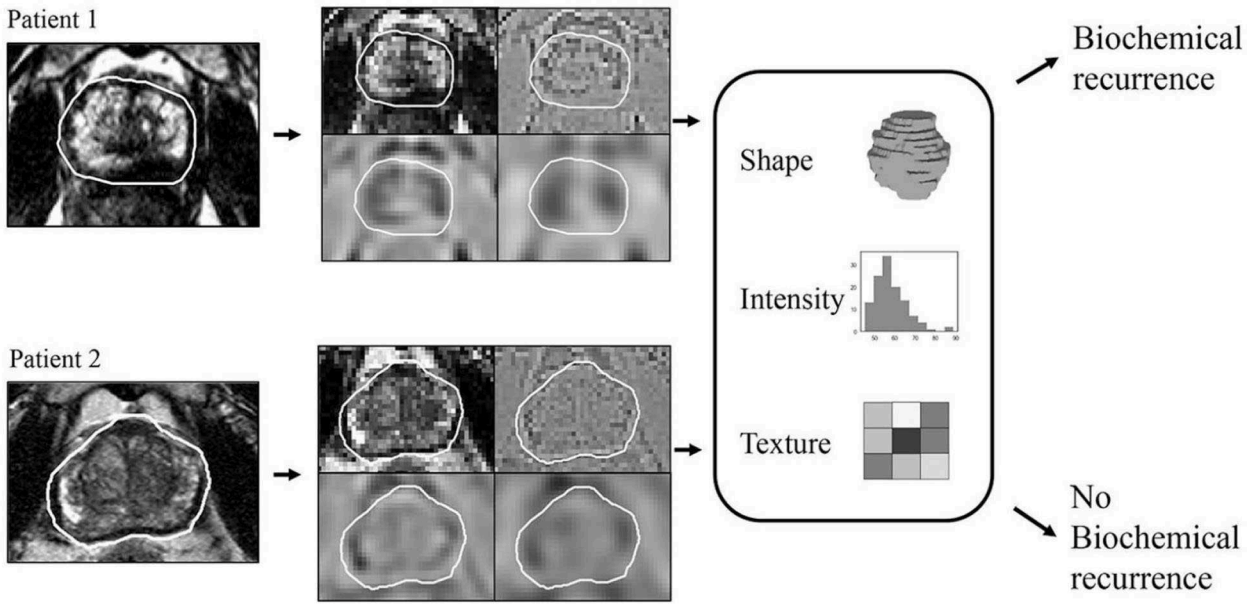


FIGURE 2 | Graphical schema of the radiomics process that involves lesion identification, drawing regions of interest, image preprocessing followed by radiomic feature extraction and classification that provides the imaging biomarker for predicting biochemical recurrence. Reused from Fernandes et al. (28) under the Creative Commons License.

predict biochemical relapse as well as BCRPCa-free survival after prostatectomy [area under the curve (AUC) 0.76] (24). Furthermore, MR radiomic signatures [using T2W and apparent diffusion coefficient (ADC) images] can accurately predict the response to carbon ion radiotherapy (CIRT) for prostate cancer as well (25). Recently, radiomics has been shown to predict Decipher score (an mRNA-based genomic test that predicts the occurrence of prostate cancer metastasis after radical prostatectomy) by differentiating between low and intermediate/high scores (with an AUC of 0.92) (26, 27).

Imaging is central to BCRPCa treatment decisions. Current practice in the US should be reformed to use 18F-fluciclovine and moving to a PSMA-based radiotracer as currently approved in Europe once FDA approved in the USA in conjunction with mpMRI or as PET/MR where available. The future is bright in the fight against BCRPCa with growing research in imaging-based precision medicine practices including radiomics-based imaging biomarkers.

AUTHOR CONTRIBUTIONS

FS contributed to manuscript writing, focusing on PET/CT and radiomics. DD-R, JD, and OK contributed to manuscript writing and provided technical input. MQ contributed to manuscript writing focusing on clinical, PET/CT, and MRI. All authors contributed to the article and approved the submitted version.

REFERENCES

1. McCormick BZ, Mahmoud AM, Williams SB, Davis JW. Biochemical recurrence after radical prostatectomy: current status of its use as a treatment endpoint and early management strategies. *Indian J Urol.* (2019) 35:6–17. doi: 10.4103/iju.IJU_355_18
2. Spratt DE, McHugh DJ, Morris MJ, Morgans AK. Management of biochemically recurrent prostate cancer: ensuring the right treatment of the right patient at the right time. *Am Soc Clin Oncol Educ Book.* (2018) 38:355–62. doi: 10.1200/EDBK_200319
3. Jemal A, Fedewa SA, Ma J, Siegel R, Lin CC, Brawley O, et al. Prostate cancer incidence and PSA testing patterns in relation to USPSTF screening recommendations. *JAMA.* (2015) 314:2054–61. doi: 10.1001/jama.2015.14905
4. Gandaglia G, Albers P, Abrahamsson PA, Briganti A, Catto JW, Chapple CR, et al. Structured population-based prostate-specific antigen screening for prostate cancer: the European Association of Urology position in (2019). *Eur Urol.* (2019) 76:142–50. doi: 10.1016/j.eururo.2019.04.033
5. *Prostate Cancer: Britain's Growing Problem.* ORCHID - Bayer (2018). Available online at: https://orchid-cancer.org.uk/wp-content/uploads/2018/04/April-2018-Report-Spread-2.pdf (accessed February 2, 2020).
6. De Bari B, Mazzola R, Aiello D, Fersino S, Gregucci F, Alongi P, et al. Could 68-Ga PSMA PET/CT become a new tool in the decision-making strategy

of prostate cancer patients with biochemical recurrence of PSA after radical prostatectomy? A preliminary, monocentric series. *La Radiol Med.* (2018) 123:719–25. doi: 10.1007/s11547-018-0890-7
7. Stephenson AJ, Kattan MW, Eastham JA, Dotan ZA, Bianco Jr FJ, Lilja H, et al. Defining biochemical recurrence of prostate cancer after radical prostatectomy: a proposal for a standardized definition. *J Clin Oncol.* (2006) 24:3973–8. doi: 10.1200/JCO.2005.04.0756
8. N.C.C.N. *Prostate Cancer (Version 1.2020).* Available online at: https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf (accessed February 2, 2020).
9. Kane CJ, Amling CL, Johnstone PA, Pak N, Lance RS, Thrasher JB, et al. Limited value of bone scintigraphy and computed tomography in assessing biochemical failure after radical prostatectomy. *Urology.* (2003) 61:607–11. doi: 10.1016/S0090-4295(02)02411-1
10. Artibani W, Porcaro AB, De Marco V, Cerruto MA, Siracusano S. Management of biochemical recurrence after primary curative treatment for prostate cancer: a review. *Urol Int.* (2018) 100:251–62. doi: 10.1159/000481438
11. Langsteger W, Rezaee A, Pirich C, Beheshti M. 18F-NaF-PET/CT and 99mTc-MDP bone scintigraphy in the detection of bone metastases in prostate cancer. *Semin Nucl Med.* (2016). 46:491–501. doi: 10.1053/j.semnuclmed.2016.07.003
12. Jadvar H, Desai B, Ji L, Conti PS, Dorff TB, Groshen SG, et al. Prospective evaluation of 18F-NaF and 18F-FDG PET/CT in detection of occult metastatic disease in biochemical recurrence of prostate cancer. *Clin Nucl Med.* (2012) 37:637–43. doi: 10.1097/RLU.0b013e318252d829
13. Akin-Akintayo OO, Jani AB, Odewole O, Tade FI, Nieh PT, Master VA, et al. Change in salvage radiotherapy management based on guidance with FACBC (fluciclovine) PET-CT in post-prostatectomy recurrent prostate cancer. *Clin Nucl Med.* (2017) 42:e22. doi: 10.1097/RLU.0000000000001379
14. Lovec P, Savir-Baruch B, Gupta G, Wagner R, Gabriel MS, Harkenrider MM, et al. Positive findings on 18F-fluciclovine PET/CT in patients with suspected recurrent prostate cancer and PSA levels$\leq$ 0.5 and$\leq$ 0.3 ng/ml. *Int J Radiat Oncol Biol Phys.* (2018) 102:S161. doi: 10.1016/j.ijrobp.2018.07.015
15. Nanni C, Zanoni L, Pultrone C, Schiavina R, Brunocilla E, Lodi F, et al. 18 F-FACBC (anti1-amino-3-18 F-fluorocyclobutane-1-carboxylic acid) versus 11 C-choline PET/CT in prostate cancer relapse: results of a prospective trial. *Eur J Nucl Med Mol Imaging.* (2016). 43:1601–10. doi: 10.1007/s00259-016-3329-1
16. van Griethuysen JJ, Bus EM, Hauptmann M, Lahaye MJ, Maas M, ter Beek LC, et al. Gas-induced susceptibility artefacts on diffusion-weighted MRI of the rectum at 1.5 T–effect of applying a micro-enema to improve image quality. *Eur J Radiol.* (2018) 99:131–7. doi: 10.1016/j.ejrad.2017.12.020
17. Messiou C, Cook G, Desouza NM. Imaging metastatic bone disease from carcinoma of the prostate. *Br J Cancer.* (2009) 101:1225–32. doi: 10.1038/sj.bjc.6605334
18. Mertan FV, Greer MD, Borofsky S, Kabakus IM, Merino MJ, Wood BJ, et al. Multi-parametric magnetic resonance imaging of recurrent prostate cancer. *Top Magnet Reson Imaging.* (2016) 25:139–47. doi: 10.1097/RMR.0000000000000088
19. Calais J, Ceci F, Nguyen K, Gartmann J, Eiber M, Reiter RE, et al. Prospective head-to-head comparison of 18F-fluciclovine and 68Ga-PSMA-11 PET/CT for localization of prostate cancer biochemical recurrence after primary prostatectomy. *J Clin Oncol.* (2019) 37:15. doi: 10.1200/JCO.2019.37.7_suppl.15
20. Song H, Harrison C, Duan H, Guja K, Hatami N, Franc BL, et al. Prospective evaluation of 18F-DCFPyL PET/CT in biochemically recurrent prostate cancer in an academic center: a focus on disease localization and changes in management. *J Nucl Med.* (2020) 61:546–51. doi: 10.2967/jnumed.119.231654
21. Duan H, Song H, Baratto L, Khalaf M, Hatami N, Franc B, et al. Prospective comparison of ^{18}F-DCFPyL PET/CT with ^{18}F-NaF PET/CT for detection of skeletal metastases in biochemically recurrent prostate cancer. *J Nucl Med.* (2019) 60(Suppl. 1):1584. doi: 10.2967/jnumed.119.231654
22. Shiradkar R, Ghose S, Jambor I, Taimen P, Ettala O, Purysko AS, et al. Radiomic features from pretreatment biparametric MRI predict prostate cancer biochemical recurrence: preliminary findings. *J Magnet Res Imaging.* (2018) 48:1626–36. doi: 10.1002/jmri.26178
23. Bourbonne V, Vallières M, Lucia F, Doucet L, Visvikis D, Tissot V, et al. Validation of an MRI-derived radiomics model to guide patients selection for adjuvant radiotherapy after prostatectomy for high-risk prostate cancer. *Int J Radiat Oncol Biol Phys.* (2019) 105. E266–7. doi: 10.1016/j.ijrobp.2019.06.1879
24. Bourbonne V, Vallières M, Lucia F, Doucet L, Visvikis D, Tissot V, et al. MRI-derived radiomics to guide post-operative management for high-risk prostate cancer. *Front Oncol.* (2019) 9:807. doi: 10.3389/fonc.2019.00807
25. Wu S, Jiao Y, Zhang Y, Ren X, Li P, Yu Q, et al. Imaging-based individualized response prediction of carbon ion radiotherapy for prostate cancer patients. *Cancer Manage Res.* (2019) 11:9121–31. doi: 10.2147/CMAR.S214020
26. Li L, Shiradkar R, Leo P, Purysko A, Algohary A, Klein EA, et al. Association of radiomic features from prostate bi-parametric MRI with Decipher risk categories to predict risk for biochemical recurrence post-prostatectomy. *J Clin Oncol.* (2019) 37(Suppl. 15):e16561–e16561. doi: 10.1200/JCO.2019.37.15_suppl.e16561
27. Wang J, Wu CJ, Bao ML, Zhang J, Wang XN, Zhang YD. Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer. *Eur Radiol.* (2017) 27:4082–90. doi: 10.1007/s00330-017-4800-5
28. Fernandes CD, Dinh CV, Walraven I, Heijmink SW, Smolic M, van Griethuysen JJ, et al. Biochemical recurrence prediction after radiotherapy for prostate cancer with T2w magnetic resonance imaging radiomic features. *Phys Imaging Radiat oncol.* (2018) 7:9–15. doi: 10.1016/j.phro.2018.06.005

Conflict of Interest: FS, DD-R, JD, and OK are employees of Image Analysis Group, Philadelphia, PA, USA.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 04 September 2020
doi: 10.3389/fonc.2020.01634

Diagnostic Accuracy and Safety of Coaxial System in Oncology Patients Treated in a Specialist Cancer Center With Prospective Validation Within Clinical Trial Data

Khurum Khan[1], Reyes Gonzalez-Exposito[2], David Cunningham[2]*, Dow-Mu Koh[3], Andrew Woolston[4], Louise Barber[4], Beatrice Griffiths[4], Kyriakos Kouvelakis[2], Vanessa Calamai[2], Monia Bali[3], Nasir Khan[3], Annette Bryant[2], Claire Saffery[2], Charles Dearman[2], Ruwaida Begum[2], Sheela Rao[2], Naureen Starling[2], David Watkins[2], Ian Chau[2], Chiara Braconi[2], Nicola Valeri[2,5], Marco Gerlinger[2,4] and Nicos Fotiadis[3]**

[1] Department of Gastrointestinal Oncology, UCL Cancer Institute, University College NHS Foundation Trust, London, United Kingdom, [2] Department of Medicine, The Royal Marsden NHS Trust, London, United Kingdom, [3] Cancer Research UK Cancer Imaging Centre, Division of Radiotherapy and Imaging, The Institute of Cancer Research and Royal Marsden Hospital, London, United Kingdom, [4] Translational Oncogenomics Laboratory, Centre for Evolution and Cancer, The Institute of Cancer Research, London, United Kingdom, [5] Division of Molecular Pathology, The Institute of Cancer Research, London, United Kingdom

Background: Image-guided tissue biopsies are critically important in the diagnosis and management of cancer patients. High-yield samples are also vital for biomarker and resistance mechanism discovery through molecular/genomic analyses.

Patients and Methods: All consecutive patients who underwent plugged image-guided biopsy at Royal Marsden from June 2013 until September 2016 were included in the analysis. In the next step, a second cohort of patients prospectively treated within two clinical trials (PROSPECT-C and PROSPECT-R) were assessed for the DNA yield from biopsies assessed for complex genomic analysis.

Results: A total of 522 plugged core biopsies were performed in 457 patients [men, 52%; median age, 63 years (range, 17–93)]. Histological diagnosis was achieved in 501 of 522 (96%) performed biopsies. Age, gender, modality, metastatic site, and seniority of the interventionist were not found to be significant factors associated with odds of failure on a logistic regression. Seventeen (3.3%) were admitted due to biopsy-related complications; nine, three, two, one, one, and one were admitted for grade I/II pain control, sepsis, vasovagal syncope, thrombosis, hematuria, and deranged liver functions, respectively; two patients with right upper quadrant pain after liver biopsy were found to have radiologically confirmed subcapsular hematoma requiring conservative treatment. One patient (0.2%) developed grade III hemorrhage following biopsy of a gastric gastrointestinal stromal tumor (GIST). Overall molecular analysis was successful in 89% (197/222 biopsies). Prospective validation in 62 biopsies gave

OPEN ACCESS

Edited by:
Randy Yeh,
Memorial Sloan Kettering Cancer Center, United States

Reviewed by:
Andrea Zivi,
Imperial College London, United Kingdom
Awais Jalil,
Southend University Hospital NHS Foundation Trust, United Kingdom

***Correspondence:**
Khurum Khan
khurum.khan1@nhs.net
David Cunningham
David.Cunningham@rmh.nhs.uk
Nicos Fotiadis
nicos.fotiadis@rmh.nhs.uk

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *14 April 2020*
Accepted: *27 July 2020*
Published: *04 September 2020*

Citation:
Khan K, Gonzalez-Exposito R, Cunningham D, Koh D-M, Woolston A, Barber L, Griffiths B, Kouvelakis K, Calamai V, Bali M, Khan N, Bryant A, Saffery C, Dearman C, Begum R, Rao S, Starling N, Watkins D, Chau I, Braconi C, Valeri N, Gerlinger M and Fotiadis N (2020) Diagnostic Accuracy and Safety of Coaxial System in Oncology Patients Treated in a Specialist Cancer Center With Prospective Validation Within Clinical Trial Data. Front. Oncol. 10:1634. doi: 10.3389/fonc.2020.01634

success rates of 92.06 and 79.03% for DNA extraction of >1 μm and tmour content of >20%, respectively.

Conclusion: The probability of diagnostic success for complex molecular analysis is increased with plugged large coaxial needle biopsy technique, which also minimizes complications and reduces hospital stay. High-yield DNA acquisition allows genomic molecular characterization for personalized medicine.

Keywords: coaxial core-needle biopsy system, tissue biopsies, formalin-fixed paraffin-embedded, clinical trials, genomic analysis

INTRODUCTION

While cancer management and treatment options have significantly improved during the last few years, our knowledge and understanding about mechanisms of response, and/or resistance to anticancer therapies remain relatively sparse. To date, this relative lack of understanding is partially due to difficulties in accessing prospectively collected tissue and blood samples from systemic anticancer therapy (SACT)-resistant tumors.

Image-guided tissue biopsies are not just important in establishing an accurate histopathological diagnosis and standard cancer management; high-yield samples are also vital in understanding the molecular and genomic characteristics of tumors. Genomic analyses on tumor samples broadly fall into two categories including (1) targeted approaches investigating a limited number of genes that are known to influence clinical decision making and (2) whole exome or genome sequencing frequently adopted in exploratory research studies to learn about new mechanisms of response or resistance to SACT (1, 2). Conventional formalin-fixed paraffin-embedded (FFPE) samples obtained during diagnostic procedures may not be sufficient for such analyses to be realized. For instance, the data from The Cancer Genome Atlas (TCGA) studies showed that fresh frozen material from primary tumor resection specimens was associated with a tumor content of 60% (3). Moreover, using FFPE DNA for large-scale genomic studies may demonstrate mutations that have occurred as a result of the fixation process, which makes it difficult to distinguish real tumor variants from these fixation artifacts. Furthermore, low-quality fragmented DNA can fail quality control in the preanalytical stage impairing success rates.

While a number of retrospective studies have demonstrated the safety and accuracy of diagnostic biopsies (4–6), data interpretation from such studies has often been hampered by small numbers, the lack of information on yield for molecular/genomic characterization of tumors, and the lack of prospective validation. At Royal Marsden (RM), we have been using coaxial core-needle biopsy (CNB) system and a preformed gelatin sponge sealing device to conduct solid organ core biopsies in order to minimize the number of passes and reduce the risk of complications, respectively. We present here the largest dataset demonstrating the safety and accuracy of this approach. Moreover, we took the opportunity to utilize a cohort of patients from two prospective clinical trials to validate tumor yields from biopsies in these translational studies.

MATERIALS AND METHODS

Study Design

All consecutive patients who underwent plugged image-guided biopsy at RM from June 2013 until September 2016 were included in the analysis. Data including gender, age, primary tumor, biopsy site, needle gage, interval between biopsy and discharge, incidence of complications, and biopsy success were collected. The study was approved by the RM Institutional review board.

Biopsy Technique

The biopsies were performed by a Consultant Interventional Radiologist (IR) or an IR fellow under supervision. Ultrasound and CT guidance was used based on the location of the lesion. Conscious sedation was administered along with local anesthesia, when required, to maximize cooperation and improve patient experience (**Figure 1**). A 15- or 17-G coaxial needle was inserted under direct image guidance at the edge of the lesion and two to six cores were obtained with a 16- or 18-G automatic core-biopsy needle (True-Core II, Argon Medical Devices, Frisco, TX, United States), respectively. Different areas inside the lesion were sampled by changing the angle and position of the coaxial needle. After the samples were collected, one to four preformed 16- or 18-G gelatin foam pledgets (Hunter biopsy-sealing device, Vascular Solutions Inc., Minneapolis, MN, United States) were deployed through the coaxial along the tract of the needle to facilitate hemostasis. The gelatin resorbs completely within 12 weeks.

Validation Cohort

In the second step, a validation cohort of patients prospectively treated within two clinical trials was used to assess the DNA yield utilized for genomic analysis. The two trials included *PROSPECT-C* [clinical trials.gov number (NCT02994888)] (2, 7) and *PROSPECT-R* [clinical trials.gov number (NCT03010722)] (1); phase II, open label, non-randomized studies of antiepidermal growth factor receptor (anti-EGFR) monoclonal antibodies and regorafenib in patients with *RAS* wild type and *RAS* mutant refractory metastatic colorectal cancer (CRC), respectively. All participants in both studies were required

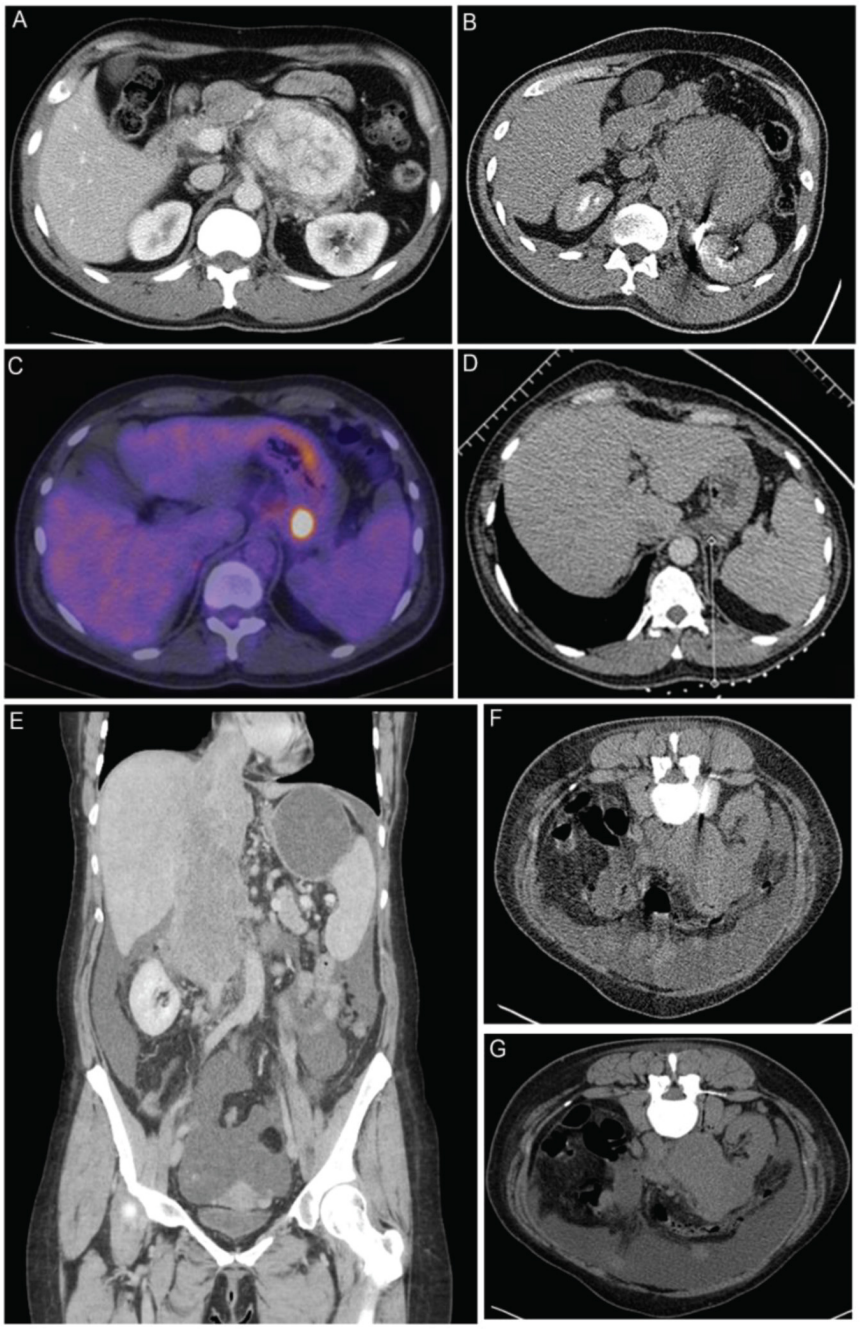


FIGURE 1 | Biopsy examples of patients within the study **(A)** A computed tomography (CT) of a patient with a highly vascular retroperitoneal mass thought to be too high risk to biopsy at the local hospital. Surgery was also considered to be high risk of R1/R2 resection, and a CT-guided biopsy was recommended by our multidisciplinary team (MDT). **(B)** Biopsy was performed with a 15-G/16-G coaxial needle. The tract was plugged with 16-G Hunter plugs, and there were no complications. The biopsy showed an inflammatory myofibroblastic tumor, which responded well on steroids and an operation was avoided. **(C)** PET/CT of a 57-year-old patient with relapsed Hodgkin's lymphoma after six cycles of ABVD chemotherapy. There was response in all sites of disease with the exception of a plaque of tissue behind the fundus of the stomach, which appear [18F]-fluorodeoxyglucose (FDG) avid on PET scan. A decision of the MDT was made to biopsy the lesion in order to exclude transformation of lymphoma. **(D)** The 17-G coaxial needle was placed medial to the left adrenal and above the splenic vessels adjacent to the lesion. Three cores were taken, and the tract was plugged. There were no complications. The biopsy showed Hodgkin's lymphoma, which responded well to systemic therapy and consolidation RT. **(E)** Coronal CT images of a 48-year-old patient with a large tumor of the inferior vena cava (IVC) extending from the level of the renal veins to the right atrium. Occluded hepatic veins and ascites can be seen on the scan. **(F)** The lesion was biopsied with a 15-G/16-G coaxial needle. **(G)** The tract was plugged with three gelfoam pledgets. A diagnosis of leiomyosarcoma of the IVC was made, and the procedure had no complications.

to have mandatory pretreatment biopsies (6 cores), biopsies at partial response in *PROSPECT-C* and stable disease at 2 months in *PROSPECT-R* (6 cores), and at the time of progression (6–12 cores from two suitable progressing metastatic sites).

Prospective Tissue Collection Procedures

Fresh frozen and FFPE tissue samples were obtained, and plasma collection was conducted as per the study protocols at the clinically relevant defined time points. Sixteen-gage core biopsy was used to collect three or four fresh biopsy specimens and one or two specimens fixed in formalin and paraffin embedded. Within the trials, approximately 25% of the total length of a core was detached for primary culture, and the remaining ∼75% of the core was snap frozen and used for genomic analysis. One core was transported to establish tumor-derived organoids and targeted panel validation (8). One core was used for genomic analysis after being placed into cryovials and immediately snap frozen in liquid nitrogen. The remaining two cores were placed straight into formalin and embedded in paraffin wax. Primary morphological and immunohistochemical analysis was performed by the histopathologist on the FFPE specimen for confirmation of diagnosis. The samples were then stored in the GI and Lymphoma Research Bank of the RM, anonymized by trial number and time point.

Tissue Sample Processing

Biopsy cores were snap frozen in liquid nitrogen at the time of collection. Genomic gDNA and mRNA were co-extracted from cores using the Qiagen All-Prep kit. DNA was also isolated from whole blood samples using the Qiagen QIAamp DNA Blood Mini kit (**Figure 2**).

Whole Exome Sequencing

A minimum of 500 ng of gDNA was prepared for whole exome sequencing (WES) using the Agilent SureSelect Human All Exon v5 capture library, according to the manufacturers' protocol. The resulting libraries were sequenced to a mean depth of 100× using paired-end 100 reads on an Illumina HiSeq 2500. High-quality reads were aligned to the National Center for Biotechnology information (NCBI) reference genome (hg19) using BWA (v0.7.12) and SAMtools (v0.1.19) to remove duplicates. Tumor content was estimated based on the CNVkit (v0.8.1) copy number profile.

Sanger Sequencing

For patients with a known tumour variant, PCR was performed on 20 nanograms of gDNA using M13F/R-tailed mutation specific primers (Life Technologies; **Supplementary Table 1**) and Q5 High-Fidelity 2 × Master Mix (NEB) on an Eppendorf Mastercycler Nexus GSX1. Primer-specific annealing temperatures for Q5 polymerase were established using the NEB online Tm calculator. PCR products were cleaned using Qiaquick PCR purification kit (Qiagen), and 15 ng DNA was submitted for M13F and M13R sequencing using the Mix2Seq service (Eurofins Genomics). Ab1 traces were visualized and compared to the reference sequence using ApE software[1]. Sample tumour content was estimated from the relative abundance of wild-type and variant peaks (**Figure 3** and **Supplementary Table 1**).

Statistical Design

The success rate of biopsies was determined by the ability to perform standard molecular testing on tissue specimens and safety determined by frequency of complications and extended hospital stay. Encrypted data were collected in a password-protected Excel file and statistical analysis performed using STATA13. Chi-squared analysis was undertaken to identify baseline characteristics that provided independent association with failure and success rates.

RESULTS

Overall Safety of Image-Guided Biopsies and Cox Regression Analysis

A total of 522 tissue biopsies were performed in 457 patients [men, 48%; median age, 63 years (range, 23–86)] (**Supplementary Table 2**). Two, three, and four biopsies were obtained from 51 (11.2%), 13 (2.8%), and 1 (0.2%) patients, respectively, at different time points as part of clinical trial protocols. Histological diagnosis was achieved in 501 of 522 (96%) performed biopsies. Same-day discharge was achieved for 444 (85.1%) procedures as outpatients, 35 (6.7%) and 17 (3.3%) had planned inpatient and elective procedures, respectively, and 8 (1.5%) patients were kept in for overnight observation after a late evening procedure. Seventeen (3.3%) were admitted with the following biopsy-related complications: grade I/II pain control (nine), sepsis (one), vasovagal syncope (two), thrombosis (one), hematuria (one), and deranged liver functions (one). Two patients with right upper quadrant pain had radiologically confirmed subcapsular hematoma requiring conservative treatment. One patient (0.2%) developed grade III hemorrhage requiring transfusion of 2 U of packed red blood cells following biopsy of a gastric gastrointestinal stromal tumor (GIST). In 21 of 522 biopsies, diagnosis was not achieved due to sampling error during needle placement. These were small lesions not well visualized with ultrasound and CT, and normal tissue adjacent to the lesion was consequently biopsied. When patients were divided into two groups including those who underwent "liver biopsy" (n = 284 biopsies from 231 patients) and all other biopsies except liver, i.e., "others" (n = 238 biopsies from 228 patients). Success rates of 95.02 and 98.32% were observed in the two groups, respectively (**Supplementary Tables 3A,B**).

Chi-Squared Tests to Assess Covariates of Failure

Results from chi-squared tests showed that the covariates of age category at earliest biopsy date, gender, modality

[1]https://openwetware.org/mediawiki/index.php?title=ApE_-_A_Plasmid_Editor_(software_review)&oldid=753142

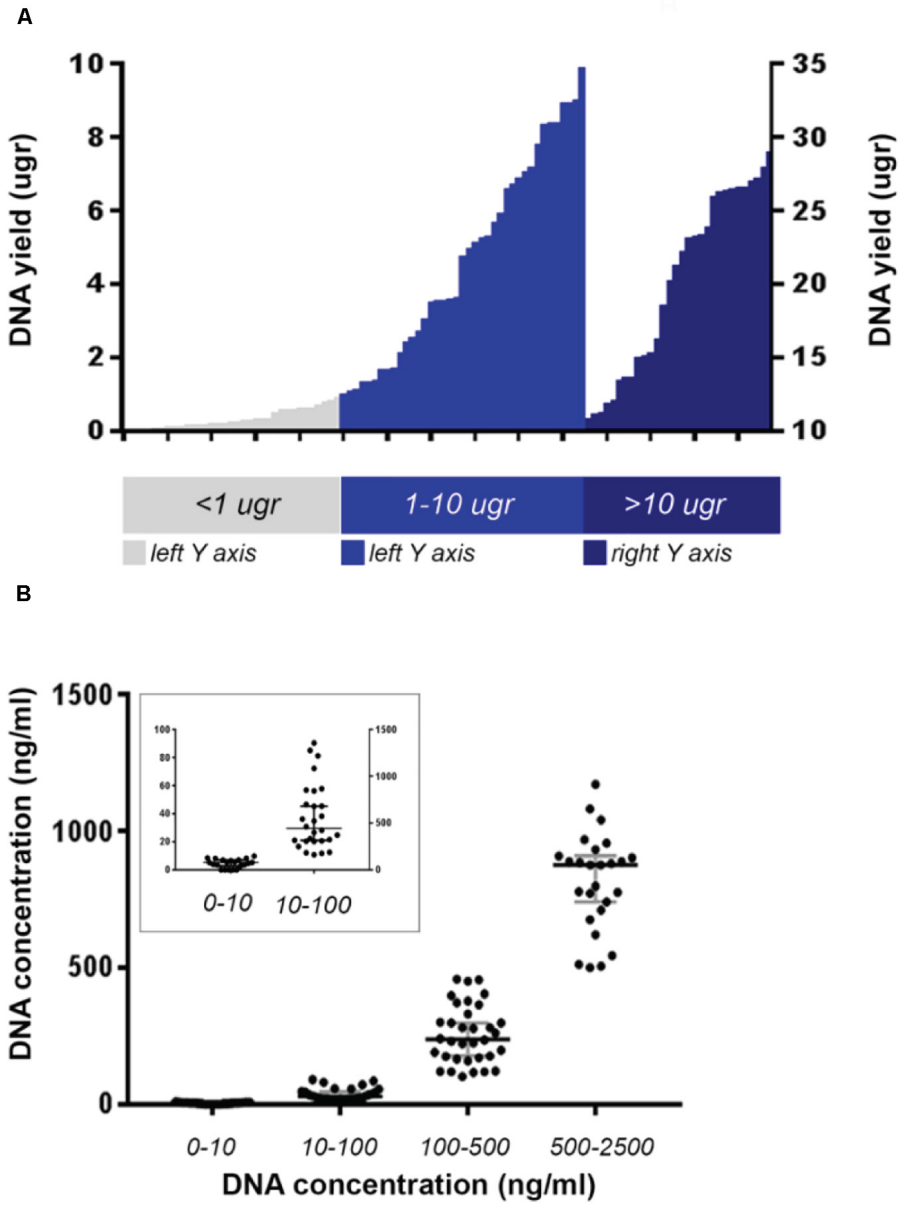


FIGURE 2 | Maximum DNA yield of the whole analyzed cohort from PROSPECT-C and PROSPECT-R patients. **(A)** Cases with DNA yield >10 μg are plotted against right *Y*-axis. **(B)** Cases divided according to their DNA concentration. Median value and with 95% CI represented in gray bars. In the small square cases with DNA concentration < 100 ng/ml are plotted against the left *Y*-axis.

of image guidance, metastatic site, and seniority of the interventionist were not associated with the occurrence of failure. Association of site of biopsy (others vs. liver), however, showed a significant trend in favor of other organs vs. liver, although the difference was not found to be numerically and clinically of significant impact (p = 0.053). Patients who had biopsy within clinical trials (n = 163 biopsies) vs. those who underwent routine clinical diagnostic biopsies (n = 338) showed a success rate of 98.79 and 95.48%, respectively. Chi-squared test demonstrated significance in favor of patients treated within clinical trials (p = 0.07; **Supplementary Table 4**).

Validated Genomic Testing in Patients With Metastatic Colorectal Cancer

Given that metastatic colorectal cancer (mCRC) patients underwent genomic profiling for clinically actionable mutations such as *KRAS*, *NRAS*, and *BRAF* analysis routinely with a clinically validated COBAS panel, we rationalized separating this

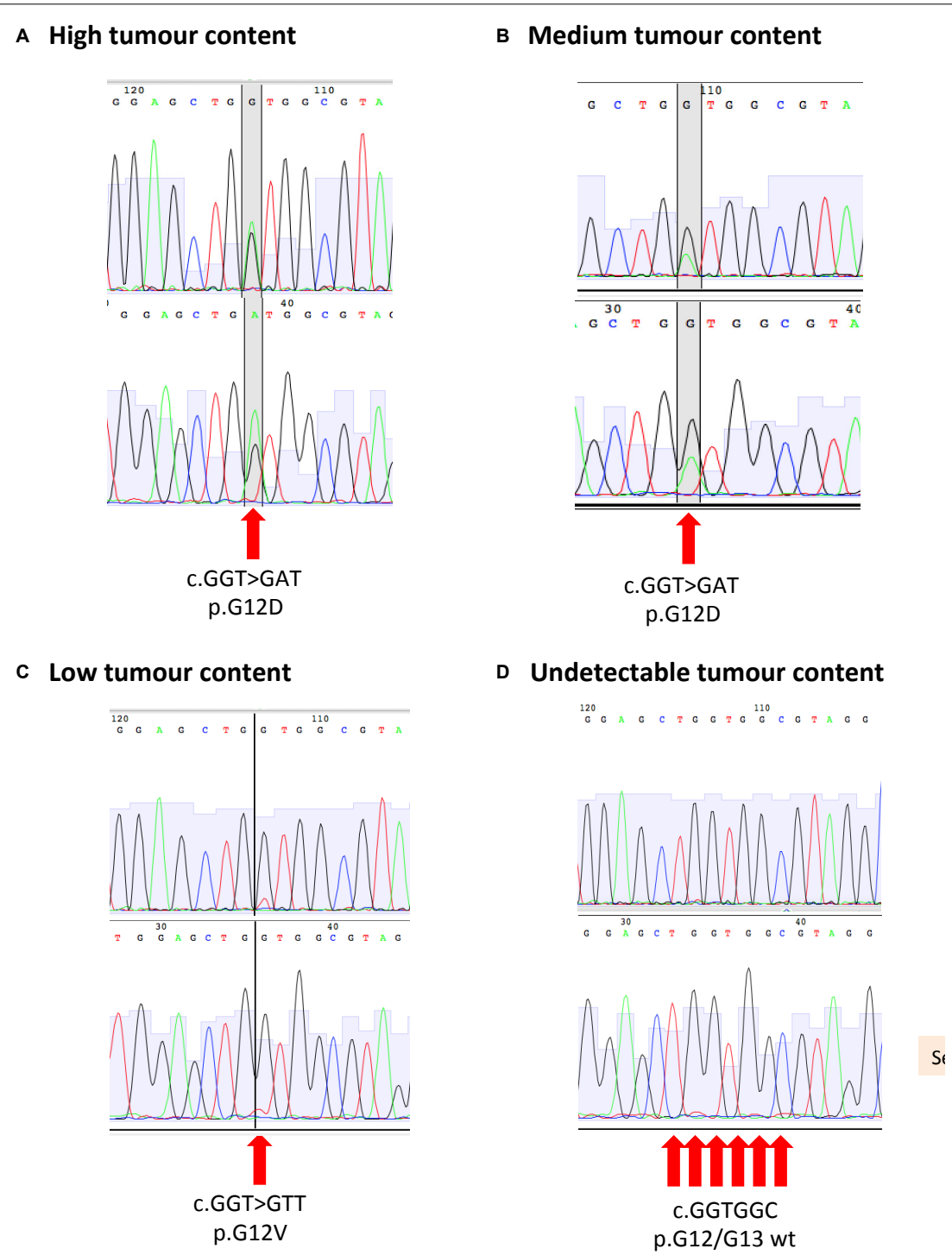


FIGURE 3 | Sanger sequencing for KRAS p.G12/G13 mutation with example of **(A)** high tumour content, **(B)** medium tumour content, **(C)** low tumour content, and **(D)** undetectable tumour content.

cohort from the remaining patients. Of the total 144 patients with mCRC, 17 repeat charts and 38 patients who were referred from other hospitals were excluded. Of the remaining 89 patients, 2 (2.25%) had a failed molecular analysis due to insufficient DNA extraction—29 (32.58%), 6 (6.74%), and 3 (3.37%) were found to have *KRAS* exons 2–4, *NRAS* exons 2–4, and *BRAF v600* mutations, respectively. Moreover, 36 patients were tested for *TP53* and *PIK3CA* mutation; 26 (72.22%); and 5 (13.89%) were

found to have these mutations, respectively. These results are largely consistent with previously published literature.

gDNA Extraction From Biopsy: A Cohort of PROSPECT-C and PROSPECT-R Studies

DNA was extracted from 62 biopsies taken from our prospective PROSPECT-C and PROSPECT-R trials and in 65% of cases; sufficient gDNA for WES was achieved from a single core. Two or three tissue cores were needed to yield sufficient DNA for 27 and 8% of the biopsy time points, respectively (**Supplementary Figure 1**). When required, utilizing all available tissue cores allowed gDNA extraction rate of 100%. Tumor content was determined for 62 biopsies (75.61%) in the analyzed cohort and was estimated as >20% in 79.03% of cases (**Table 1**).

Assessment of Biopsy Tumor Content

All patients entering the PROSPECT trials were tested for *KRAS*/NRAS mutations in the archival tumor biopsy by standard COBAS methodology, as this precluded entry into PROSPECT-C study (2). As a result, all patients entering the PROSPECT-R trial had a cataloged *KRAS*/*NRAS* variant that could be used to investigate the tumor content of the respective biopsy samples (1). Mutation profiles for *KRAS*/*NRAS* have previously been shown to be highly concordant between samples from the same colorectal tumor (9). We therefore estimated the cancer cell content of biopsy samples using Sanger sequencing to detect the likely truncal *KRAS*/*NRAS* mutations identified previously by clinical sequencing assays. Samples were scored according to the following criteria: "high" tumor content if the variant base was detected at an intensity exceeding or equal to the wild-type base; "medium" if the variant base was detected at >25% of the intensity of the wild-type base; "low" if the variant base was clearly detected above background but at <25% of the intensity of the wild-type base; and "not detected" if the variant base could only be detected within the background noise or not at all (**Table 2** and **Figure 3**). Further cores were extracted and sequenced if the first had low or no detectable tumor content (**Table 2**). In five cases, an additional core had medium tumor content where the first tested core has low/not detectable tumor content. Out of the 49 samples tested, 39 were scored as medium or high tumor content (**Table 2**).

TABLE 1 | DNA extraction and estimated tumor content.

	PROSPECT-C	PROSPECT-R	Total
Number of BL	15	31	46
Number of PD	15	21	36
Number of pairs (BL/PD)	7	19	26
Attempted DNA extraction (total)	30	33	63
Attempted DNA extraction (BL)	15	31	46
Attempted DNA extraction (PD)	15	2	17
Attempted DNA extraction (pairs BL/PD)	7	2	9
DNA yield > 1 μg (total)	27	31	58
DNA yield > 1 μg (BL)	13	29	42
DNA yield > 1 μg (PD)	14	2	16
DNA yield > 1 μg (pairs BL/PD)	6	2	8
Estimated tumor content (total)	30	32	62
≥20%	22	27	49
<20%	8	5	13
Estimated tumor content (BL)	15	31	46
≥20%	11	27	38
<20%	4	4	8
Estimated tumor content (PD)	15	1	16
≥20%	11	0	11
<20%	4	1	5
Estimated tumor content in BL/PD pairs	7	1	8
≥20%	4	0	4
<20%	3	1	4

BL,baseline; PD, progressive disease.

DISCUSSION

Tissue biopsies are often considered as the gold standard for diagnostic and research purposes; however, there are many logistical, technical, and ethical challenges in the successful appliance of tissue biopsies in the clinic. To our knowledge, we present the largest dataset of tissue biopsies with a prospective validation cohort demonstrating high tumor yield and ability to perform genomic analysis via image-guided tissue sampling (10, 11).

Biomarker discovery requires validation in prospective clinical trials; however, tissue collection procedures need to be optimized such that the valuable tissue obtained during trials is processed successfully (12, 13). Moreover, even within a resource-friendly environment, molecular profiling studies have often suffered due to inadequacy of samples; failure rates reportedly vary between 15 and 33% (14–17). Keeping these issues in view, we ensured that prebiopsy scans were discussed in person with a radiologist, and only the most amenable lesions were chosen for pretreatment biopsies; experienced radiologists were then able to target multiple cores (6) from the periphery of the chosen lesions. The current study demonstrates that a strong infrastructure and good communication allows high-quality tumor samples to be obtained in a time-efficient manner. The coaxial biopsy technique used at the RM has the advantage of puncturing the capsule of solid organs (liver, kidney, spleen) only once, minimizing

TABLE 2 | Sample tumor content estimated from Sanger sequencing.

	PROSPECT-R
Number of samples sequenced:	49
Number of cores sequenced per sample:	
1 core:	35
2 cores:	10
3 cores:	4
Number of high tumor content cores:	12
Number of medium tumor content cores:	27
Number of low tumor content cores:	6
Number of cores with no detectable tumor content:	22

the injury to normal tissues, improving patient experience, and at the same time acquiring multiple large cores for diagnosis and molecular analysis. The application of preformed gelatin sponge sealing device at the biopsy tract provides a mechanical matrix that facilitates clotting. Gelfoam pledgets, due to their bulk, surface-acting hemostatic agents, slow the flow of blood, protect the forming clot, and offer a framework for the deposition of the cellular elements of blood, decreasing the risk of major bleeding (18). The grade III hemorrhage in our series was only 0.2%, which compares favorably with the 0.5–2% seen in large series in the literature using Tru-Cut needles with or without coaxial technique (19, 20).

Common concerns about trials mandating research biopsies include the lack of patient understanding about the purpose of such studies and the potential risks associated with additional interventional procedures within the research protocols (21–24). In the case of our PROSPECT-C trial (1, 2), patients included in the study had access to anti-EGFR antibody treatment via the cancer drug fund (CDF) independent of the research biopsy findings, which meant that the research biopsies were of no direct patient benefit. In order to ensure that patients clearly understood the purpose of their participation in PROSPECT-C and other research studies, a prospective patient-based survey at the RM was performed. Remarkably, it showed that most patients who consented to a research biopsy gave an altruistic reason (e.g., to help research and/or others) as to why they agreed to participate (25). A common concern regarding trial-related invasive intervention is procedure-related complications. Notably, the biopsy complication rates in more than 500 patients in our cohort (including patients on PROSPECT studies) were extremely low and compared favorably with published literature (26, 27). The technical reasons for success can be attributed to the use of large gage coaxial needles, which enable multiple tissue cores to be sampled with a single pass. Subsequent application of gelatin foam pledgets via a coaxial cannula at withdrawal effectively seals the biopsy track and minimizes hemorrhage (<1%), thus enabling safe same-day discharge in the majority of patients. This technique, however, needs to be carefully considered in appropriate patients; for example, any attempt to biopsy lung parenchyma would carry a significant risk of the gelfoam pledget deploying in a pulmonary vein resulting in systemic embolus. We, however, acknowledge that the exceptional safety observed in our cohort may not be reproducible in a less resource-friendly environment, as it is highly operator dependent, and thus, clinicians are encouraged to audit their own data when determining the need for requesting tissue biopsies.

Following the safe acquisition of biopsy material, the processing of tumor samples has its own challenges. First, the acquired sample contains a mixture of cancer cells and stroma (connective tissue, blood vessels, and inflammatory cells). It is well established that stromal infiltration may lead to problems in interpreting genomic data (28, 29). In contemporaneous studies conducted at the RM (e.g., FOrMAT study), sample failure rates were high with only 16% of samples showing tumor content of >50% (30). The FOrMAT study collected a range of GI tumor samples including pancreatic cancers, which are more likely to be dominated by inflammatory and stromal cells (31), but it relied on using only FFPE tissues. FFPE tissue has limitations for complex genomic studies, as the DNA yield and quality are affected by the process of fixation and paraffin embedding (32–36). The PROSPECT studies benefited from parallel analysis using both FFPE and fresh frozen tissue, where the former was used for pathological assessment and the latter for molecular characterization and genomic analysis. By utilizing all available tissue cores as required, we achieved a gDNA extraction of >90% and an estimated tumor content of >20% in 87.27% of the cases. These data compare favorably with a recent large-scale study comprising of >10,000 patients, who were subjected to a hybridization-based next-generation sequencing (NGS) panel capable of detecting all-protein coding mutations, copy number alterations, and selected promoter mutations and structural rearrangements (37).

We next took into account the limitations of tumor estimates generated by subjective pathological assessment of tumor morphology and cellularity estimates. Cellularity can be estimated by quantifying the mutant alleles using technologies, such as Sanger or Ion Torrent sequencing, but this requires prior knowledge of the mutation (29, 38); in PROSPECT-R, Sanger sequencing was used to assess tumor content, as *RAS* mutation was a prerequisite for entry into the study. However, only patients with no known *RAS* pathway mutation could participate in PROSPECT-C, so alternative techniques were required for tumor cellularity estimates. An unbiased statistical approach that directly measures tumor content from the DNA sample, therefore, allowed us to take into account factors such as tumor ploidy and intratumor heterogeneity (ITH). This study highlights the safety of tissue biopsies and has significant clinical implications in the management of various malignancies—repeat biopsies should be considered in clinically relevant cases, for example at the time of progression on targeted therapies. Moreover, recent data by our group (1, 2) and others have demonstrated strong concordance between solid and liquid biopsies, and thus, the latter can be considered where a clinically validated panel is available and answers the relevant clinical question.

CONCLUSION

Oncologic management and clinical trial participation require accurate histological and molecular characterization. Image-guided biopsies using large gage coaxial needles enable multiple tissue cores to be obtained with a single pass. This increases the probability of diagnostic success for complex molecular analysis. Applying gelatin foam pledgets via the coaxial cannula following biopsy to seal the track reduces hemorrhagic risk and enables safe same-day discharge in the majority of patients. By successfully obtaining sufficient number of tumor tissue samples within prospective trials, such studies can further the

understanding of tumor biology and help develop biomarkers of clinical and translational relevance. Ultimately, this will enhance the application of personalized medicine in the clinic.

AUTHOR'S NOTE

This work was presented in ESMO 2017 and the abstract was published in Annals of Oncology in supplements.

DATA AVAILABILITY STATEMENT

The accession number for the DNA and RNA sequencing data reported in this manuscript is European Genome-phenome Archive: EGAS00001003367 and EGAD00001004501.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Royal Marsden Research and Ethics Committee. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version.

FUNDING

The authors acknowledge support from the National Institute for Health Research Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and The Institute of Cancer Research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01634/full#supplementary-material

REFERENCES

1. Khan K, Rata M, Cunningham D, Koh DM, Tunariu N, Hahne JC, et al. Functional imaging and circulating biomarkers of response to regorafenib in treatment-refractory metastatic colorectal cancer patients in a prospective phase II study. *Gut.* (2017) 67:1484–92. doi: 10.1136/gutjnl-2017-314178
2. Khan KH, Cunningham D, Werner B, Vlachogiannis G, Spiteri I, Heide T, et al. Longitudinal liquid biopsy and mathematical modeling of clonal evolution forecast time to treatment failure in the PROSPECT-C phase II colorectal cancer clinical trial. *Cancer Discov.* (2018) 8:1270–85. doi: 10.1158/2159-8290.cd-17-0891
3. Cancer Genome Atlas Network. *TCGA Tissue Sample Requirements: High Quality Requirements Yield High Quality Data*. (2016). Available online at: https://www.cancerresearchuk.org/sites/default/files/smp1_booklet_1.2_-_no_marks.pdf (accessed August 14, 2020).
4. Hsu MY, Pan KT, Chen CM, Lui KW, Chu SY, Lin YY, et al. CT-guided percutaneous core-needle biopsy of pancreatic masses: comparison of the standard mesenteric/retroperitoneal versus the *trans*-organ approaches. *Clin Radiol.* (2016) 71:507–12. doi: 10.1016/j.crad.2016.02.021
5. Olson MC, Atwell TD, Harmsen WS, Konrad A, King RL, Lin Y, et al. Safety and accuracy of percutaneous image-guided core biopsy of the spleen. *AJR Am J Roentgenol.* (2016) 206:655–9. doi: 10.2214/ajr.15.15125
6. Ocak S, Duplaquet F, Jamart J, Pirard L, Weynand B, Delos M, et al. Diagnostic accuracy and safety of CT-guided percutaneous transthoracic needle biopsies: 14-Gauge versus 22-gauge needles. *J Vascul Int Radiol JVIR.* (2016) 27:674–81. doi: 10.1016/j.jvir.2016.01.134
7. Woolston A, Khan K, Spain G, Barber LJ, Griffiths B, Gonzalez-Exposito R, et al. Genomic and transcriptomic determinants of therapy resistance and immune landscape evolution during Anti-EGFR treatment in colorectal cancer. *Cancer Cell.* (2019) 36:35–50.e9. doi: 10.1016/j.ccell.2019.05.013
8. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernandez-Mateos J, Khan K, et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science (New York N Y).* (2018) 359:920–6.
9. Brannon AR, Vakiani E, Sylvester BE, Scott SN, McDermott G, Shah RH, et al. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome Biol.* (2014) 15:454.
10. Khan K, Gonzalez Exposito R, Cunningham D, Koh DM, Woolston A, Kouvelakis K, et al. Diagnostic accuracy and safety of coaxial core-needle biopsy (CNB) system in a predominanty gastrointestinal oncology patient population, treated at the Royal Marsden (RM) Hospital. *Ann Oncol.* (2017) 28(Suppl. 3):iii6. doi: 10.1093/annonc/mdx263.014
11. Khurum K, Gonzalez-Exposito R, Cunningham D, Koh D, Woolston A, Barber L, et al. Diagnostic accuracy and safety of coaxial core-needle biopsy (CNB) system in Oncology patients treated in a specialist cancer centre with prospective validation within clinical trial data. *medRxiv* (2020). [Preprint]. doi: 10.1101/2020.04.17.20065458
12. Diamandis M, White NM, Yousef GM. Personalized medicine: marking a new epoch in cancer patient management. *Mol Cancer Res.* (2010) 8:1175–87. doi: 10.1158/1541-7786.mcr-10-0264
13. Wistuba II, Gelovani JG, Jacoby JJ, Davis SE, Herbst RS. Methodological and practical challenges for personalized cancer therapies. *Nat Rev Clin Oncol.* (2011) 8:135–41. doi: 10.1038/nrclinonc.2011.2
14. Ferté C, Massard C, Ileana E, Hollebecque A, Lacroix L, Ammari S, et al. Abstract CT240: Molecular screening for cancer treatment optimization (MOSCATO 01): a prospective molecular triage trial; Interim analysis of 420 patients. *Cancer Res.* (2014) 74:CT2140.
15. Andre F, Bachelot T, Commo F, Campone M, Arnedos M, Dieras V, et al. Comparative genomic hybridisation array and DNA sequencing to direct treatment of metastatic breast cancer: a multicentre, prospective trial (SAFIR01/UNICANCER). *Lancet Oncol.* (2014) 15:267–74. doi: 10.1016/s1470-2045(13)70611-9
16. Meric-Bernstam F, Brusco L, Shaw K, Horombe C, Kopetz S, Davies MA, et al. Feasibility of large-scale genomic testing to facilitate enrollment onto genomically matched clinical trials. *J Clin Oncol Off J Am Soc Clin Oncol.* (2015) 33:2753–62. doi: 10.1200/jco.2014.60.4165
17. Le Tourneau C, Delord JP, Goncalves A, Gavoille C, Dubot C, Isambert N, et al. Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. *Lancet Oncol.* (2015) 16:1324–34.
18. Lungren MP, Lindquester WS, Seidel FG, Kothary N, Monroe EJ, Shivaram G, et al. Ultrasound-guided liver biopsy with gelatin sponge pledget tract embolization in infants weighing less than 10 kg. *J Pediatr Gastroenterol Nutrit.* (2016) 63:e147–51. doi: 10.1097/mpg.0000000000001429

19. Atwell TD, Smith RL, Hesley GK, Callstrom MR, Schleck CD, Harmsen WS, et al. Incidence of bleeding after 15,181 percutaneous biopsies and the role of aspirin. *AJR Am J Roentgenol.* (2010) 194:784–9. doi: 10.2214/ajr.08.2122
20. Lipnik AJ, Brown DB. Image-guided percutaneous abdominal mass biopsy: technical and clinical considerations. *Radiol Clin North Am.* (2015) 53: 1049–59.
21. Pentz RD, Harvey RD, White M, Farmer ZL, Dashevskaya O, Chen Z, et al. Research biopsies in phase I studies: views and perspectives of participants and investigators. *Irb.* (2012) 34:1–8.
22. Olson EM, Lin NU, Krop IE, Winer EP. The ethical use of mandatory research biopsies. *Nat Rev Clin Oncol.* (2011) 8:620–5. doi: 10.1038/nrclinonc.2011.114
23. Saggese M, Dua D, Simmons E, Lemech C, Arkenau H. Research biopsies in the context of early phase oncology studies:clinical and ethical considerations. *Oncol Rev.* (2013) 7:e5. doi: 10.4081/oncol.2013.232
24. Peppercorn J, Shapira I, Collyar D, Deshields T, Lin N, Krop I, et al. Ethics of mandatory research biopsy for correlative end points within clinical trials in oncology. *J Clin Oncol Off J Am Soc Clin Oncol.* (2010) 28:2635–40. doi: 10.1200/jco.2009.27.2443
25. Moorcraft SY, Begum R, Cunningham D, Peckitt C, Baratelli C, Gillbanks A, et al. Attitudes of patients with gastrointestinal cancers toward research biopsies. *Clin Colorectal Cancer.* (2016) 16:e181–9. doi: 10.1016/j.clcc.2016.09.008
26. Tacher V, Le Deley MC, Hollebecque A, Deschamps F, Vielh P, Hakime A, et al. Factors associated with success of image-guided tumour biopsies: results from a prospective molecular triage study (MOSCATO-01). *Eur J Cancer.* (2016) 59:79–89. doi: 10.1016/j.ejca.2016.02.006
27. Howlett DC, Drinkwater KJ, Lawrence D, Barter S, Nicholson T. Findings of the UK national audit evaluating image-guided or image-assisted liver biopsy. Part II. Minor and major complications and procedure-related mortality. *Radiology.* (2013) 266:226–35. doi: 10.1148/radiol.12120224
28. Laird PW. Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet.* (2010) 11:191–203. doi: 10.1038/nrg2732
29. Song S, Nones K, Miller D, Harliwong I, Kassahn KS, Pinese M, et al. qpure: A tool to estimate tumor cellularity from genome-wide single-nucleotide polymorphism profiles. *PLoS One.* (2012) 7:e45835. doi: 10.1371/journal.pone.0045835
30. Moorcraft SY, Gonzalez de Castro D, Cunningham D, Jones T, Walker BA, Peckitt C, et al. Investigating the feasibility of tumour molecular profiling in gastrointestinal malignancies in routine clinical practice. *Ann Oncol Off J Eur Soc Med Oncol.* (2018) 29:230–6. doi: 10.1093/annonc/mdx631
31. Erkan M, Reiser-Erkan C, Michalski CW, Kleeff J. Tumor microenvironment and progression of pancreatic cancer. *Exp Oncol.* (2010) 32:128–31.
32. Koshiba M, Ogawa K, Hamazaki S, Sugiyama T, Ogawa O, Kitajima T. The effect of formalin fixation on DNA and the extraction of high-molecular-weight DNA from fixed and embedded tissues. *Pathol Res Pract.* (1993) 189:66–72. doi: 10.1016/s0344-0338(11)80118-4
33. Howat WJ, Wilson BA. Tissue fixation and the effect of molecular fixatives on downstream staining procedures. *Methods.* (2014) 70:12–9. doi: 10.1016/j.ymeth.2014.01.022
34. Zsikla V, Baumann M, Cathomas G. Effect of buffered formalin on amplification of DNA from paraffin wax embedded small biopsies using real-time PCR. *J Clin Pathol.* (2004) 57:654–6. doi: 10.1136/jcp.2003.013961
35. Cree IA, Deans Z, Ligtenberg MJ, Normanno N, Edsjo A, Rouleau E, et al. Guidance for laboratories performing molecular pathology for cancer patients. *J Clin Pathol.* (2014) 67:923–31. doi: 10.1136/jclinpath-2014-202404
36. Cancer Research UK. *CRUK Stratified Medicine Program*. (2016). Available online at: http://wwwcancerresearchukorg/sites/default/files/smp1_booklet_12_-_no_markspdf (accessed May 14, 2016).
37. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med.* (2017) 23:703–13.
38. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature.* (2011) 475:348–52.

Conflict of Interest: DC received research funding from Roche, Amgen, Celgene, Sanofi, Merck Serono, Novartis, AstraZeneca, Bayer, Merrimack, and MedImmune. IC has had advisory roles with Merck Serono, Roche, Sanofi Oncology, Bristol Myers Squibb, Eli-Lilly, Novartis, and Gilead Science. He has received research funding from Merck-Serono, Novartis, Roche and Sanofi Oncology, and honoraria from Roche, Sanofi-Oncology, Eli-Lilly, and Taiho. KKh has advisory role with Bayer Oncology group. NV received honoraria from Merck Serono, Bayer, and Eli-Lilly.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

PERSPECTIVE
published: 18 September 2020
doi: 10.3389/fonc.2020.566297

Dissociated Response in Metastatic Cancer: An Atypical Pattern Brought Into the Spotlight With Immunotherapy

Olivier Humbert [1,2]* *and David Chardin* [1,2]

[1] Department of Nuclear Medicine, Centre Antoine-Lacassagne, Université Côte d'Azur, Nice, France, [2] TIRO-UMR E 4320, Université Côte d'Azur, Nice, France

When evaluating metastatic tumor response to systemic therapies, dissociated response is defined as the coexistence of responding and non-responding lesions within the same patient. Although commonly observed on interim whole-body imaging, the current response criteria in solid cancer do not consider this evolutive pattern, which is, by default, assimilated to progression. With targeted therapies and chemotherapies, dissociated response is observed with different frequencies, depending on the primary cancer type, treatment, and imaging modality. Because FDG PET/CT can easily assess response on a lesion-by-lesion basis, thus quickly revealing response heterogeneity, a PET/CT dissociated response has been described in up to 48% of women treated for a metastatic breast cancer. Although some studies have underlined a specific prognostic of dissociated response, it has always ended up being described as an unfavorable prognostic pattern and therefore assimilated to the "Progressive Disease" category of RECIST/PERCIST. This dichotomous imaging report (response vs. progression) provides a simple information for clinical decision-support, which probably explains the relatively low consideration for the dissociated response pattern to chemotherapies and targeted therapies until now. With immune checkpoint inhibitors, this paradigm is quickly changing. Dissociated response is observed in around 10% of advanced lung cancer patients and appears to be associated to treatment efficiency. Indeed, for this subset of patients, a clinical benefit of immunotherapy and favorable prognosis are usually observed. This specific pattern should therefore be considered in the future immunotherapy-adapted criteria for response evaluation using CT and PET/CT, and specific clinical managements should be evaluated for this response pattern.

Keywords: dissociated, response, immunotherapy, metastatic cancer, heterogeneous response, imaging

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Egesta Lopci, University of Milan, Italy
Ryogo Minamimoto, National Center for Global Health and Medicine, Japan

***Correspondence:**
Olivier Humbert
olivier.humbert@univ-cotedazur.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *27 May 2020*
Accepted: *12 August 2020*
Published: *18 September 2020*

Citation:
Humbert O and Chardin D (2020) Dissociated Response in Metastatic Cancer: An Atypical Pattern Brought Into the Spotlight With Immunotherapy. Front. Oncol. 10:566297. doi: 10.3389/fonc.2020.566297

INTRODUCTION

When evaluating tumor response to systemic therapies in the metastatic setting, dissociated response (also termed mixed response, or heterogeneous response) is usually defined as the coexistence of responding and non-responding lesions within the same patient (**Figure 1**). Although dissociated response is a commonly observed evolutive pattern to systemic therapies,

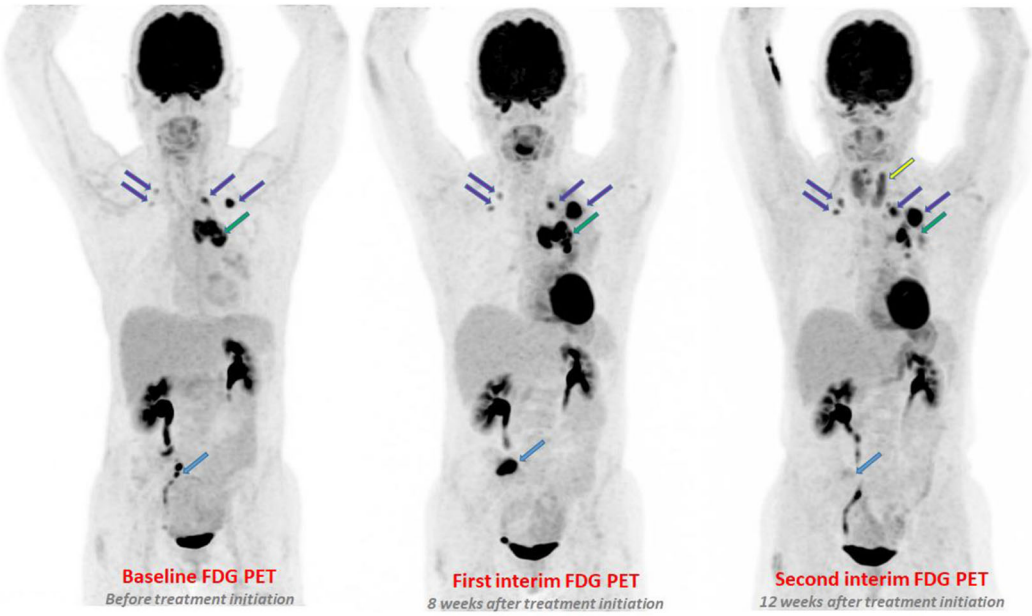


FIGURE 1 | A 60-year-old women with metastatic adenocarcinoma treated with pembrolizumab. On first PETinterim at 8 weeks of treatment (three cycles), the maximum intensity projection image (MIP) showed a metabolic progression of a bone lesion of the sacrum (light blue arrow), of a left mediastino-hilar mass (green arrow) and of lung nodules (purple arrows). The subsequent PETinterim, performed at 12 weeks of treatment (5 cycles), showed a dissociated response with disappearance of the bone lesion of the sacrum (light blue arrow), clear metabolic regression of the left mediastino-hilar mass (green arrow) and metabolic progression of the lung nodules (purple arrows). A thyroiditis was also observed (yellow arrow). Because, the clinical status of the patient remained stable, the treatment was continued and the patient finally obtained a prolonged durable clinical benefit of immunotherapy (16 months of treatment).

little is known about the biological specifications or the prognostic significance of this atypical pattern. This review aims to report what we already know about dissociated response in the setting of targeted therapies and chemotherapies, and highlight the new knowledge gained with the appearance of immunotherapy.

PHYSIO-PATHOLOGICAL HYPOTHESES

A combination of factors may explain the underlying biological mechanisms of a dissociated tumor response.

Firstly, because genomic instability occurs during the clonal evolution of solid cancers, multiple coexisting metastases can arise from genetically different tumor clones (1, 2). Indeed, cancerous cells not only undergo clonal evolution from a single progenitor cell into more aggressive and therapy resistant cells, but also exhibit branched evolution, whereby each tumor develops and preserves multiple distinct sub-clonal populations (2). This genotypic and phenotypic heterogeneity is an unfavorable prognostic factor for survival and can explain a dissociated response, particularly when using targeted therapies due to their selective pressure on tumor evolution (1–3).

Secondly, micro-environmental differences among metastatic sites can also induce heterogeneous responses. For instance, systemic therapies have a lower diffusion in bone tissue which is due to highly complex and variable interactions between tumor cells, bone cells and the bone matrix and can lead to a lower efficiency. Concerning brain metastasis, the blood-brain barrier can be a critical obstacle to the diffusion of certain drugs even though they are effective in other organs. Moreover, the heterogeneity of the immune environment of the lesions can actively influence therapeutic response and therefore explain different responses across lesions (4).

Some authors have suggested exercising caution when observing a dissociated response (5). Indeed, some unrelated processes may be pitfalls and erroneously mimic a mixed response. These pitfalls include: Synchronous neoplasms, inflammatory processes observed on FDG PET/CT (fat necrosis, diverticulitis…) or treatment-related effects (radiation-induced inflammation…). Thus, Clark et al. have defined an apparent dissociated response on FDG PET/CT to be a red flag that should lead to reconsider whether all findings are metastases of the same cancer. But we believe that radiologists and nuclear physicians are aware of this risk and that this pitfall rarely explain the heterogeneous response pattern assessed on PET/CT.

IMPACT OF THE IMAGING MODALITY ON RESPONSE ASSESSMENT

What is striking concerning the few papers that have studied dissociated response is that most were performed using FDG PET/CT. At baseline, PET/CT is a highly sensitive imaging technique that can quickly depict the whole metastatic pool

of lesions since hypermetabolic lesions appear with a high contrast. After initiating systemic treatment, FDG-PET/CT provides a whole-body quantitative assessment of treatment-induced changes in tumoral glycolysis and can be used to assess response on a lesion-by-lesion basis early on. Thus, a unique lesion with discordant evolution within the whole tumor load can easily be detected, revealing response heterogeneity (6–8). This explains why, even though the dissociated response pattern can also be assessed with CT, it is mostly described in PET/CT studies.

It's worth noting that dissociated response can be observed at any time during the imaging follow-up of a patients treated with systemic therapies. Despite the fact that most studies have evaluated dissociated response occurrence during the first 2 or 3 months of treatment (**Table 1**), it is currently unknown if it is more likely to be observed at early or late evaluation time points.

IMPACT OF THE PRIMARY CANCER TYPE AND TREATMENT TYPE

A dissociated response can be observed with chemotherapies, targeted-therapies or immunotherapies, but its frequency varies across treatment types and primary cancer types. It seems more common in cancers with heterogeneous molecular profiles between metastases, such as breast cancer (16), and has recently been described with treatments targeting the immune response (**Table 1**). Surprisingly, few studies have evaluated the prognostic significance of such an atypical pattern.

Chemotherapies + Targeted Therapies in Solid Cancers

Metastatic Breast Cancer (mBC)

In mBC, monitoring the treatment response with CT scan is hampered by the high frequency of bone metastases (70%) for which the apparent size does not necessarily change with treatment response (17, 18). The Response Evaluation Criteria in Solid Tumors (RECIST V1.1) specify that only lytic or mixed bone lesions with soft tissue components can be considered as measurable lesions (19). Because FDG PET/CT does not have this limitation, mBC is one of the first solid cancers for which PET has been routinely used to assess response. This is also the first metastatic setting in which the dissociated response was described. In 2010, Huyge et al. showed the intra-individual variability of the PET/CT metabolic response among lesions in bone-dominant metastatic breast cancer patients (12). These women were treated with different systemic therapies: chemotherapy (78%), hormone therapy (35%), anti-HER2 targeted therapy (4%). The metabolic response was analyzed according to the European Organization for Research and Treatment of Cancer (EORTC) criteria. Dissociated metabolic response occurred in 48% of patients, concerning mostly bone lesions, and tended to be associated with a better outcome than homogeneous non-response ($p = 0.07$). This result may suggest that, in case of dissociated disease evolution, the prognosis will depend on the number, the localization and the intrinsic aggressivity of the progressing lesion. This is, to our knowledge, the only available study on this topic in mBC.

Non-small-cell Lung Cancer (NSCLC)

In 2014, using CT scan, Lee et al. published a retrospective study including 68 patients with NSCLC who received second line EGFR-TKIs after a progression under systemic treatment (11). They observed that 32% of patient showed a dissociated response, and that this pattern was more frequent than homogeneous progression (19% of patients). Dong et al. published a larger retrospective study in 2017, including 246 consecutive patients with NSCLC and a response assessment with FDG PET/CT (3). The overall incidence of dissociated response was 21.5% and tended to occur more often in patients with advanced NSCLC (IIIB-IV) than those with earlier disease (I-IIIA) (30.0 vs. 5.8%, $p < 0.001$) and in patients treated with targeted therapies (EGFR-TKI) compared to those treated with chemotherapy (47.2 vs. 28.0%, $p = 0.008$). A dissociated response was an independent unfavorable prognostic factor for PFS ($p = 0.04$) and OS ($p = 0.006$) compared to patients with homogeneous evolution (homogeneous response or non-response). Interestingly, patients having a dissociated response were further categorized into those with "efficacious" dissociated response (i.e., only local PET-based disease progression and few clinical symptoms) and those with "inefficacious" dissociated response (the other patients). Most of the patients with efficacious dissociated response (65%) maintained prior regimens, with or without local intervention, whereas most patients with inefficacious dissociated response (63%) switched to next-line regimens. Compared to patients with an inefficacious dissociated response, the patients with an efficacious one showed a significant improvement in progression-free survival (9.4 vs. 3.8 months; $p = 0.012$) and overall survival (26.5 vs. 9.5 months; $p = 0.027$). This result underlines the need to recognize different patterns of dissociated response in NSCLC to improve outcome prediction for these patients.

Metastatic Colorectal Cancer (mCRC)

In a prospective multicentric study including 92 patients with a mCRC treated with a combination of sorafenib and capecitabine, a dissociated response was observed in one third of patients on interim FDG PET/CT (7). The presence of at least one metabolically non-responding lesion was associated with a poor outcome compared to patients without any metabolically non-responding lesions. But no prognostic difference was observed between patients with a dissociated response and patients with a homogeneous non-response. Therefore, the study concluded that the presence, but not the number, of non-responding lesions was the most important prognostic determinant. In a smaller study, only including nine patients with advanced KRAS wild-type mCRC treated with anti-EGFR therapy, a dissociated response was observed in nearly half of the patients (10).

Melanoma

In a Phase I monocentric trial evaluating the metabolic response of 23 patients with a BRAF mutant metastatic melanoma treated with dabrafenib, an heterogeneous PET response was observed in 26% of patients on the first interim PET performed 2 weeks after treatment initiation, and was associated with a shorter time-to-progression compared to homogeneous response (9). No patients with homogeneous lesion progression was observed which is

TABLE 1 | Summary of publications concerning dissociated response in solid cancers.

References	First author	No. of patients	Primary cancer	Study design	Drug class	Imaging modality	% of DR	Prognostic significance of DR	Timing after treatment initiation
CHEMOTHERAPIES AND TARGETED THERAPIES IN SOLID CANCERS									
(9)	M. S. Carlino	23	Melanoma (≥ 2 lesions)	Monocentric Prospective	BRAF inhibitor	PET/CT	26%	DR = Shorter TTP than homogeneous responder	Day 15
(7)	A. Hendlisz	92	Metastatic colorectal cancer	Multicentric Prospective	Sorafenib + Capecitabine.	PET/CT	32%	The presence of at least one non-responding lesion is associated with a poorer out-come	Day 21
(10)	E. J. van Helden	9	Advanced KRAS wild-type colorectal adenocarcinoma	Monocentric Prospective	EGFR inhibitor	PET/CT	43%	-	Week 4
(11)	Y. Lee	68	Advanced NSCLC (IIIB-IV)	Monocentric Retrospective	EGFR-TKIs	CT and others	32%	-	Mostly at Week 8
(3)	Z. Y. Dong	246	Advanced or metastatic NSCLC	Monocentric Retrospective	Chemotherapy or EGFR-TKIs	PET/CT	21%	DR = an independent unfavorable prognostic factor for PFS and OS	NK
(12)	V. Huyge	25	Bone-dominant metastatic breast cancer	Monocentric Retrospective	Different systemic therapies	PET/CT	48%	TTP tends to be higher in patients with DR compared to those with a homogeneous non-response.	≤ 12 months
IMMUNOTHERAPIES IN SOLID CANCER									
(13)	M. Tazdait	160	Advanced NSCLC	Monocentric Retrospective	PD-1/L1 inhibitors	CT	7.5%	DR = better overall survival than true progression	Mostly at Week 6
(14)	T. Tozuka	62	Advanced NSCLC	Monocentric Retrospective	PD-1/L1 inhibitors	CT	9%	DR = favorable prognosis compared to homogeneous progression	≤2 months
(15)	O. Humbert	50	Advanced NSCLC (III-IV)	Monocentric Prospective	PD-1 inhibitors	PET/CT	10%	DR is associated with a clinical benefit of immunotherapy	Month 3

DR, dissociated response; TTP, time-to progression; CT, Computed tomography; PET/CT, Positron emission tomography/computed tomograpy; NSCLC, Non-Small Cell Lung Cancer; EGFR, Epidermal growth factor receptor; EGFR-TKI, epidermal growth factor receptor-tyrosine kinase inhibitor; PD1, programmed cell death-1; PD-L1, programmed death-ligand 1.

consistent with the high level of activity of dabrafenib in these patients. It is worth noting that the more commonly reported measure of ΔSUVmax (metabolic change in the hottest lesion) was not found to be a prognostic biomarker in this study.

Immunotherapies in Solid Cancer

Immunotherapy with immune checkpoints inhibitors represent a recent breakthrough in the treatment of various metastatic cancers, showing a benefit in overall survival (OS) across a broad range of cancer types (20–23). Indeed, a subset of patients with metastatic cancer can demonstrate a clinical benefit that can last several years even after stopping the treatment (20–23). Nonetheless, most patients do not exhibit response to immunotherapy and identifying patients that will benefit from immunotherapy as early as possible remains a crucial issue.

Because tumor shrinkage is not the unique pattern of tumor response anymore, immunotherapy has raised new challenges in the evaluation of tumor response, as much with CT than with PET/CT. Indeed, some responding-patients can have a transient increase in tumor burden and metabolism, or appearance of new lesions, followed by a delayed response or stability. This specific immune-related response pattern is termed 'pseudo-progression' (PsPD) and is explained by the immune infiltration of tumors that can both induce a morphologic and metabolic increase of lesions (24, 25). This has been integrated in new immunotherapy-adapted criteria for CT to maintain the treatment in patients beyond a first imaging progression: the iRECIST (26). When using FDG PET, new immune-related response criteria for solid tumors have also been proposed, but without any consensus.

Beyond pseudo-progression, three recent studies have shown that a dissociated response is another atypical evolutive pattern of response to immunotherapy in advanced NSCLC with prognostic significance. Dissociated response is defined as the coexistence of responding and non-responding lesions within the same patient. Using CT imaging, Tazdait et al. retrospectively evaluated 160 patients with NSCLC treated with anti-PD1/PD-L1 drugs (13). They applied different morphologic imaging criteria (RECIST V1.1, iRECIST, irRECIST) and found, on the first CT evaluation, 7.5% of patients exhibiting a dissociated response. Atypical patterns (pseudo-progression + dissociated response) were associated with a better overall survival than true progressions. Another retrospective study including 62 NSCLC patients also observed a dissociated response in 9.2% of patients treated with PD-1/L1 inhibitors, and confirmed the improved OS associated to this pattern compared to homogenous progression (14.0 vs. 6.6 months) (14).

Using FDG PET/CT, our team recently published a prospective study including 50 patients with NSCLC treated with pembrolizumab/nivolumab and demonstrating that 12% of the population had a pseudo-progression and 10% had a dissociated response (15). Unlike what had been done in previous studies, the dissociated response was not defined on the first PET/CT evaluation showing a PERCIST disease progression, but on the subsequent confirmatory PET evaluation performed a few weeks later (3 months after treatment initiation). Because all these patients with dissociated response had a preserved clinical status and a limited number of progressive lesions, the patients' oncologists decided to maintain the therapy. A 6-months clinical benefit of immunotherapy was reached for all of them. Thus, it's worth noting that a dissociated response, contrary to pseudo-progression, is not only described on the first PET exam showing a metabolic tumor progression but can be described at later time-points of disease progression.

To sum up, a dissociated response appears to be a common evolutive pattern during immunotherapy (around 10% of treated patients), as frequent as the well-described pseudo-progression pattern. As Tazdait et al. have mentioned, this profile can be difficult to identify when using the conventional RECIST assessment, and requires a deep analysis of CT images (13), whereas this pattern can be easier to identify with PET/CT due to its ability to analyze the whole pool of lesions with great sensitivity. These studies also underlined that a dissociated response corresponds to a sign of treatment efficacy rather than failure, with a favorable outcome compared to homogeneous progression. Yet the prognostic value of the other atypical evolutive pattern, i.e., pseudo-progression, still needs to be explored. Furthermore, the best time point to assess these evolutive patterns will need to be defined: on the first or on the subsequent imaging evaluation?

CURRENT STRATEGIES AND FUTURE DIRECTIONS

Studying tumoral heterogeneity requires assessing the response of the whole baseline metastatic tumor load without restriction in number of lesions nor sites. However, both morphological (RECIST V1.1, iRECIST…) and metabolic (EORTC, PERCIST) response criteria only consider a limited number of operator-selected target lesions (19, 26, 27). Thus, the clinical scenario of a dissociated response of a single metastatic lesion has currently neither been integrated to morphological nor to metabolic criteria of response. More concerning, a proper, consensual name has not even been given to this pattern yet, as it can be referred to as "mixed response," "heterogeneous response," or "dissociated response."

There is no clear consensus concerning the cut-off of changes in lesion size or metabolism (SUV) to define response or progression at a lesion level and therefore to define dissociated response at a patient level. In our point of view, dissociated response on CT exam should be inspired of RECIST V1.1 criteria and defined as a concomitant relative decrease in size >30% in some lesions and relative increase in size >20% in others (and/or presence of new lesions). On PET/CT, dissociated response definition should be inspired by PERCIST criteria and defined as a concomitant relative decrease >30% in some tumor lesions metabolism (ΔSUV) and relative metabolic increase >30% in others (and/or new hypermetabolic lesions).

The low consideration for this evolutive pattern may be because the response assessment is usually reported dichotomously by the radiologist (progression vs. response) to ease the clinician that needs to take a decision to continue

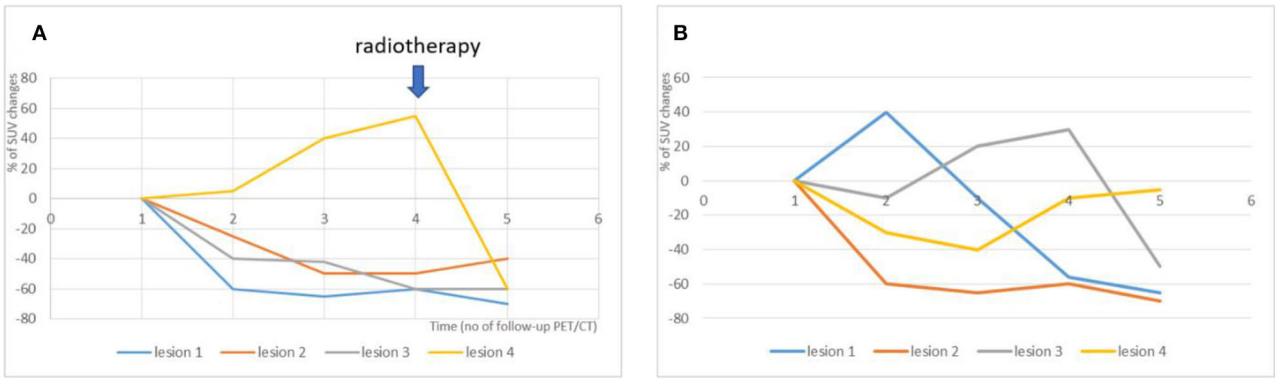


FIGURE 2 | Illustration of two different dissociated response patterns in patients with NSCLC treated with immunotherapy and subsequent clinical management. **(A)** A continuous progression of one isolated metastatic lesion was observed across successive PET/CT exams (yellow line), while the 3 other metastatic lesions were responsive to treatment (orange, blue and gray lines). The Multidisciplinary Tumor Board decided to continue immunotherapy with concomitant local radiotherapy on the progressive lesion (blue arrow). **(B)** Transient immune activation "one after the other" of various metastatic lesions was observed. This pattern is very close to the standard pseudo-progressive pattern. In this pattern, the Multidisciplinary Tumor Board decided to maintained immunotherapy, with no need of local treatment.

or change the treatment. Although a specific prognostic value of dissociated response has been underlined in some types of solid cancers, it was always considered as an unfavorable pattern of response. Dissociated response was therefore included in the "Progressive Disease" category of RECIST/PERCIST, based on the assumption that "one progressive lesion is enough to define progression."

With immune checkpoints inhibitors, the paradigm is quickly evolving: dissociated response is becoming a favorable prognostic pattern that absolutely needs to be distinguished from true progression. When a dissociated response to immunotherapy is observed, the continuation of the immune checkpoint inhibitors can provide a durable response (15). In our experience, different patterns of dissociated response can be observed in the immunotherapy setting (**Figure 2**):

- the continuous progression of oligo-metastatic lesions across successive exams, while the rest of the metastatic disease is under control. Because these few lesions show consistent resistance to immune therapies in consecutive exams, adding local ablative treatments to the growing lesions while pursuing the immune check-points inhibitors, may be a way to restore the prognosis.
- transient immune activation "one after the other" of various metastatic lesions, a pattern that is very close to the standard pseudo-progressive pattern. In this pattern, immunotherapy should be maintained, with no need of local treatment.

Thus, a dissociated response requires a specific categorization and should be captured in the future immunotherapy-adapted guidelines and criteria for CT and PET/CT, as is pseudo-progression in iRECIST. Further prospective works will be necessary to study the frequency and prognostic significance of the different dissociated patterns and optimize the best clinical management for each of them.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

1. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature.* (2013) 501:338–45. doi: 10.1038/nature12625
2. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* (2012) 72:4875–82. doi: 10.1158/0008-5472.CAN-12-2217
3. Dong Z, Zhai H, Hou Q, Su J, Liu S, Yan H, et al. Mixed responses to systemic therapy revealed potential genetic heterogeneity and poor survival in patients with non-small cell lung cancer. *The Oncologist.* (2017) 22:61–9. doi: 10.1634/theoncologist.2016-0150
4. Junttila MR, Sauvage FJ de. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature.* (2013). 501:346–54. doi: 10.1038/nature12626
5. Clark MS, Packard AT, Johnson DR, Johnson GB. Pitfalls of a mixed metabolic response at PET/CT. *RadioGraphics.* (2019) 39:1461–75. doi: 10.1148/rg.2019180093
6. Hendlisz A, Golfinopoulos V, Garcia C, Covas A, Emonts P, Ameye L, et al. Serial FDG-PET/CT for early outcome prediction in patients with metastatic colorectal cancer undergoing chemotherapy. *Ann Oncol Off J Eur Soc Med Oncol.* (2012) 23:1687–93. doi: 10.1093/annonc/mdr554
7. Hendlisz A, Deleporte A, Delaunoit T, Maréchal R, Peeters M, Holbrechts S, et al. The prognostic significance of metabolic response heterogeneity

in metastatic colorectal cancer. *PLOS ONE.* (2015) 10:e0138341. doi: 10.1371/journal.pone.0138341
8. Rousseau C, Devillers A, Sagan C, Ferrer L, Bridji B, Campion L, et al. Monitoring of early response to neoadjuvant chemotherapy in stage II and III breast cancer by [^{18} f]fluorodeoxyglucose positron emission tomography. *J Clin Oncol.* (2006) 24:5366–72. doi: 10.1200/JCO.2006.05.7406
9. Carlino MS, Saunders CAB, Haydu LE, Menzies AM, Martin Curtis C, Lebowitz PF, et al. 18F-labelled fluorodeoxyglucose–positron emission tomography (FDG–PET) heterogeneity of response is prognostic in dabrafenib treated BRAF mutant metastatic melanoma. *Eur J Cancer.* (2013). 49:395–402. doi: 10.1016/j.ejca.2012.08.018
10. van Helden EJ, Hoekstra OS, Boellaard R, Roth C, Mulder ER, Verheul HMW, et al. Early 18F-FDG PET/CT evaluation shows heterogeneous metabolic responses to anti-EGFR therapy in patients with metastatic colorectal cancer. *PLOS ONE.* (2016) 11:e0155178. doi: 10.1371/journal.pone.0155178
11. Lee Y, Kim HY, Lee S-H, Lim KY, Lee GK, Yun T, et al. Clinical significance of heterogeneity in response to retreatment with epidermal growth factor receptor tyrosine kinase inhibitors in patients with lung cancer acquiring secondary resistance to the drug. *Clin Lung Cancer.* (2014) 15:145–51. doi: 10.1016/j.cllc.2013.11.008
12. Huyge V, Garcia C, Alexiou J, Ameye L, Vanderlinden B, Lemort M, et al. Heterogeneity of metabolic response to systemic therapy in metastatic breast cancer patients. *Clin Oncol R Coll Radiol G B.* (2010) 22:818–27. doi: 10.1016/j.clon.2010.05.021
13. Tazdait M, Mezquita L, Lahmar J, Ferrara R, Bidault F, Ammari S, et al. Patterns of responses in metastatic NSCLC during PD-1 or PDL-1 inhibitor therapy: comparison of RECIST 1.1, irRECIST and iRECIST criteria. *Eur J Cancer.* (2018) 88:38–47. doi: 10.1016/j.ejca.2017.10.017
14. Tozuka T, Kitazono S, Sakamoto H, Yoshida H, Amino Y, Uematsu S, et al. Dissociated responses at initial computed tomography evaluation is a good prognostic factor in non-small cell lung cancer patients treated with anti-programmed cell death-1/ligand 1 inhibitors. *BMC Cancer [Internet].* (2020) doi: 10.1186/s12885-020-6704-z
15. Humbert O, Cadour N, Paquet M, Schiappa R, Poudenx M, Chardin D, et al. 18FDG PET/CT in the early assessment of non-small cell lung cancer response to immunotherapy: frequency and clinical significance of atypical evolutive patterns. *Eur J Nucl Med Mol Imaging.* (2020). 47:1158–67. doi: 10.1007/s00259-019-04573-4
16. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature.* (2000) 406:747–52. doi: 10.1038/35021093
17. Hamaoka T, Madewell JE, Podoloff DA, Hortobagyi GN, Ueno NT. Bone imaging in metastatic breast cancer. *J Clin Oncol Off J Am Soc Clin Oncol.* (2004) 22:2942–53. doi: 10.1200/JCO.2004.08.181
18. Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. *Clin Cancer Res Off J Am Assoc Cancer Res.* (2006) 12(20 Pt 2):6243s−9s. doi: 10.1158/1078-0432.CCR-06-0931
19. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer Oxf Engl.* (1990) (2009). 45:228–47. doi: 10.1016/j.ejca.2008.10.026
20. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med.* (2015) 372:2521–32. doi: 10.1056/NEJMoa1503093
21. Herbst RS, Baas P, Kim D-W, Felip E, Pérez-Gracia JL, Han J-Y, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet Lond Engl.* (2016) 387:1540–50. doi: 10.1016/S0140-6736(15)01281-7
22. Ferris RL, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med.* (2016) 10;375:1856–67. doi: 10.1056/NEJMoa1602252
23. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med.* (2015) 373:1803–13. doi: 10.1056/NEJMoa1510665
24. Aide N, De Pontdeville M, Lopci E. Evaluating response to immunotherapy with 18F-FDG PET/CT: where do we stand? *Eur J Nucl Med Mol Imaging.* (2020) 47:1019–21. doi: 10.1007/s00259-020-04702-4
25. Carter BW, Bhosale PR, Yang WT. Immunotherapy and the role of imaging. *Cancer.* (2018) 124:2906–22. doi: 10.1002/cncr.31349
26. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* (2017) 18:e143–52. doi: 10.1016/S1470-2045(17)30074-8
27. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving Considerations for PET response criteria in solid tumors. *J Nucl Med Off Publ Soc Nucl Med.* (2009) 50 (Suppl. 1):122S−50S. doi: 10.2967/jnumed.108.057307

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

OPINION
published: 23 September 2020
doi: 10.3389/fonc.2020.559946

The Role of Imaging Biomarkers Derived From Advanced Imaging and Radiomics in the Management of Brain Tumors

Faiq Shaikh[1], Diana Dupont-Roettger[1], Jamshid Dehmeshki[1,2], Omer Awan[3], Olga Kubassova[1] and Sotirios Bisdas[4]*

[1] Image Analysis Group, Philadelphia, PA, United States, [2] Department of Computer Science, Kingston University, Kingston-upon-Thames, United Kingdom, [3] Department of Radiology, University of Maryland Medical Center, Baltimore, MD, United States, [4] Department of Neuroradiology, University College London, London, United Kingdom

Keywords: imaging, biomarkers, radiomics, brain, tumors

OPEN ACCESS

Edited by:
Fatima-Zohra Mokrane, Université Toulouse III Paul Sabatier, France

Reviewed by:
Yuming Jiang, Stanford University, United States
Guolin Ma, China-Japan Friendship Hospital, China

****Correspondence:***
Faiq Shaikh
faiq.shaikh@ia-grp.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *07 May 2020*
Accepted: *13 August 2020*
Published: *23 September 2020*

Citation:
Shaikh F, Dupont-Roettger D, Dehmeshki J, Awan O, Kubassova O and Bisdas S (2020) The Role of Imaging Biomarkers Derived From Advanced Imaging and Radiomics in the Management of Brain Tumors. Front. Oncol. 10:559946. doi: 10.3389/fonc.2020.559946

INTRODUCTION

Significant advances have been made in the realm of medical image analysis in the past few decades, aimed at improving our understanding of the disease—how it develops, behaves, and responds to treatment. Advanced imaging strategies using magnetic resonance imaging (MRI) and positron emission tomography (PET) provide structural and functional phenotypic biomarkers that correlate with key disease processes. Radiomics-based biomarkers provide a deeper analysis of pathophysiologic processes and insights to better diagnose, classify, stratify, and prognosticate brain tumors, and to assess their response to therapy.

Radiomics in Neuro-Oncology

Radiomics is an imaging analysis methodology that involves the extraction of quantifiable features, which serve as biomarkers for structural changes as well as pathophysiological processes in disease entities. Applying radiomics yields a numerical dataset that can be parsed, processed, and analyzed using machine learning methods (1). Radiomics-based biomarkers can provide key insights in the diagnosis, classification, and therapeutic management of various solid tumors. It is also beginning to have an impact in the management of neuro-oncological diseases, including low-grade gliomas, glioblastoma multiforme (GBM), and brain metastases (2). There is a wide spectrum of radiomics applications in this field, ranging from accurate classification of brain lesions (gliomas vs metastases, IDH-wild type vs. -mutant tumors), therapy planning (radiation therapy response prediction), and immunotherapy response assessment.

Methodology

Radiomics analysis may be performed on computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT). Lesion identification and image segmentation are performed as the first steps and can be a manual or automated process, followed by 3D reconstruction performed on these regions of interest.

The next step is that of feature extraction and classification (FE/FC). These features are categorized as shape features (morphology-based), first-order statistics (histogram-based), and second-order statistics (texture analysis) features (3). Furthermore, higher-order statistics may also be extracted using mathematical transforms (such as Minkowski functionals, Laplace features, wavelet transforms, etc.) (4). Feature extraction produces several numerical values (depending on the imaging modality and the library used for extraction), which are then analyzed using advanced statistical or machine learning (ML) approaches, which may be supervised or unsupervised, and

include cluster analysis, support vector machine (SVM), random forest, convolutional neural network (CNN), and deep learning neural network (DLNN) (see **Figure 1**) (5). The main purpose is to train a model to identify radiomics features that can serve as imaging biomarkers for disease processes. This is followed by model validation and includes methods such as *k*-fold cross-validation to test the skill of the ML model. More recent works in neuro-oncology involving DLNN have revolved around automated tumor segmentation, quantification of disease burden, pseudoprogression assessment, multi-omics-based disease characterization, and prognostication.

DISCUSSION

The Current Imaging Biomarker Landscape in Neuro-Oncology

A noninvasive imaging biomarker may be described as a characteristic feature identifiable on an imaging study that indicates a key disease process. The key step is to establish these new biomarkers through correlation with ground truths, which could be the previously imaging-based "gold standards," clinical outcomes, or pathologic evidence. There is an increasing fund of quantitative imaging biomarkers (QIB) that are catalyzing the practice of precision medicine (6). In clinical trials, the QIBs are being used as surrogate endpoints, which can significantly reduce the time and incurred costs (7). QIBs are being explored as predictive classifiers for clinical trials, which can be used for patient selection/recruitment and in the timely determination of responders vs. nonresponders.

Brain lesions are structurally and functionally complex, and there is a growing focus on noninvasive methods to study this complexity to assess the disease status. Gliomas are a heterogeneous set of tumors, based on their issue, cellular, and molecular characteristics. The role of nonimaging biomarkers in gliomas and GBMs is well known, i.e., IDH1 mutation (8) and methylguanine-DNA methyltransferase (MGMT) promoter methylation (9). However, the role of imaging biomarkers in disease stratification or management guidance of GBM is less established.

Multiple imaging biomarkers have been identified for brain metastasis from various primary tumors. Multiparametric MRI, which includes apparent diffusion coefficient (ADC) and perfusion-weighted sequences, is used extensively in the clinical management of brain tumors. Perfusion-weighted and permeability MRI have been used for detection, delineation, and therapy response assessment of malignant brain lesions (10). Dynamic susceptibility contrast-enhanced MRI (DSC-MRI) deriving relative cerebral blood volume (rCBV) and cerebral blood flow (rCBF) values have led the quantifiable image biomarker discovery (11). Higher rCBV in the peritumoral edema, which may contain infiltrating angiogenic tumor cells, is indicative of primary intrinsic tumor as opposed to pure vasogenic edema seen in metastatic disease (12). However, the evidence for ADC to do the same is weak. Also, rCBV measurement from the solid tumoral region is another established discriminative biomarker for distinguishing GBM from the other tumor types (13).

Magnetic resonance spectroscopy (MRS) allows us to assess tissue metabolites noninvasively and has yielded several biomarkers of interest, such as choline (Cho)/creatinine (Cr) ratio, which is, for example, lower in cerebral metastases than in GBMs (14). Similarly, the peritumoral Cho/NAA ratio has also been shown to be useful to that effect (15). Furthermore, decreased creatine/phosphocreatine (Cr) values in patients with low-grade gliomas (WHO grade II) have been shown to correlate with better prognosis in terms of longer progression-free times and later malignant transformation (16). High levels of glycine have been reported in biopsies of patients with GBM (17). These, among other metabolites such as lactate, have been implicated as important MRS-based biomarkers for brain tumors.

18F-Fluorodeoxyglucose (FDG) PET/CT and more recently PET/MR has traditionally had a limited role in the management of primary brain tumors, primarily due to FDG biodistribution in the brain, hence there is an increasing role for amino-acid PET tracers in neuro-oncology. 18F-Fluoro-ethyl-tyrosine (FET) has been shown to detect recurrence in previously treated glioblastomas and is influenced by MGMT promoter methylation status (18). FET-PET-based biological tumor volume in newly diagnosed GBM has been shown to be a prognostic imaging biomarker for survival, independent of MGMT promoter methylation (19). However, it is important to note that the role of this biomarker for survival outcomes modeling has not been established. High tumoral amino-acid uptake using 11C-methionine (MET) PET is another well-studied biomarker for malignant gliomas and is independently associated with poor prognosis (20). α-[^{11}C]Methyl-L-tryptophan PET has also been shown to predict longer overall survival (21).

Radiomics-Based Imaging Biomarkers in Neuro-Oncology: A Novel Paradigm

Radiomic signatures are providing the next-generation imaging biomarkers that have implications in the management of brain tumors. These signatures are based on combinations of first-order histogram-based features (Haralik features, kurtosis, and entropy) and second-order texture analysis features (such as gray-scale run lengths). Key areas in neuro-oncology where radiomics has been initially applied are the following:

- ***Precision diagnostics and disease stratification/classification:*** Since primary and metastatic brain tumors are histologically and genetically heterogeneous, and it is important to understand the role tumor heterogeneity plays in the natural history of cancer, its response to therapy, and prognosis/outcomes (22–24). The extracted radiomics features provide a numerical value for the heterogeneous tumor microenvironment changes (25). GBM is a notoriously aggressive cancer, given its therapeutic resistance and high recurrence rate, both of which have underpinnings in its molecular heterogeneity (26). Radiomics has provided insights into the tissue and molecular heterogeneity and correlated with the underlying genetic alterations (27–29). Furthermore, molecular heterogeneity of GBM at the transcriptomic level

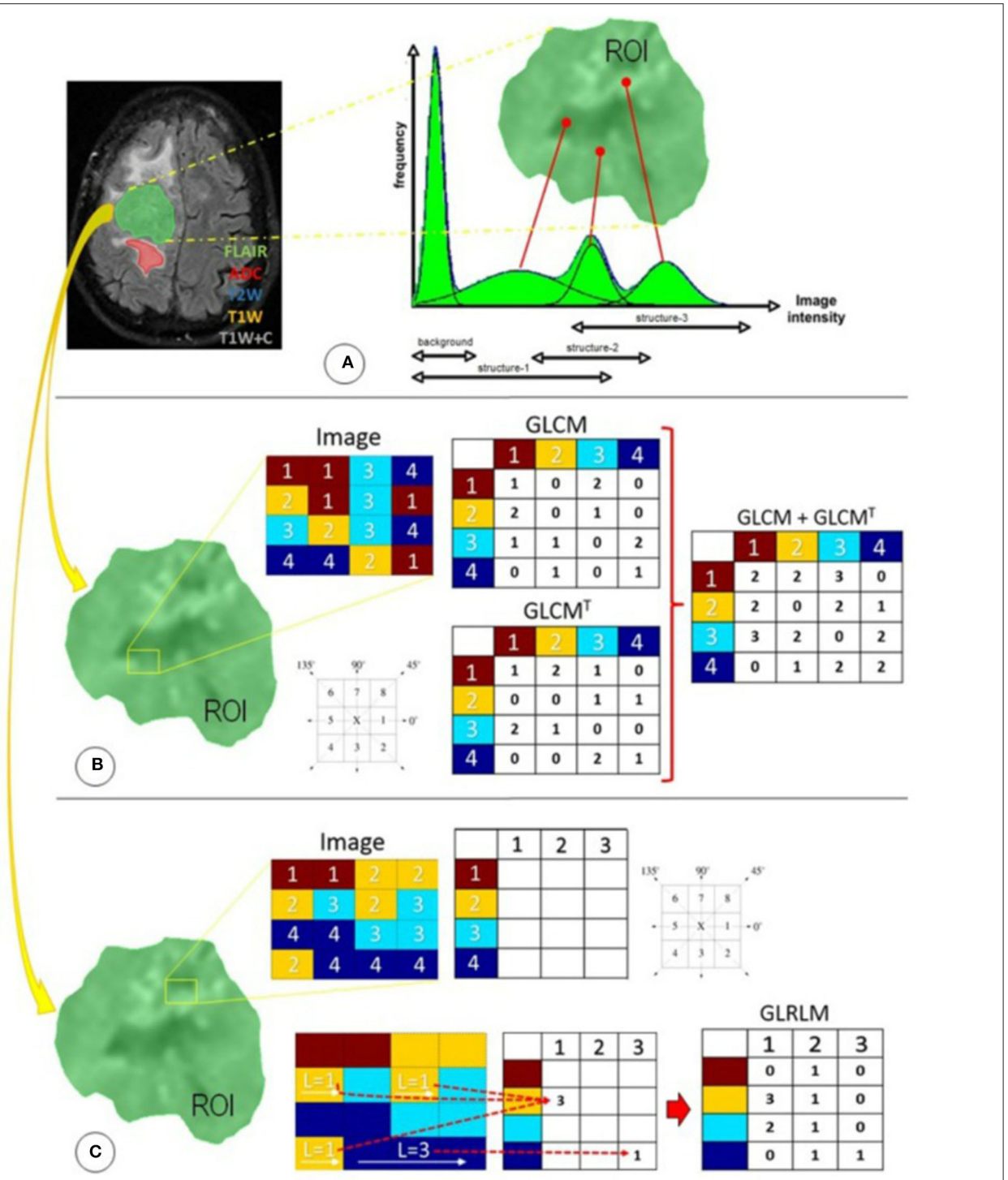


FIGURE 1 | Radiomics features used in this study were distributed in three different techniques focused primarily on statistical approaches: **(A)** first-order statistics, **(B)** second-order statistics through the GLCM, and **(C)** higher-order statistics through the GLRLM. ADC, apparent diffusion coefficient; FLAIR, fluid-attenuated inversion recovery; GLCM, gray-level co-occurrence matrix; GLCMT, gray-level co-occurrence matrix transpose; GLRLM, gray-level run-length matrix; L, length of homogeneous runs for each gray level; ROI, region of interest; T1W, T1-weighted precontrast; T1W+C, T1-weighted postcontrast; T2W, T2-weighted. (Reused from Florez E, Nichols T, E Parker E, T Lirette S, Howard CM, Fatemi A. Multiparametric magnetic resonance imaging in the assessment of primary brain tumors through radiomic features: a metric for guided radiation treatment planning. *Cureus.* (2018) 10:e3426. doi: 10.7759/cureus.3426, under the CC-BY license).

can be assessed using radiomics and may provide a framework to classify/stratify GBMs (30). Shofty et al. (31) demonstrated the ability of radiomics analysis of multiparametric MRI to stage 1p/19q co-deleted low-grade gliomas with sensitivity, specificity, and accuracy of 92, 83, and 87%, respectively (31). There is an interest in reclassifying many cancer types from the conventional histological basis to that based on radiogenomic signatures shedding light into various tissue heterogeneity patterns as they are better aligned with therapeutic responsiveness (32, 33).

- ***Disease prognostication and prediction modeling:*** There are multiple prognostic determinants for brain tumors, including the histologic subtype, specific genetic mutations, degree of anaplasia, degree of necrosis of fibrosis, degree of de-differentiation, local infiltration, vasculogenesis and resulting vascular scavenging, and hypoxia. For most of these processes, radiomics analysis can provide some degree of quantification, such as wavelet transforms for the degree of vascularity or Minkowski functionals for the degree of necrosis (34). Zhang et al. (35) demonstrated the use of Minkowski features among others to help differentiate radiation necrosis from tumor progression in patients with brain metastases undergoing gamma-knife surgery. MR-based radiomics analysis has been shown to predict overall survival and progression-free survival in GBM (36). Radiomics signatures correlate with and predict the expression of key molecular biomarkers in brain tumors, such as Ki-67 expression in low-grade gliomas or IDH mutation in GBM (37, 38). These early predictive models may provide bases of re-classifying cancers based on their progression and prognosis, allowing indolent cancers to be managed more conservatively while reserving more aggressive therapeutic approaches for more aggressive cancers. This is exemplified by a study by Davatzikos et al. where they showed that molecular features depicted by radiomics provided better risk stratification of GBM beyond the WHO classification (39). Furthermore, radiomics can help in the assessment of medical complications associated with brain tumors, such as epilepsy in patients with low-grade gliomas, which facilitates better disease management (40).
- ***Therapy response assessment and monitoring:*** Radiomics-based phenotype assessment of cancer lesions is an effective tool in the sensitivity profiling against therapeutic options (such as quantifying hypoxia to determine chemosensitivity), as well as an early assessment of therapy response (41–43). The standard visual assessment of radiological images for this purpose has been plagued by the confounding pseudoprogression. Current MRI techniques and human-based interpretation are tedious and prone to high interpersonal variability for accurate classification and prognostication of gliomas (44). The current Response Assessment in Neuro-oncology (RANO) criteria are used for GBM and the immunotherapy RANO (iRANO) criteria have been introduced to address the issue of pseudoresponse/pseudoprogression for both conventional chemoradiation and immunotherapies (45). Novel approaches using multiparametric MR and/or PET imaging combined with radiomics-based texture analysis can help evaluate subtle microstructural as well as functional changes at earlier time points than standard imaging (46). These can be quantifiable harbingers of true therapy response assessment and debunking pseudo-progression more accurately and earlier than conventional approaches. A multicenter study performed by Elshafeey et al. (47) using MR-based radiomics analysis for immunotherapy response assessment in GBM yielded an accuracy, specificity, and sensitivity of 91, 91, and 88%, respectively.

Radiomic tumor signatures can be incorporated into a multidimensional, multi- "omics" model, which uses genetic/molecular determinants to create a holistic genotype–phenotype the landscape of cancer and have the potential for informing the prognosis and accurately predicting/assessing therapy response (48). New approaches, such as using Multi Assay Experiment (MAE) as the container for multi-omics analysis, facilitates the process of data compilation and integration required for such complex analyses (49). Furthermore, quantitative scoring scales based on such ML-based analytical models have applications in clinical management as well as "go/no-go" decision-making in clinical trials (50, 51).

Generally speaking, there are factors that hinder the full-scale application and widespread acceptance of radiomics in the field of neuro-oncology. This includes the lack of user-friendly, FDA-approved software programs that perform radiomic analysis, the lack of a generalizable model to use for predictions, and the lack of a prospective study to show the added value of radiomics compared with the conventional ROI and histogram analysis.

Specific to its clinical application in neuro-oncologic management, there are certain gaps where radiomics has yet to make an impact. These include providing biomarkers for the precision guidance of the therapeutic management of brain tumors, particularly GBM. Genetic markers that are implicated in the prognostication and therapy guidance of GBM include gene amplification of epidermal growth factor receptor (EGFR), TP53, and PTEN mutation, among others (52, 53). Developing radiomic signatures that correlate with these genetic markers can help develop noninvasive imaging biomarkers for risk/severity stratification, survival outcomes, and therapy response prediction and assessment for these patients. Furthermore, having such radiomic biomarkers can catalyze the development of novel therapeutics using these genetic markers as targets.

The scope of radiomics applications is growing. When vetted through robust statistical analyses and real-world applications, it can augment the shift toward personalized, precision-based practices in neuro-oncology.

AUTHOR CONTRIBUTIONS

FS: conceptualization, subject expertise (cancer imaging + radiomics methodology). SB: subject matter expertise (clinical neuroradiology). OK: subject matter expertise

(mathematics, statistics). OA: subject matter expertise (clinical imaging). JD: subject matter expertise (computer science, machine learning methodology). DD-R: subject matter expertise (computer vision methodology). All authors contributed to the article and approved the submitted version.

REFERENCES

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
2. Zhou M, Scott J, Chaudhury B, Hall L, Goldgof D, Yeom KW, et al. Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches. *Am J Neuroradiol.* (2018) 39:208–16. doi: 10.3174/ajnr.A5391
3. Parmar C, Velazquez ER, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One.* (2014) 9:e0102107. doi: 10.1371/journal.pone.0102107
4. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson imaging.* (2012) 30:1234–48. doi: 10.1016/j.mri.2012.06.010
5. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJ. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep.* (2015) 5:13087. doi: 10.3389/fonc.2015.00272
6. Larue RT, Defraene G, De Ruysscher D, Lambin P, Van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol.* (2017) 90:20160665. doi: 10.1259/bjr.20160665
7. O'Connor JP, Jackson A, Asselin MC, Buckley DL, Parker GJ, Jayson GC. Quantitative imaging biomarkers in the clinical development of targeted therapeutics: current and future perspectives. *Lancet Oncol.* (2008) 9:766–76. doi: 10.1016/S1470-2045(08)70196-7
8. Turkalp Z, Karamchandani J, Das S. IDH mutation in glioma: new insights and promises for the future. *JAMA Neurol.* (2014) 71:1319–25. doi: 10.1001/jamaneurol.2014.1205
9. Weller M, Stupp R, Reifenberger G, Brandes AA, Van Den Bent MJ, Wick W, et al. MGMT promoter methylation in malignant gliomas: ready for personalized medicine?. *Nat Rev Neurol.* (2010) 6:39. doi: 10.1038/nrneurol.2009.197
10. Lacerda S, Law M. Magnetic resonance perfusion and permeability imaging in brain tumors. *Neuroimaging Clin.* (2009) 19:527–57. doi: 10.1016/j.nic.2009.08.007
11. Paulson ES, Schmainda KM. Comparison of dynamic susceptibility-weighted contrast-enhanced MR methods: recommendations for measuring relative cerebral blood volume in brain tumors. *Radiology.* (2008) 249:601–13. doi: 10.1148/radiol.2492071659
12. Chiang IC, Kuo YT, Lu CY, Yeung KW, Lin WC, Sheu FO, et al. Distinction between high-grade gliomas and solitary metastases using peritumoral 3-T magnetic resonance spectroscopy, diffusion, and perfusion imagings. *Neuroradiology.* (2004) 46:619–27. doi: 10.1007/s00234-004-1246-7
13. Xing Z, You RX, Li J, Liu Y, Cao DR. Differentiation of primary central nervous system lymphomas from high-grade gliomas by rCBV and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Clin Neuroradiol.* (2014) 24:329–36. doi: 10.1007/s00062-013-0255-5
14. Tsougos I, Svolos P, Kousi E, Fountas K, Theodorou K, Fezoulidis I, et al. Differentiation of glioblastoma multiforme from metastatic brain tumor using proton magnetic resonance spectroscopy, diffusion and perfusion metrics at 3 T. *Cancer Imaging.* (2012) 12:423. doi: 10.1102/1470-7330.2012.0038
15. Server A, Josefsen R, Kulle B, Mæhlen J, Schellhorn T, Gadmar Ø, et al. Proton magnetic resonance spectroscopy in the distinction of high-grade cerebral gliomas from single metastatic brain tumors. *Acta Radiol.* (2010) 51:316–25. doi: 10.3109/02841850903482901
16. Hattingen E, Raab P, Franz K, Lanfermann H, Setzer M, Gerlach R, et al. Prognostic value of choline and creatine in WHO grade II gliomas. *Neuroradiology.* (2008) 50:759–67. doi: 10.1007/s00234-008-0409-3
17. Choi C, Ganji SK, DeBerardinis RJ, Dimitrov IE, Pascual JM, Bachoo R, et al. Measurement of glycine in the human brain *in vivo* by 1H-MRS at 3 T: application in brain tumors. *Magn Reson Med.* (2011) 66:609–18. doi: 10.1002/mrm.22857
18. Munck af Rosenschold P, Costa J, Engelholm SA, Lundemann MJ, Law I, Ohlhues L, et al. Impact of [18F]-fluoro-ethyl-tyrosine PET imaging on target definition for radiation therapy of high-grade glioma. *Neuro Oncol.* (2015) 17:757–63. doi: 10.1093/neuonc/nou316
19. Suchorska B, Jansen NL, Linn J, Kretzschmar H, Janssen H, Eigenbrod S, et al. Biological tumor volume in 18FET-PET before radiochemotherapy correlates with survival in GBM. *Neurology.* (2015) 84:710–9. doi: 10.1212/WNL.0000000000001262
20. De Witte O, Goldberg I, Wikler D, Rorive S, Damhaut P, Monclus M, et al. Positron emission tomography with injection of methionine as a prognostic factor in glioma. *J Neurosurg.* (2001) 95:746–50. doi: 10.3171/jns.2001.95.5.0746
21. John F, Bosnyák E, Robinette NL, Amit-Yousif AJ, Barger GR, Shah KD, et al. Multimodal imaging-defined subregions in newly diagnosed glioblastoma: impact on overall survival. *Neuro Oncol.* (2019) 21:264–73. doi: 10.1093/neuonc/noy169
22. Fidler IJ. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res.* (1978) 38:2651–60.
23. Charles NA, Holland EC, Gilbertson R, Glass R, Kettenmann H. The brain tumor microenvironment. *Glia.* (2011) 59:1169–80. doi: 10.1002/glia.21136
24. Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A.* (2013) 110:4009–14. doi: 10.1073/pnas.1219747110
25. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
26. Parker NR, Khong P, Parkinson JF, Howell VM, Wheeler HR. Molecular heterogeneity in glioblastoma: potential clinical implications. *Front Oncol.* (2015) 5:55. doi: 10.3389/fonc.2015.00055
27. Cui Y, Tha KK, Terasaka S, Yamaguchi S, Wang J, Kudo K, et al. Prognostic imaging biomarkers in glioblastoma: development and independent validation on the basis of multiregion and quantitative analysis of MR images. *Radiology.* (2016) 278:546–53. doi: 10.1148/radiol.2015150358
28. Yang D, Rao G, Martinez J, Veeraraghavan A, Rao A. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. *Med Phys.* (2015) 42:6725–35. doi: 10.1118/1.4934373
29. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol.* (2017) 72:3–10. doi: 10.1016/j.crad.2016.09.013
30. Kong DS, Kim J, Ryu G, You H-J, Sung JK, Han YH, et al. Quantitative radiomic profiling of glioblastoma represents transcriptomic expression. *Oncotarget.* (2018) 9:6336–45. doi: 10.18632/oncotarget.23975
31. Shofty B, Artzi M, Bashat DB, Liberman G, Haim O, Kashanian A, et al. MRI radiomics analysis of molecular alterations in low-grade gliomas. *Int J Comput Assist Radiol Surg.* (2018) 13:563–71. doi: 10.1007/s11548-017-1691-5
32. Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1
33. Colen RR, Hassan I, Elshafeey N, Zinn PO. Shedding light on the 2016 World Health Organization Classification of Tumors of the Central Nervous System in the era of radiomics and radiogenomics. *Magn Reson Imaging Clin.* (2016) 24:741–9. doi: 10.1016/j.mric.2016.07.001

34. Sanduleanu S, Woodruff HC, De Jong EE, Van Timmeren JE, Jochems A, Dubois L, et al. Tracking tumor biology with radiomics: a systematic review utilizing a radiomics quality score. *Radiother Oncol.* (2018) 127:349–60. doi: 10.1016/j.radonc.2018.03.033
35. Zhang Z, Yang J, Ho A, Jiang W, Logan J, Wang X, et al. A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images. *Eur Radiol.* (2018) 28:2255–63. doi: 10.1007/s00330-017-5154-8
36. Chaddad A, Sabri S, Niazi T, Abdulkarim B. Prediction of survival with multi-scale radiomic analysis in glioblastoma patients. *Med Biol Eng Comput.* (2018) 56:2287–300. doi: 10.1007/s11517-018-1858-4
37. Li Y, Qian Z, Xu K, Wang K, Fan X, Li S, et al. Radiomic features predict Ki-67 expression level survival in lower grade gliomas. *J Neurooncol.* (2017) 135:317–24. doi: 10.1007/s11060-017-2576-8
38. Li ZC, Bai H, Sun Q, Zhao Y, Lv Y, Zhou J, et al. Multiregional radiomics profiling from multiparametric MRI: Identifying an imaging predictor of IDH1 mutation status in glioblastoma. *Cancer Med.* (2018). 7:5999–6009. doi: 10.1002/cam4.1863
39. Rathore S, Akbari H, Rozycki M, Abdullah KG, Nasrallah MP, Binder ZA, et al. Radiomic MRI signature reveals three distinct subtypes of glioblastoma with different clinical and molecular characteristics, offering prognostic value beyond IDH1. *Sci Rep.* (2018) 8:1–2. doi: 10.1038/s41598-018-22739-2
40. Liu Z, Wang Y, Liu X, Du Y, Tang Z, Wang K, et al. Radiomics analysis allows for precise prediction of epilepsy in patients with low-grade gliomas. *NeuroImage Clin.* (2018) 19:271–8. doi: 10.1016/j.nicl.2018.04.024
41. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* (2014) 5:1–9. doi: 10.1038/ncomms5644
42. Vallières M, Kay-Rivest E, Perrin LJ, Liem X, Furstoss C, Aerts HJ, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Sci Rep.* (2017) 7:1–4. doi: 10.1038/s41598-017-10371-5
43. Aerts HJ. The potential of radiomic-based phenotyping in precision medicine: a review. *JAMA Oncol.* (2016) 2:1636–42. doi: 10.1001/jamaoncol.2016.2631
44. Brandsma D, Stalpers L, Taal W, Sminia P, van den Bent MJ. Clinical features, mechanisms, and management of pseudoprogression in malignant gliomas. *Lancet Oncol.* (2008) 9:453–61. doi: 10.1016/S1470-2045(08)70125-6
45. Okada H, Weller M, Huang R, Finocchiaro G, Gilbert MR, Wick W, et al. Immunotherapy response assessment in neuro-oncology: a report of the RANO working group. *Lancet Oncol.* (2015) 16:e534–42. doi: 10.1016/S1470-2045(15)00088-1
46. Da Cruz LH, Rodriguez I, Domingues RC, Gasparetto EL, Sorensen AG. Pseudoprogression and pseudoresponse: imaging challenges in the assessment of posttreatment glioma. *Am J Neuroradiol.* (2011) 32:1978–85. doi: 10.3174/ajnr.A2397
47. Elshafeey N, Kotrotsou A, Hassan A, Elshafei N, Hassan I, Ahmed S, et al. Multicenter study demonstrates radiomic features derived from magnetic resonance perfusion images identify pseudoprogression in glioblastoma. *Nat Commun.* (2019) 10:1–9. doi: 10.1038/s41467-019-11007-0
48. Bera K, Beig N, Tiwari P. Opportunities and advances in radiomics and radiogenomics in neuro-oncology. In: *International Workshop on Radiomics and Radiogenomics in Neuro-oncology*. Cham: Springer (2019), p. 12–23.
49. Zanfardino M, Franzese M, Pane K, Cavaliere C, Monti S, Esposito G, et al. Bringing radiomics into a multi-omics framework for a comprehensive genotype-phenotype characterization of oncological diseases. *J Transl Med.* (2019) 17:337. doi: 10.1186/s12967-019-2073-2
50. Pinker K, Shitano F, Sala E, Do RK, Young RJ, Wibmer AG, et al. Background, current role, and potential applications of radiogenomics. *J Magn Reson Imaging.* (2018) 47:604–20. doi: 10.1002/jmri.25870
51. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin.* (2019) 69:127–57. doi: 10.3322/caac.21552
52. Hill C, Hunter SB, Brat DJ. Genetic markers in glioblastoma: prognostic significance and future therapeutic implications. *Adv Anat Pathol.* (2003) 10:212–7. doi: 10.1097/00125480-200307000-00004
53. Schmidt MC, Antweiler S, Urban N, Mueller W, Kuklik A, Meyer-Puttlitz B, et al. Impact of genotype and morphology on the prognosis of glioblastoma. *J Neuropathol Exp Neurol.* (2002). 61:321–8. doi: 10.1093/jnen/61.4.321

Conflict of Interest: FS, DD-R, OK, JD, and SB are full-time/part-time employees of Image Analysis Group.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 25 September 2020
doi: 10.3389/fonc.2020.573279

Feasibility of Optical Surface-Guidance for Position Verification and Monitoring of Stereotactic Body Radiotherapy in Deep-Inspiration Breath-Hold

Patrick Naumann[1,2,3], Vania Batista[1,2,3], Benjamin Farnia[4], Jann Fischer[1,2,3], Jakob Liermann[1,2,3], Eric Tonndorf-Martini[1,2,3], Bernhard Rhein[1,2,3,5] and Jürgen Debus[1,2,3,5,6,7]*

[1] Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany, [2] Heidelberg Institute of Radiation Oncology (HIRO), Heidelberg, Germany, [3] National Center for Tumor diseases (NCT), Heidelberg, Germany, [4] Department of Radiation Oncology, University of Miami, Miami, FL, United States, [5] Heidelberg Ion-Beam Therapy Center (HIT), Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany, [6] German Cancer Consortium (DKTK), Heidelberg, Germany, [7] Clinical Cooperation Unit Radiation Oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany

Background: Reductions in tumor movement allow for more precise and accurate radiotherapy with decreased dose delivery to adjacent normal tissue that is crucial in stereotactic body radiotherapy (SBRT). Deep inspiration breath-hold (DIBH) is an established approach to mitigate respiratory motion during radiotherapy. We assessed the feasibility of combining modern optical surface-guided radiotherapy (SGRT) and image-guided radiotherapy (IGRT) to ensure and monitor reproducibility of DIBH and to ensure accurate tumor localization for SBRT as an imaging-guided precision medicine.

Methods: We defined a new workflow for delivering SBRT in DIBH for lung and liver tumors incorporating SGRT and IGRT with cone beam computed tomography (CBCT) twice per treatment fraction. Daily position corrections were analyzed and for every patient two points retrospectively characterized: an anatomically stable landmark (predominately Schmorl's nodes or spinal enostosis) and a respiratory-dependent landmark (predominately surgical clips or branching vessel). The spatial distance of these points was compared for each CBCT and used as surrogate for intra- and interfractional variability. Differences between the lung and liver targets were assessed using the Welch *t*-test. Finally, the planning target volumes were compared to those of free-breathing plans, prepared as a precautionary measure in case of technical or patient-related problems with DIBH.

Results: Ten patients were treated with SBRT according this workflow (7 liver, 3 lung). Planning target volumes could be reduced significantly from an average of 148 ml in free breathing to 110 ml utilizing DIBH ($p < 0.001$, paired *t*-test). After SGRT-based patient set-up, subsequent IGRT in DIBH yielded significantly higher mean corrections for liver targets compared to lung targets (9 mm vs. 5 mm, $p = 0.017$). Analysis of spatial

OPEN ACCESS

Edited by:
Laurent Dercle,
Columbia University Irving Medical Center, United States

Reviewed by:
Cyril Jaudet,
Centre François Baclesse, France
Antoine Schernberg,
Hôpital Tenon AP-HP, France

***Correspondence:**
Patrick Naumann
patrick.naumann@med.uni-heidelberg.de

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 16 June 2020
Accepted: 08 September 2020
Published: 25 September 2020

Citation:
Naumann P, Batista V, Farnia B, Fischer J, Liermann J, Tonndorf-Martini E, Rhein B and Debus J (2020) Feasibility of Optical Surface-Guidance for Position Verification and Monitoring of Stereotactic Body Radiotherapy in Deep-Inspiration Breath-Hold. Front. Oncol. 10:573279. doi: 10.3389/fonc.2020.573279

distance between the fixed and moveable landmarks confirmed higher interfractional variability (interquartile range (IQR) 6.8 mm) than intrafractional variability (IQR 2.8 mm). In contrast, lung target variability was low, indicating a better correlation of patients' surface to lung targets (intrafractional IQR 2.5 mm and interfractional IQR 1.7 mm).

Conclusion: SBRT in DIBH utilizing SGRT and IGRT is feasible and results in significantly lower irradiated volumes. Nevertheless, IGRT is of paramount importance given that interfractional variability was high, particularly for liver tumors.

Keywords: stereotactic body radiation, Deep-inspiration breath-hold, surface-guided radiation therapy, Image-guided radiation therapy, precision radiation oncology, lung tumor, liver metastasis

INTRODUCTION

Focused delivery of high radiation doses to an extracranial tumor in few fractions is defined as stereotactic body radiotherapy (SBRT). It has become a commonly available and recognized treatment option for early stage primary tumors of or oligometastases from liver and lung primaries with a high rate of local control, often comparable to surgical resection (1–3). Due to technological advances in radiotherapy over the last decade, radiation plans with highly conformal dose distributions are widely available. This sculpted delivery of radiation dose is dependent on three-dimensional on-board imaging allowing image-guided radiotherapy (IGRT), which is now standard in modern linear accelerators. These improvements facilitate precise patient positioning and a safe and accurate characterization of dose deposition that are mandatory for SBRT. Nevertheless, moving targets are still challenging and respiratory motion management is the most crucial aspect for safe and effective utilization of SBRT (4).

In order to compensate for target motion, the International Commission on Radiation Units and Measurements (ICRU) introduced the concept of an internal margin to account for respiratory-induced changes in size, shape and position of a clinical target volume (CTV) (5). The addition of these internal margins to the CTV results in an internal target volume (ITV) to which further external margins for planning uncertainties are added to obtain a final planning target volume (PTV). This PTV is the volume that in the end receives the prescribed radiation dose. Yet, when this motion-encompassing approach is used for SBRT, the final PTV may become large or close to organs at risk (OAR), impeding the delivery of the high radiation doses needed for effective treatment. As a consequence, motion mitigation techniques were developed, which include: abdominal compression, beam-gating and breath-hold (6). Abdominal compression significantly reduces movement of the diaphragm and enables a good set-up accuracy for SBRT of liver and lower lung lobe targets in free-breathing (7, 8). For gated treatments an individual part of the respiratory cycle, usually the end-exhale phase, is chosen as a treatment window and further mitigates respiratory motion (4, 9). Tracking of targets with the beam is another method to manage respiratory motion, but requires real-time imaging during the treatment delivery (4, 10). The most reduction of respiratory movement, however, is achieved by breath-hold techniques that primarily target deep inspiration, which is tolerated longer than the end-exhale breath-hold phase and also carries less residual motion than gating of free breathing (4, 11).

For precise and accurate SBRT delivery, the reproducibility of deep inspiration breath-hold (DIBH) is crucial and has to be confirmed for several breath-hold cycles that are needed for a SBRT session. Depending on the PTV size, the beam-on time for SBRT typically varies between 2 and 5 min, despite already higher dose rates obtained by omission of the commonly used flattening filter (12). Reproducibility of DIBH can be assessed without additional radiation dose by either optical surface imaging solutions, such as AlignRT^{TM} (Vision RT, London, United Kingdom), or active breathing control devices, such as ABC^{TM} (Elekta, Stockholm, Sweden), which both reduce residual spatial uncertainties to 1–2 mm in standard radiotherapy of breast and lung cancer (13, 14).

Nevertheless, only limited data on SGRT and DIBH for SBRT is available. Here, we present our initial experience in utilizing a combination of SGRT and IGRT for patient positioning and treatment monitoring during SBRT of lung and liver targets.

MATERIALS AND METHODS

Study Design and Patient Selection

Ten patients transferred for lung or liver SBRT to our department in 2018/19 were included in this pilot study. Only patients who could hold their breath for at least 30 s were eligible for study inclusion. All patients gave informed consent for an individualized SBRT approach using DIBH instead of treatment in free-breathing and abdominal compression which is standard of care. For monitoring of correct DIBH reproducibility the hard- and software tools of Vision RT Ltd. (London, United Kingdom) were used. These commercial solutions are officially approved and licensed for this purpose. The analysis was approved by the local ethics committee (S-063/2019).

Radiotherapy Planning

Planning simulation was performed for all patients in vacuum cushion immobilization (BlueBAG BodyFIX^{TM}, Innovative Technologie Völp (IT-V), Innsbruck, Austria). The edges of the cushion near to the target area were folded and smoothed to prevent shadowing and concealment of parts of the body surface for appropriate optical surface-guidance. Abdominal

compression was not used to mitigate free breathing motion as it would prevent DIBH, recognition of the body surface by shadowing and could potentially modify the patient's surface in an unreproducible manner. The AZ-733V Respiratory Gating System (Anzai Medical Co., Ltd., Japan) was utilized for registration of breathing motion and recording of time resolved, four-dimensional (4D) CTs. After 4D-CT acquisition, the respiratory belt was unstrapped and additional CT series in free-breathing as well as DIBH were recorded.

All CT data were transferred to our institutional treatment planning system (RayStation 6B, RaySearch Laboratories, Stockholm, Sweden). Target volumes were delineated using available contrast-enhanced imaging in both free-breathing and DIBH sequences. Planning target volumes (PTV) were generated in DIBH by adding a 3–5 mm safety margin to the clinical target volume (CTV). The CTV was created from the gross tumor volume (GTV) with an isotropic margin of 5–7 mm to account for microscopic spread. In contrast, the free-breathing PTV was created by a 2–3 mm expansion of an internal target volume (ITV) that integrated all motion information of the CTV extracted from the 4D-CT data. Treatment plans were calculated for both DIBH and free-breathing PTV, with the latter prepared as a precautionary measure in case of technical issues or patient-related problems with breath-hold.

Treatment Delivery

The SBRT workflow of this study incorporates positioning and monitoring of patients in DIBH with surface-guidance (**Figure 1A**). A screenshot of the software (AlignRT^{TM} version 5.1.1, Vision RT, London, United Kingdom) used for DIBH positioning and monitoring is shown in **Figure 1B**.

In brief, the software allows for accurate patient set-up in DIBH and a radiation-free real-time feedback of DIBH positioning during the treatment session. Reproducibility of DIBH position was validated in-room by one to two repetitions. To verify the surface-guided position, a fast cone-beam computed tomography (CBCT), which lasts a breath-hold of 30 s, was acquired in DIBH. After registration to the planning CT, the couch was moved in DIBH accordingly and this new image guided position directly captured as a new surface reference in the SGRT software. To validate the reproducibility of tumor position in different breath-hold sequences, an additional fast CBCT was acquired. If a new registration resulted in shifts >2 mm, the couch was moved in DIBH as needed and the reference surface updated. Finally, orthogonal 2D-MV portal imaging was acquired in DIBH to confirm correct isocenter positioning.

After correct patient positioning the treatment was delivered in DIBH by a linear accelerator (Versa HD, Elekta, Stockholm, Sweden) using a flattening filter free (FFF) technique to reduce beam-on time whenever possible. We decided to set the individual defined region of interest (ROI), for which the SGRT software calculates differences in patient's actual surface and the reference surface, to the lower thorax to measure thoracic motion during DIBH and to avoid large distances between the treatment isocenter and the ROIs' centroid in liver SBRT. During treatment delivery the maximum allowed position error was 3 mm and 2° for translational and rotational differences, respectively. For patients' comfort the treatment delivery was also stopped every 20–40 s to allow for breaks and prevent patients from becoming out of breath. The overall treatment time including positioning and IGRT varied between 20 and 60 min depending on various patient specific factors concerning breath-hold (maximum tolerated breath-hold duration, time until normalization of the respiratory rate, reproducibility of breath-hold) resulting in different numbers of breath-holds needed.

Data Analysis

Differences in couch coordinates were assessed during the course of daily CBCT imaging (intrafractional) and compared to the derived shifts from subsequent treatment days (interfractional). For each patient and CBCT, a static, breathing-independent anatomic landmark (mostly Schmorl's nodes or enostosis of the spine at the level of the gross tumor) and a moving, breathing-dependent landmark (such as clips near the liver target lesion or a pulmonary vessel branching adjacent to the lung target lesion) were defined and their intra- and interfractional positions compared.

Statistical Analysis

Prism 7.04 software (GraphPad, San Diego, CA, United States) was used to perform statistical analysis and graphical plotting of results. Results were considered significant when the two-tailed p-value was less than 0.05. D'Agostino and Pearson omnibus K2 normality tests suggested a Gaussian distribution for the continuous variables of the vector length of position correction and spatial distance of fix and moving points, gross tumor volumes (GTV) and planning target volumes as well as the differences in planning target volumes in free-breathing and in DIBH for the same patient. Statistical differences between volumes of lung and liver targets were assessed using unpaired t-tests with Welch correction (Welch t-tests), which are more reliable for populations with unequal sample sizes and different variances than Student's t-test. Paired t-tests were conducted for comparison of PTVs in DIBH and in free-breathing of the same patients. Frequencies of needs for a second position correction were grouped by target location and analyzed in a 2×2 contingency table and differences analyzed using Fisher's exact test.

RESULTS

As a proof of concept analysis, we performed SBRT in DIBH in a total of 10 patients and for 41 treatment fractions using a combination of CBCT image- and surface-guidance for position verification and monitoring of DIBH. Details of our workflow are depicted in **Figure 1** and explained in the methods section. Targets were either a primary tumor or single metastasis of the lungs and liver in 3 and 7 patients, respectively. Details on patients' demographics and treatment delivery are shown in **Table 1**. Lung targets were mostly small peripheral lesions and could be treated in 3 fractions. In contrast, the liver tumors differed in size and some lesions were adjacent to organs at

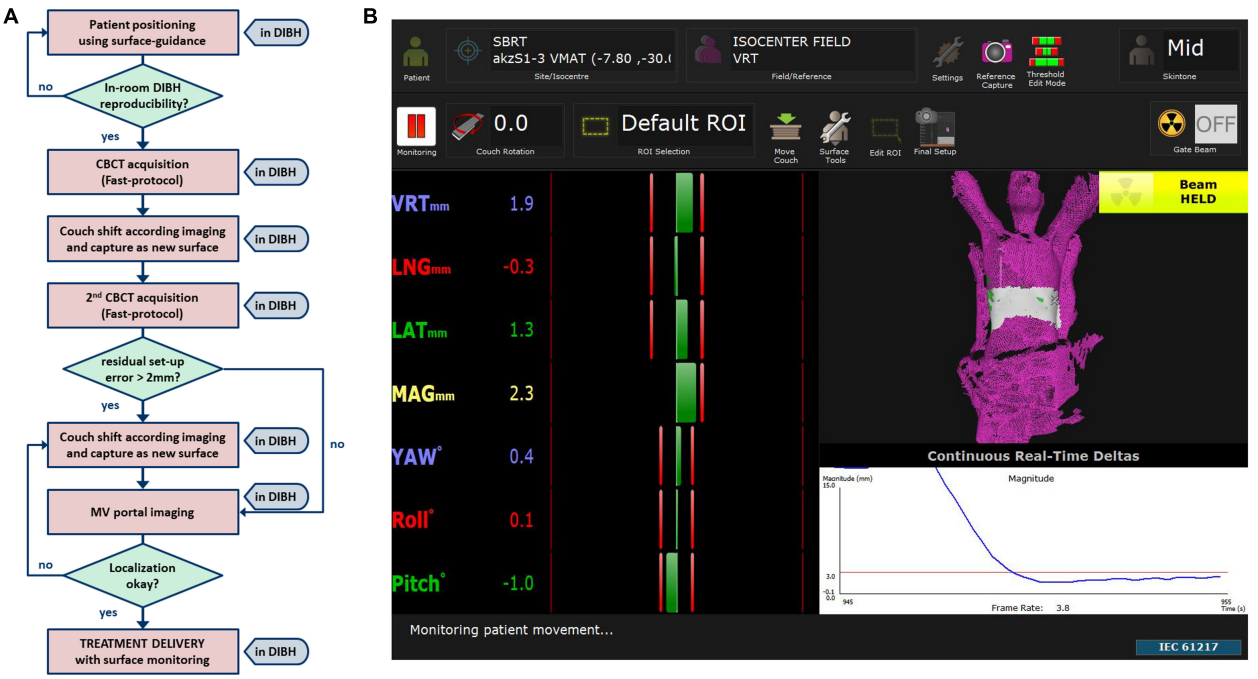


FIGURE 1 | (A) Work flow of patient set-up and localization until treatment delivery together with **(B)** an example screenshot of one of our SBRT patients showing a typical region of interest (ROI) placed to the lower thorax for DIBH monitoring of a liver target.

TABLE 1 | Patient demographics and treatment patterns.

No.	Sex, age	Primary tumor	Target lesion	Irradiation technique	Dose and fractionation	Prescription isodose†
L1	f, 38	Breast cancer	Lung metastasis (segment 8 left)	3DCRT (8 beams, FFF)	15 Gy ×3	65%
L2	f, 69	NSCLC	primary (segment 3 left)	3DCRT (8 beams, FFF)	15 Gy ×3	65%
L3	f, 59	NSCLC	Lung metastasis (segment 9 right)	3DCRT (9 beams, FFF)	15 Gy ×3	65%
H1	f, 56	Breast cancer	Liver metastasis (segment IVa/VIII)	IMRT – VMAT (3 arcs, FF)	7.5 Gy ×8	80%
H2	m, 76	HCC	Liver metastasis (segment VII)	IMRT – VMAT (2 arcs, FF)	15 Gy ×3	65%
H3	m, 74	HCC	Liver metastasis (segment VI)	IMRT – VMAT (2 arcs, FFF)	7.5 Gy ×8	80%
H4	m, 48	Pancreas	Liver metastasis (segment VI/VII)	IMRT – VMAT (2 arcs, FFF)	15 Gy ×3	65%
H5	m, 53	Prostate	Liver metastasis (segment VI)	IMRT – VMAT (2 arcs, FF)	6 Gy ×6	80%
H6	m, 62	Pancreatic NEC	Liver metastasis (segment VIII)	IMRT – VMAT (2 arcs, FF)	9 Gy ×3	80%
H7	f, 54	Breast cancer	Liver metastasis (segment IVa)	IMRT – VMAT (2 arcs, FFF)	15 Gy ×3	65%

†*PTV-encompassing isodose. 3DCRT, 3D conformal radiotherapy; FF, flattening filter; FFF, flattening filter free; HCC, Hepatic cellular carcinoma; IMRT, intensity modulated radiotherapy; NSCLC, Non-small lung cancer; PTV, Planning Target Volume; VMAT, volumetric arc therapy.*

risk such as the gastrointestinal tract requiring individual dose fractionations between 3 and 8 fractions.

Daily Position Correction

Despite correct alignment of patients' surface to the reference surface in DIBH, a subsequent CBCT-based image correction was required for most treatment fractions. These couch shifts ranged from −7 to 9 mm, −25 to 12 mm, and −13 to 12 mm in lateral, longitudinal and vertical direction, respectively. Nevertheless, median values of all shift directions of first CBCT were close to zero. Grouping patients by target location revealed higher shifts for tumors in the liver compared to lung targets. Highest ranges were observed for shifts in cranial-caudal direction along the *y*-axis (**Figure 2A**).

The mean lengths of applied position correction vectors were 5.4 mm (2.0–10.2) and 8.8 mm (1.7–26.9) for lung and liver targets, respectively. Compared to liver targets, the correction vectors of lung targets were significantly smaller ($p = 0.0172$, Welch's *t*-test, **Figure 2B**).

A second position verification by CBCT in DIBH utilizing the newly captured reference surface could confirm a correct set-up in most cases. A new couch shift was required in 1 of 9 (11%) and 7 of 34 (21%) treatment sessions for lung and liver targets, respectively. Despite higher rates for liver SBRT, there was no significant difference in the frequency of the need for a second position correction for lung and liver targets ($p = 1.0$, Fisher's exact test).

Intra- and Interfractional Variability

The planning CT and the two CBCTs acquired for every treatment session and each patient were retrospectively analyzed for intra- and interfractional difference. **Figure 3** shows the calculated spatial distances from an individually selected, breathing-independent point (mostly bony structures of the spine at the level of the target) to another individually chosen, breathing-dependent point near the gross tumor (mostly surgical clips or a vessel branching) for each patient and CBCT.

In case of a perfect match, the distances between the breathing-independent point and the breathing-dependent point were equal for every imaging series (planning CT and CBCT). Equality of distances of first and second CBCT as well as of planning CT and first CBCT would correspond to no intra- and no interfractional difference, respectively.

Intra-fraction distances of both lung and liver targets were close to the line of equality (**Figure 3A**). On average, the intrafractional differences were 1.6 mm (-3.9 to 0.5) and 1.2 mm (-27.8 to 28.3) for targets in the lungs and liver, respectively. Despite higher ranges for liver targets compared to lung targets the interquartile ranges (IQR) for both lung and liver targets were similar (2.5 mm vs. 2.8 mm). There were no significant differences of intrafractional spatial distances between lung and liver target location ($p = 0.12$, Welch's *t*-test).

In contrast, mean interfractional differences for lung targets were 0.9 mm (0.7–2.1; IQR 1.7 mm), whereas in liver targets the average difference was 3.8 mm (−32.7 to 13.8) with a higher IQR of 6.8 mm (**Figure 3B**). These interfractional differences were statistically significant ($p = 0.01$, Welch's *t*-test), indicating higher variability between treatment sessions for liver SBRT compared to lung SBRT.

Differences in Treated Volumes

The mean gross tumor volumes (GTV) of lung targets, liver targets and all lesions were 1.9, 67, and 48 ml, respectively, with significantly smaller GTVs for lung targets ($p = 0.018$, Welch's *t*-test). The mean internal target volumes (ITV) which comprise the clinical target volumes (CTV) with additional margins obtained from 4D-CT data in free-breathing was 108 ml for all lesions. By using DIBH these margins could be omitted and the CTV in DIBH was 28% smaller (77 ml, $p < 0.01$, paired *t*-test, **Figure 4A**). Consequently, planning target volumes of all targets could be reduced significantly by 26% from an average of 148 ml in free-breathing (FB) to 110 ml utilizing DIBH ($p < 0.01$, paired *t*-test). This reduction of target volumes by using DIBH was seen for both lung and liver targets. However, statistical significance was not reached for the smaller sub-group of lung targets (**Figure 4B**).

DISCUSSION

We share a new work-flow for SBRT in DIBH by combining SGRT and IGRT in patients with targets in the lungs and liver and demonstrate feasibility. Utilizing DIBH for SBRT significantly reduced the irradiated volumes compared to free-breathing treatment plans. SGRT enabled initial couch shifts based on first CBCT in DIBH to be close to zero, but the ranges and absolute correction vectors of liver targets were significantly higher when compared to lung targets. This interfractional variability resulted in a daily need for position correction and for an updated reference surface for most patients for every treatment session. Intrafractional movements on the other hand, were quite low with second CBCTs in DIBH rarely requiring correction of patient positioning in both liver and lung SBRT. Most liver SBRT patients exhibited reproducible intrafractional positioning in DIBH. In contrast to targets in the liver, both intra- and interfractional variability of lung targets was small indicating a better correlation of patients' surface to lung targets rather than to liver targets.

Deep inspiration breath-hold was introduced many years ago for radiotherapy of left sided breast cancer patients to reduce the irradiated heart volume by enlarging the distance of the heart to the chest well (15). Reproducibility was ensured by the active breathing coordinator^{TM} (ABC^{TM}, Elekta, Stockholm, Sweden), a commercial device consisting of a nose clamp and a mouthpiece that is connected to a breathing tube containing a valve which closes once a pre-defined target air volume is inhaled by the patient. The prospective United Kingdom HeartSpare Study compared this ABC^{TM}-assisted breath-hold with a voluntary breath-hold technique and found that the latter was non-inferior in terms of reproducibility and normal tissue sparing. Moreover, voluntary breath-hold was faster and preferred by both patients and therapists (16).

Surface-guided radiotherapy describes the use of commercially available surface imaging solutions that were

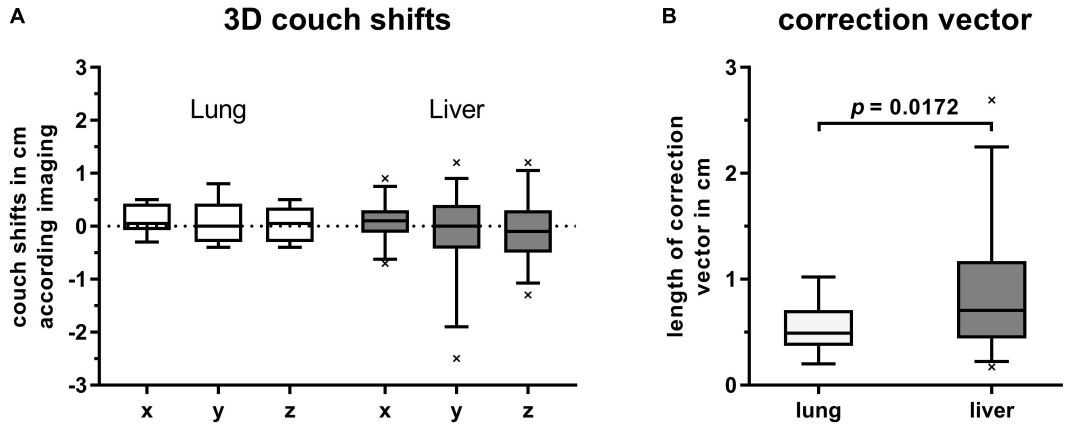


FIGURE 2 | Box plots with boxes extending from 25th to 75th percentiles. Whiskers are drawn down to 5th and up to 95th percentiles, respectively. **(A)** Couch shifts are shown in left (+) to right (–), cranial (+) to caudal (–) and anterior (+) to posterior (–) directions along the *x*-, *y*-, and *z*-axis, respectively. Values were grouped by target location and derived from first CBCTs for position verification after alignment of current DIBH surface to that of the reference surface. **(B)** Length of correction vectors of first CBCTs for image-guidance in DIBH and grouped by target location. Statistical difference was calculated by Welch's *t*-test.

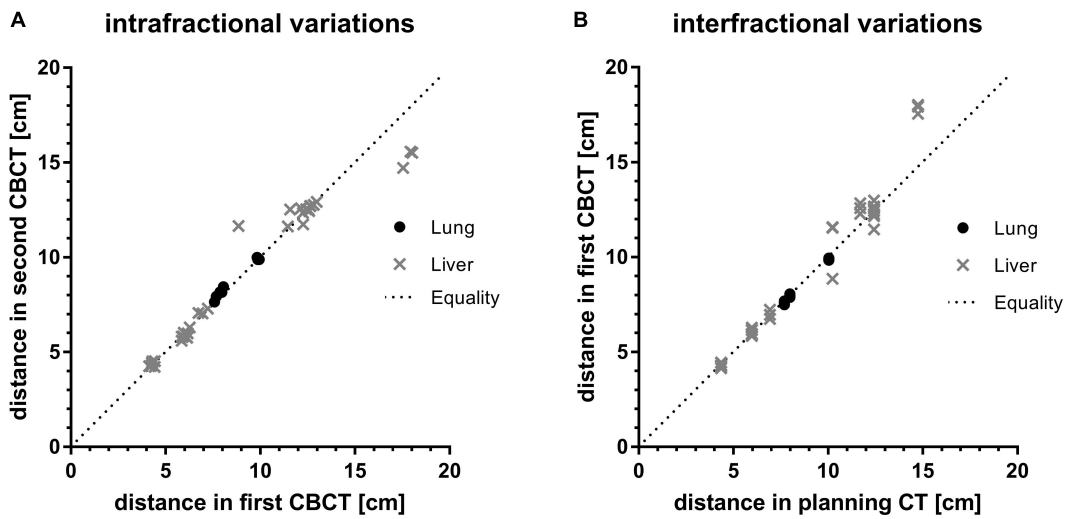


FIGURE 3 | Assessment of **(A)** intra-and **(B)** interfractional variations by comparison of absolute distances between a breathing independent point to a breathing dependent movable point of first compared to second CBCT and of planning CT compared to first CBCT for every fraction, respectively. Values are grouped by target location. The dotted line indicates theoretical equality of distances in imaging.

primarily developed to assist with patient set-up before radiotherapy delivery. These tools precisely indicate the spatial difference of a region of interest (ROI) on the patients' body surface and a corresponding reference surface generated from the external contour of the planning simulation CT, facilitating a fast and accurate patient set-up with six degrees of freedom (17). SGRT is non-invasive, patients are not exposed to additional radiation dose, it basically does not rely on skin marks and allows monitoring of DIBH with most data available for SGRT with DIBH for adjuvant radiotherapy in left sided breast-cancer (18). Since SGRT allows not only monitoring of the patient-setup and breathing before, but also during treatment it is of particular interest for SBRT to ensure a safe, accurate and precise dose delivery during the entire treatment session.

For SBRT in targets that move with respiration, an abdominal compression is traditionally employed to mitigate breathing motion and to decrease the residual range of motion, as assessed during time resolved CT acquisition. These 4D-CT data are used to create a PTV that encompasses the tumor during the entire breathing cycle or to define a specific gating window which results in reduced irradiation volumes but, on the other hand, extend treatment delivery time (4). Several reports, however, have commented that the 4D-CT information from planning is not necessarily representative of the motion amplitude during treatment (19, 20). Moreover, breathing patterns may change both within and between treatments (intra- and interactionally) (21, 22). Therefore, we tested the feasibility of combining SGRT with IGRT to enable SBRT in DIBH.

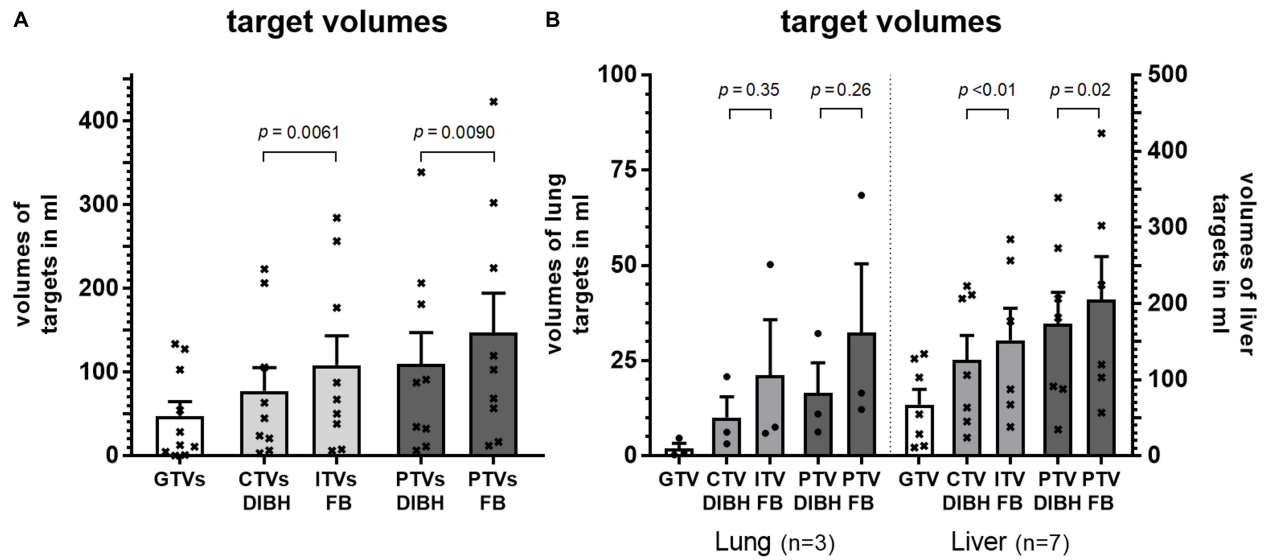


FIGURE 4 | Volumes of gross tumor (GTV), clinical target volume (CTV), internal target volume (ITV) and planning target in DIBH (CTV DIBH and PTV DIBH) as well as in free-breathing (ITV FB and PTV FB) for **(A)** all patients and **(B)** grouped by target location. Statistical differences were calculated by paired *t*-test.

Earlier data on SGRT and SBRT in free-breathing suggests that the pre-imaging treatment set-up for SBRT can be improved by SGRT compared to in-room laser localization of skin tattoos or skin marks (23). Moreover, SGRT reliably detects intrafractional shifts during treatment delivery when deviations extend 2 mm, as confirmed by CBCT (24). In addition, DIBH improves image quality and reduces craniocaudal registration uncertainties compared to free-breathing in lung cancer radiotherapy (25). Furthermore, using breath-hold for SBRT delivery reduces motion artifacts that is especially important for small lung tumors that are poorly visualized on imaging even with modern linear accelerators (26). Yet to perform IGRT in combination with DIBH, imaging within a breath-hold of approximately 30 s was required. We achieved this through our technique by not performing a complete gantry rotation and instead using a higher than standard gantry rotation speed without appreciable loss of image information. As previously suggested, faster CBCTs show no significant registration differences compared to standard CBCTs and a higher gantry velocity with fewer projections produces fewer reconstruction artifacts (27).

We observed highest interfractional variability in cranio-caudal direction and lowest in left-right direction, underpinning the idea that respiratory motion, which is mostly performed by the diaphragm, impacts most on target localization. Indeed, using SGRT for positioning of breast cancer patients in free-breathing showed least errors for lateral set-up compared to imaging (28). Our data further suggest a higher correlation of thoracic DIBH surface to intrathoracic targets than to abdominal targets, despite the fact that we chose to set the ROI for DIBH monitoring to the lower thorax for both thoracic and abdominal targets. Indeed, the correlation of skin to tumor is not necessarily constant, especially for liver and pancreas as previously reported (29, 30). In addition, lung volumes in DIBH may not necessarily always be the same for every breath-hold, although SGRT may confirm a match within a ROI on the patient's body surface to the corresponding reference surface. Such variations in lung volumes during breath-hold were recently reported to have an impact on target localization (31).

Previous work on employing active breathing coordinator (ABC^{TM})-controlled breath-hold for lung and liver SBRT reported good intrafractional reproducibility of liver position in the majority of patients. However, interfractional reproducibility was worse, suggesting a need for daily image guidance (31–33). Another study, conducted by Lu et al., also found higher interfractional than intrafractional motion but observed clinically significant intrafractional motion >3 mm in 26 and 47% of patients with liver and lung cancer, respectively (34). This intrafractional differences could be explained by an intra-breath-hold residual motion of the diaphragm that was recently estimated by an ultrasound-based monitoring of the diaphragm dome during ABC^{TM}-controlled breath-holds to be <2 mm and <5 mm in 59 and 95% of 385 DIBHs in 13 patients (35).

The most elegant solution for precise and accurate SBRT in DIBH is probably possible with MR-guidance since on-board magnetic resonance (MR) imaging in linear accelerators (MR-LINAC) for set-up and treatment delivery is non-invasive, exposes the patient to no additional radiation and allows for direct target localization and real-time visualization (36). In contrast, SGRT remains an indirect visualization of the patient's surface although it gives feedback of localization in real-time, too. Nevertheless, MR-LINACs are currently rarely available and there are still some issues concerning reliability of gating and tracking procedures, the additional time needed for dose-optimization and the dose delivery time, that is yet mostly slower than in conventional linear accelerators (37). Thus, a SBRT session in DIBH at a MR-LINAC would require a patient to perform more breath-holds. In addition, a MR LINAC treatment is not an

option for every patient depending on body size, claustrophobia and metal implants or implanted electronic devices.

CONCLUSION

Surface-guided radiotherapy in DIBH for lung and liver tumors using a combination of SGRT and IGRT is feasible. This approach is easy to incorporate in contemporary practice and does not require any breathing tubes connected to the patient. Nevertheless, daily 3D imaging is of paramount importance given that interfractional variability is high, particularly in DIBH for liver SBRT.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because legal regulations disallow and as a requirement of local ethics committee clinical raw data must not be made available to others than those involved in the study. Requests to access the datasets should be directed to patrick.naumann@med.uni-heidelberg.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission Heidelberg. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PN and VB were responsible for conceptualization, data curation and formal analysis. BR and JD supervised the project. PN drafted the manuscript. PN, VB, JF, BR, and ET-M participated in patient treatment. BF and JL critically reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

This research received no external funding, but we acknowledge financial support for parts of publication costs by Deutsche Forschungsgemeinschaft within the funding program Open Access Publishing, by the Baden-Württemberg Ministry of Science, Research and the Arts and by Ruprecht-Karls-Universität Heidelberg.

ACKNOWLEDGMENTS

We thank all involved patients as well as all therapists for their patience. Without their support the study would not have succeeded.

REFERENCES

1. Salama JK, Kirkpatrick JP, Yin FF. Stereotactic body radiotherapy treatment of extracranial metastases. *Nature Rev Clin Oncol.* (2012) 9:654–65. doi: 10.1038/nrclinonc.2012.166
2. Tree AC, Khoo VS, Eeles RA, Ahmed M, Dearnaley DP, Hawkins MA, et al. Stereotactic body radiotherapy for oligometastases. *Lancet Oncol.* (2013) 14:e28–37.
3. Guckenberger M, Baus WW, Blanck O, Combs SE, Debus J, Engenhart-Cabillic R, et al. Definition and quality requirements for stereotactic radiotherapy: consensus statement from the DEGRO/DGMP working group stereotactic radiotherapy and radiosurgery. *Strahlenther Onkol.* (2020) 196:417–20. doi: 10.1007/s00066-020-01603-1
4. Dieterich S, Green O, Booth J. SBRT targets that move with respiration. *Phys Med.* (2018) 56:19–24. doi: 10.1016/j.ejmp.2018.10.021
5. Landberg T, Chavaudra J, Dobbs J, Gerard J, Hanks G, Horiot J, et al. Report 62. *J ICRU.* (1999) os32.
6. Keall PJ, Mageras GS, Balter JM, Emery RS, Forster KM, Jiang SB, et al. The management of respiratory motion in radiation oncology report of AAPM task group 76. *Medical Phys.* (2006) 33:3874–900. doi: 10.1118/1.2349696
7. Herfarth KK, Debus J, Lohr F, Bahner ML, Fritz P, Hoss A, et al. Extracranial stereotactic radiation therapy: set-up accuracy of patients treated for liver metastases. *Int J Radiation Oncol Biol Phys.* (2000) 46:329–35. doi: 10.1016/s0360-3016(99)00413-7
8. Bouilhol G, Ayadi M, Rit S, Thengumpallil S, Schaerer J, Vandemeulebroucke J, et al. Is abdominal compression useful in lung stereotactic body radiation therapy? A 4DCT and dosimetric lobe-dependent study. *Phys Med.* (2013) 29:333–40. doi: 10.1016/j.ejmp.2012.04.006
9. Brandner ED, Chetty IJ, Giaddui TG, Xiao Y, Huq MS. Motion management strategies and technical issues associated with stereotactic body radiotherapy of thoracic and upper abdominal tumors: a review from NRG oncology. *Medical Phys.* (2017) 44:2595–612. doi: 10.1002/mp.12227
10. Schmitt D, Blanck O, Gauer T, Fix MK, Brunner TB, Fleckenstein J, et al. Technological quality requirements for stereotactic radiotherapy : expert review group consensus from the DGMP working group for physics and technology in stereotactic radiotherapy. *Strahlenther Onkol.* (2020) 196:421–43. doi: 10.1007/s00066-020-01583-2
11. Boda-Heggemann J, Knopf AC, Simeonova-Chergou A, Wertz H, Stieler F, Jahnke A, et al. Deep inspiration breath hold-based radiation therapy: a clinical review. *Int J Radiation Oncol Biol Phys.* (2016) 94:478–92. doi: 10.1016/j.ijrobp.2015.11.049
12. Navarria P, Ascolese AM, Mancosu P, Alongi F, Clerici E, Tozzi A, et al. Volumetric modulated arc therapy with flattening filter free (FFF) beams for stereotactic body radiation therapy (SBRT) in patients with medically inoperable early stage non small cell lung cancer (NSCLC). *Radiother Oncol.* (2013) 107:414–8. doi: 10.1016/j.radonc.2013.04.016
13. Alderliesten T, Sonke JJ, Betgen A, Honnef J, van Vliet-Vroegindeweij C, Remeijer P. Accuracy evaluation of a 3-dimensional surface imaging system for guidance in deep-inspiration breath-hold radiation therapy. *Int J Radiation Oncol Biol Phys.* (2013) 85:536–42. doi: 10.1016/j.ijrobp.2012.04.004
14. Wong VY, Tung SY, Ng AW, Li FA, Leung JO. Real-time monitoring and control on deep inspiration breath-hold for lung cancer radiotherapy–combination of ABC and external marker tracking. *Medical Phys.* (2010) 37:4673–83. doi: 10.1118/1.3476463
15. Sixel KE, Aznar MC, Ung YC. Deep inspiration breath hold to reduce irradiated heart volume in breast cancer patients. *Int J Radiation Oncol Biol Phys.* (2001) 49:199–204. doi: 10.1016/s0360-3016(00)01455-3
16. Bartlett FR, Colgan RM, Carr K, Donovan EM, McNair HA, Locke I, et al. The UK HeartSpare study: randomised evaluation of voluntary deep-inspiratory breath-hold in women undergoing breast radiotherapy. *Radiother Oncol.* (2013) 108:242–7. doi: 10.1016/j.radonc.2013.04.021
17. Schoffel PJ, Harms W, Sroka-Perez G, Schlegel W, Karger CP. Accuracy of a commercial optical 3D surface imaging system for realignment of patients for

radiotherapy of the thorax. *Phys Med Biol.* (2007) 52:3949–63. doi: 10.1088/0031-9155/52/13/019
18. Hoisak JDP, Pawlicki T. The role of optical surface imaging systems in radiation therapy. *Semin Radiat Oncol.* (2018) 28:185–93. doi: 10.1016/j.semradonc.2018.02.003
19. Worm ES, Hoyer M, Fledelius W, Hansen AT, Poulsen PR. Variations in magnitude and directionality of respiratory target motion throughout full treatment courses of stereotactic body radiotherapy for tumors in the liver. *Acta Oncol.* (2013) 52:1437–44. doi: 10.3109/0284186x.2013.813638
20. Park JC, Park SH, Kim JH, Yoon SM, Song SY, Liu Z, et al. Liver motion during cone beam computed tomography guided stereotactic body radiation therapy. *Medical Phys.* (2012) 39:6431–42. doi: 10.1118/1.4754658
21. Ge J, Santanam L, Noel C, Parikh PJ. Planning 4-dimensional computed tomography (4DCT) cannot adequately represent daily intrafractional motion of abdominal tumors. *Int J Radiation Oncol Biol Phys.* (2013) 85:999–1005. doi: 10.1016/j.ijrobp.2012.09.014
22. Dhont J, Vandemeulebroucke J, Burghelea M, Poels K, Depuydt T, Van Den Begin R, et al. The long- and short-term variability of breathing induced tumor motion in lung and liver over the course of a radiotherapy treatment. *Radiother Oncol.* (2018) 126:339–46. doi: 10.1016/j.radonc.2017.09.001
23. Leong B, Padilla L. Impact of use of optical surface imaging on initial patient setup for stereotactic body radiotherapy treatments. *J Appl Clin Med Phys.* (2019) 20:149–58. doi: 10.1002/acm2.12779
24. Heinzerling JH, Hampton CJ, Robinson M, Bright M, Moeller BJ, Ruiz J, et al. Use of surface-guided radiation therapy in combination with IGRT for setup and intrafraction motion monitoring during stereotactic body radiation therapy treatments of the lung and abdomen. *J Appl Clin Med Phys.* (2020) 21:48–55. doi: 10.1002/acm2.12852
25. Josipovic M, Persson GF, Bangsgaard JP, Specht L, Aznar MC. Deep inspiration breath-hold radiotherapy for lung cancer: impact on image quality and registration uncertainty in cone beam CT image guidance. *Br J Radiol.* (2016) 89:20160544. doi: 10.1259/bjr.20160544
26. Peng Y, Vedam S, Chang JY, Gao S, Sadagopan R, Bues M, et al. Implementation of feedback-guided voluntary breath-hold gating for cone beam CT-based stereotactic body radiotherapy. *Int J Radiation Oncol Biol Phys.* (2011) 80:909–17. doi: 10.1016/j.ijrobp.2010.08.011
27. Arns A, Wertz H, Boda-Heggemann J, Schneider F, Blessing M, Abo-Madyan Y, et al. Ultrafast single breath-hold cone-beam CT lung cancer imaging with faster linac gantry rotation. *Radiother Oncol.* (2019) 135:78–85. doi: 10.1016/j.radonc.2019.02.004
28. Hattel SH, Andersen PA, Wahlstedt IH, Damkjaer S, Saini A, Thomsen JB. Evaluation of setup and intrafraction motion for surface guided whole-breast cancer radiotherapy. *J Appl Clin Med Phys.* (2019) 20:39–44. doi: 10.1002/acm2.12599
29. Velec M, Moseley JL, Craig T, Dawson LA, Brock KK. Accumulated dose in liver stereotactic body radiotherapy: positioning, breathing, and deformation effects. *Int J Radiation Oncol Biol Phys.* (2012) 83:1132–40. doi: 10.1016/j.ijrobp.2011.09.045
30. Minn AY, Schellenberg D, Maxim P, Suh Y, McKenna S, Cox B, et al. Pancreatic tumor motion on a single planning 4D-CT does not correlate with intrafraction tumor motion during treatment. *Am J Clin Oncol.* (2009) 32:364–8. doi: 10.1097/coc.0b013e31818da9e0
31. Lee S, Zheng Y, Podder T, Biswas T, Verma V, Goss M, et al. Tumor localization accuracy for high-precision radiotherapy during active breath-hold. *Radiother Oncol.* (2019) 137:145–52. doi: 10.1016/j.radonc.2019.04.036
32. Eccles C, Brock KK, Bissonnette JP, Hawkins M, Dawson LA. Reproducibility of liver position using active breathing coordinator for liver cancer radiotherapy. *Int J Radiation Oncol Biol Phys.* (2006) 64:751–9. doi: 10.1016/j.ijrobp.2005.05.066
33. Brock J, McNair HA, Panakis N, Symonds-Tayler R, Evans PM, Brada M. The use of the active breathing coordinator throughout radical non-small-cell lung cancer (NSCLC) radiotherapy. *Int J Radiation Oncol Biol Phys.* (2011) 81:369–75.
34. Lu L, Diaconu C, Djemil T, Videtic GM, Abdel-Wahab M, Yu N, et al. Intra- and inter-fractional liver and lung tumor motions treated with SBRT under active breathing control. *J Appl Clin Med Phys.* (2018) 19:39–45.
35. Vogel L, Sihono DSK, Weiss C, Lohr F, Stieler F, Wertz H, et al. Intra-breath-hold residual motion of image-guided DIBH liver-SBRT: an estimation by ultrasound-based monitoring correlated with diaphragm position in CBCT. *Radiother Oncol.* (2018) 129:441–8.
36. Klüter S. Technical design and concept of a 0.35 T MR-Linac. *Clin Transl Radiat Oncol.* (2019) 18:98–101.
37. Corradini S, Alongi F, Andratschke N, Belka C, Boldrini L, Cellini F, et al. MR-guidance in clinical reality: current treatment challenges and future perspectives. *Radiation Oncol.* (2019) 14:92.

Conflict of Interest: In the past 5 years JD attended advisory board meetings of MERCK KGaA (Darmstadt), for which the University hospital Heidelberg received travel grants and honoraria. His department further received funding for research projects and for educational grants to the University hospital of Heidelberg by Accuray (2016), Merck KGaA (2015-open), Siemens GmbH (2015-open), Viewray (2018-open), Vision RT (2017-open). He is CEO of the HIT GmbH and also member of the kuratorium of the Physikalisch Technische Bundesanstalt (PTB). As chair of HIRO (Heidelberg Institute of Radiation Oncology) and director of the NCT (National Center for Tumor Diseases) he is responsible for collaborations with a multitude of companies and institutions. JD confirms that to the best of his knowledge none of the above funding sources was involved in the data analysis and preparation of this work.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

MINI REVIEW
published: 06 October 2020
doi: 10.3389/fonc.2020.565086

The Role of 18F-FDG PET/CT in Guiding Precision Medicine for Invasive Bladder Carcinoma

Antoine Girard[1], Helena Vila Reyes[2], Hiram Shaish[3], Jean-François Grellier[4], Laurent Dercle[5], Pierre-Yves Salaün[6], Olivier Delcroix[6] and Mathieu Rouanne[7]**

[1] Department of Nuclear Medicine, Centre Eugène Marquis, Université Rennes 1, Rennes, France, [2] Department of Urology, Columbia University Irving Medical Center – New York Presbyterian Hospital, New York, NY, United States, [3] Department of Radiology, Columbia University Medical Center, New York, NY, United States, [4] Centre Cardiologique du Nord, Nuclear Medicine, Saint-Denis, France, [5] Department of Radiology, New York Presbyterian Hospital – Columbia University Medical Center, New York, NY, United States, [6] Department of Nuclear Medicine, Centre Hospitalier Régional Universitaire de Brest, Brest cedex, France, [7] Department of Urology, Hôpital Foch, Université Versailles-Saint-Quentin-en-Yvelines, Université Paris-Saclay, Suresnes, France

OPEN ACCESS

Edited by:
Sridhar Nimmagadda, Johns Hopkins University, United States

Reviewed by:
Gad Abikhzer, Jewish General Hospital, Canada
Luigi Aloj, University of Cambridge, United Kingdom

***Correspondence:**
Antoine Girard
a.girard@rennes.unicancer.fr
Mathieu Rouanne
rouanne.mathieu@gmail.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *23 May 2020*
Accepted: *09 September 2020*
Published: *06 October 2020*

Citation:
Girard A, Vila Reyes H, Shaish H, Grellier J-F, Dercle L, Salaün P-Y, Delcroix O and Rouanne M (2020) The Role of 18F-FDG PET/CT in Guiding Precision Medicine for Invasive Bladder Carcinoma. Front. Oncol. 10:565086. doi: 10.3389/fonc.2020.565086

Bladder cancer (BC) is the 10th most common cancer worldwide. Approximately one quarter of patients with BC have muscle-invasive disease (MIBC). Muscle-invasive disease carries a poor prognosis and choosing the optimal treatment option is critical to improve patients' outcomes. Ongoing research supports the role of 2-deoxy-2-(18F)fluoro-D-glucose positron emission tomography (18F-FDG PET) in guiding patient-specific management decisions throughout the course of MIBC. As an imaging modality, 18F-FDG PET is acquired simultaneously with either computed tomography (CT) or MRI to offer a hybrid approach combining anatomical and metabolic information that complement each other. At initial staging, 18F-FDG PET/CT enhances the detection of extravesical disease, particularly in patients classified as oligometastatic by conventional imaging. 18F-FDG PET/CT has value in monitoring response to neoadjuvant and systemic chemotherapy, as well as in localizing relapse after treatment. In the new era of immunotherapy, 18F-FDG PET/CT may also be useful to monitor treatment efficacy as well as to detect immune-related adverse events. With the advent of artificial intelligence techniques such as radiomics and deep learning, these hybrid medical images can be mined for quantitative data, providing incremental value over current standard-of-care clinical and biological data. This approach has the potential to produce a major paradigm shift toward data-driven precision medicine with the ultimate goal of personalized medicine. In this review, we highlight current literature reporting the role of 18F-FDG PET in supporting personalized management decisions for patients with MIBC. Specific topics reviewed include the incremental value of 18F-FDG PET in prognostication, pre-operative planning, response assessment, prediction of recurrence, and diagnosing drug toxicity.

Keywords: PET – Positron Emission Tomography, bladder cancer, muscle invasive bladder cancer, immunotherapy, staging

INTRODUCTION

Bladder cancer (BC) is the 10th most common cancer worldwide, with approximately half a million new cases diagnosed globally and 200,000 related deaths per year (1). At diagnosis, 75% of the patients have non-muscle invasive BC whereas the remaining 25% have muscle invasive disease (MIBC). While non-muscle invasive BC is characterized by frequent recurrence (50–70%) but a relatively low propensity to progress (10–15%), MIBC has a poor prognosis with high rates of metastasis and 5-year survival <50% despite radical surgery (2).

The current standard treatment for MIBC is based on radical cystectomy (RC) with prior cisplatin-based neoadjuvant chemotherapy (NAC) in eligible patients (cT2-T4aN0M0) (3). For patients with more advanced stage disease or for those who recur after radical surgery, cisplatin-based combination chemotherapy remains the standard of care for first-line systemic treatment (3).

Recent advances in the field of immunotherapy are reshaping the therapeutic landscape for patients with BC. Specifically, immune checkpoints inhibitors (ICIs) have demonstrated promising results in both localized and metastatic settings (4–6). To date, five anti-PD(L)1 monoclonal antibodies have been approved by the US Food and Drug Administration in the second-line setting. This current shift in treatment strategy has created an unmet need to re-evaluate the clinical use of existing imaging techniques.

Although traditionally imaging of patients with BC has predominantly focused on CT (including CT urography) and MRI, 2-deoxy-2-(18F)fluoro-D-glucose (18F-FDG) positron emission tomography/computed tomography (PET/CT) may have the potential to offer additional diagnostic information due to its unique ability to image metabolism. Indeed, urothelial carcinoma, much like many other solid tumors, is characterized by alterations of glucose metabolism and overexpression of glucose transporters (GLUT-1 and GLUT-3) (7).

This review presents the most recent advances in 18F-FDG PET-guided personalized medicine in the context of MIBC. Through a review of the current literature, we present the potential clinical value of 18F-FDG PET/CT and PET/MRI for risk stratification at diagnosis, monitoring of treatment response, and detection of recurrence during follow-up.

RISK STRATIFICATION AT DIAGNOSIS

Data supporting the potential role of 18F-FDG PET/CT for initial staging of MIBC are continuously increasing. Current data suggest that 18F-FDG PET/CT improves the detection of extravesical disease and can therefore substantially improve management decisions.

Tumor Detection

2-Deoxy-2-(18F)fluoro-D-glucose positron emission tomography/computed tomography performance for tumor detection in the bladder is hampered by urinary excretion of 18F-FDG (8). Several studies have proposed using adapted protocols (hyperhydration, forced diuresis and refilling, or filling the bladder in a retrograde manner). Utilizing these techniques, authors report sensitivities between 50 and 96% (9–17). However, to date, 18F-FDG PET/CT has not be shown to improve primary tumor detection and staging, when compared to cystoscopy and morphological imaging alone performed with CT and especially MRI (18).

Lymph Node Staging

Currently, the benefit of 18F-FDG PET/CT over contrast-enhanced CT (CECT) alone for lymph node (LN) staging remains controversial, since both modalities have excellent specificity but relatively poor sensitivity.

A recent meta-analysis including 14 studies and 785 patients reported that the pooled sensitivity and specificity of 18F-FDG PET/CT for initial pelvic LN staging, in a per patient analysis, were 57% [95% CI 49–64%] and 92% [95% CI 87–95%], respectively (19). One major limitation of the current literature is the heterogeneous methodologies across published studies, notably regarding study designs, inclusion criteria, administration of NAC, injection–acquisition time in PET, the use of forced diuresis, the administration of contrast media for PET/CT, and interpretation criteria for both CECT and PET/CT (20). Studies comparing performance of CT to hybrid 18F-FDG PET/CT for pelvic LN staging with pathological analysis of extended pelvic LN dissection samples as a reference are presented in **Table 1**. Pooled sensitivities for CT and PET/CT are 38% [95% CI 29–47%] and 52% [95% CI 45–60%], respectively, while the pooled specificities are 91% [95% CI 88–94%] and 92% [95% CI 89–95%], respectively. Few studies have suggested a potential role for metabolic analysis by ruling out nodal disease in PET negative enlarged pelvic LNs (21, 22) (**Figure 1A**). Moreover, higher standardized uptake values (SUVmax) in LNs has been correlated with higher recurrence risk, independent of pathological findings (23).

To our knowledge, there is only a single study of 18 patients comparing 18F-FDG PET/CT and conventional MRI for pelvic LN staging in MIBC. With pathology as the gold standard, authors reported sensitivity and specificity of 80% and 80% for MRI, and 93 and 88% for 18F-FDG PET/CT, respectively. There was no statistically significant difference. However, this study was limited by small sample size and the lack of multiparametric MRI sequences such as diffusion-weighted imaging (DWI) (24).

Distant Metastatic Staging

Along with detecting nodal involvement, 18F-FDG PET/CT's greatest strength is probably in detecting distant metastases during initial staging. In a patient-based analysis, 18F-FDG PET/CT sensitivity ranges between 54 and 87%, while specificity ranges between 90 and 97% for the detection of distant metastases from BC (15, 25–27). One study reported 18F-FDG PET/CT to be more sensitive than CT alone for detection of distant metastases, with sensitivities of 54 versus 41%, respectively. Both modalities showed similar very high specificities of 97 and 98%, respectively (26). Thus, 18F-FDG PET/CT revealed more lesions suspected to be metastasis or second primary cancer than conventional imaging (26, 28–30). 18F-FDG PET/CT changed management

TABLE 1 | CT alone and 18F-FDG PET/CT performances for LN staging, with pathology as a gold standard, without preoperative chemotherapy.

First author	Year	CT criteria	PET/CT criteria	Sensitivity		Specificity		Accuracy	
				CT	PET/CT	CT	PET/CT	CT	PET/CT
Girard (22)	2019	SA > 8 mm	SUV_{max} > 4, and/or SA > 10 mm, and/or SUV_{max} > 2 and SA > 8 mm	7/17	8/17	38/44	42/44	45/61	50/61
Pichler (91)	2017	SA > 8 mm	SA > 10mm, and/or PET subjective analysis	5/11	7/11	54/59	52/59	59/70	59/70
Uttam (92)	2016	SA > 10 mm	SUV_{max} > 2.5	3/3	3/3	6/12	7/12	9/15	10/15
Jeong (93)	2015	SA > 10 mm or necrosis	SUV_{max} > 2.5	5/17	8/17	43/44	41/44	48/61	49/61
Aljabery (94)	2015	LA ≥ 10 mm	SUV_{max} > 2.5	7/17	7/17	33/37	32/37	40/54	39/54
Rouanne (95)	2014	SA > 10 mm	PET subjective analysis	NA	13/26	NA	74/76	NA	87/102
Goodfellow (26)	2014	SA > 8 mm	SA > 8 mm and/or PET subjective analysis	13/28	19/28	64/65	62/65	77/93	81/93
Hitier-Berthault (96)	2013	LA > 10 mm	PET subjective analysis	2/22	8/22	27/30	26/30	29/52	34/52
Swinnen (97)	2010	NA	PET subjective analysis	6/13	6/13	35/38	37/38	41/51	43/51
Kibel (33)	2009	NA	PET subjective analysis	NA	7/10	NA	30/32	NA	37/42
Pooled [95% CI]				**37.5% (48/128) [29.1–46.5%]**	**52.4% (86/164) [44.5–60.3%]**	**91.2% (300/329) [87.6–94.0%]**	**92.2% (403/437) [89.3–94.6%]**	**76.1% (348/457) [72.0–80.0%]**	**81.4% (489/601) [78.0–84.4%]**

SA, short axis; LA, long axis; LN, lymph node; SUV_{max}, maximum standardized uptake value; NA, not available.

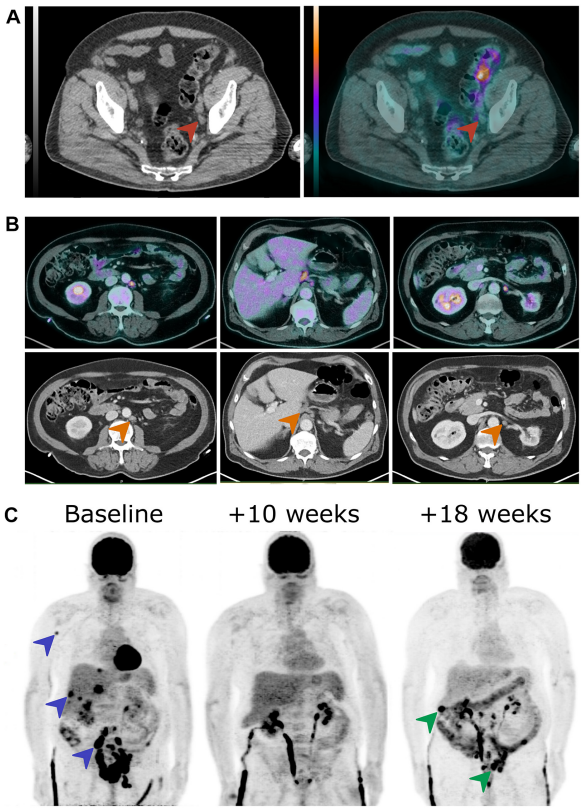


FIGURE 1 | 18F-FDG PET/CT images of three patients at different time points in the course of MIBC. **(A)** A 76-year-old man presented with an enlarged external iliac lymph node at initial staging without any 18F-FDG uptake (red arrowheads). The N0 status was confirmed by pathological examination of pelvic lymph node dissection sample. **(B)** A 67-year-old man was oligometastatic with one enlarged latero-aortic lymph node on CT at initial staging and 18F-FDG PET/CT revealed several other retroperitoneal lymph node metastases (orange arrowheads). **(C)** In a 62-year-old man who presented MIBC with osseous, hepatic, and nodal metastases at diagnosis (blue arrowheads), 18F-FDG PET/CT demonstrated complete metabolic response after three cycles of chemotherapy and then a locoregional and hepatic relapse after the fifth chemotherapy cycle (green arrowheads). 18F-FDG PET/CT, 2-deoxy-2-(18F)fluoro-D-glucose positron emission tomography/computed tomography; MIBC, muscle invasive bladder cancer.

over conventional imaging in a range of 18–68% of patients (25, 29, 31, 32), and resulted in less additional tests in 70% of patients (25). The presence of FDG-avid regional LNs or extra pelvic lesions was an independent predictor of overall survival (28, 33), whereas it was not statistically significant for the counterpart conventional CECT findings (28).

In conclusion, the current literature suggests that 18F-FDG PET/CT provides an incremental value for nodal and distant staging in MIBC at initial diagnosis. However, due to its significant cost for healthcare systems (34), prospective studies with high evidence level are still needed before its use can be formally adopted into consensus guidelines and recommendations (3, 18, 35, 36). In support of its role in distant staging, a recent consensus statement revealed that 18F-FDG PET/CT is the imaging modality of choice to avoid over-treatment in oligometastatic patients with an agreement of 88% of participants (37) (**Figure 1B**).

Metabolic Prognostic Factors

Several prognostic biomarkers can be extracted from 18F-FDG PET using routine clinical workstations. The mainstream prognostic imaging biomarkers are baseline metabolic tumor burden (TMTV), baseline total lesion glycolysis (TLG), and SUVmax. All have been associated with overall survival in a wide range of cancers and treatments (38–40). With the paradigm shift of immunotherapy, authors have evaluated the potential role of appraising host metabolism in lymphoid tissues such as the bone marrow. They have demonstrated that increased bone marrow metabolism is associated with shorter overall survival and systemic immune suppression (41, 42).

In patients with advanced BC, such prognostic factors have not been reported as of yet. Additionally, there are no studies demonstrating that these biomarkers can predict response to chemotherapy or ICIs. In a proof of concept study, Chen et al. suggested that higher 18F-FDG uptake by BC may be associated with elevated PD-1/PD-L1 expression, potentially guiding the decision to select patients for ICIs (43, 44).

RESPONSE EVALUATION

Based on a limited number of studies, 18F-FDG PET/CT appears to be a valuable tool for monitoring response to chemotherapy both in the neoadjuvant and metastatic settings. While the performance of 18F-FDG PET/CT in evaluating tumor sensitivity to immunotherapy has been demonstrated in several solid malignancies, it remains to be established for MIBC.

Neoadjuvant or Induction Treatment

There have been a few studies that have focused on the performance of 18F-FDG PET/CT in monitoring the response of BC to NAC. In terms of primary bladder tumor evaluation after NAC, 18F-FDG PET/CT has demonstrated 75% sensitivity and 90% specificity in identifying patients with complete pathologic response (45). After induction chemotherapy in pelvic LN metastatic patients, responders were distinguished from non-responders with a sensitivity of 83–100% and a specificity of 67–94% with 18F-FDG PET/CT, compared to 88 and 33%, respectively, with conventional CECT. Complete responders were correctly identified with 67–75% sensitivity and 46–90% specificity with 18F-FDG PET/CT, compared to 64 and 60% with CECT, respectively (45–47). These results suggest that 18F-FDG PET/CT may be useful to assess response to neoadjuvant or induction chemotherapy, but is of limited interest to select patients for RC due to low predictive value in detecting residual LN involvement (47, 48).

To our knowledge, there is no published study investigating the role of 18F-FDG PET/CT in evaluating response after preoperative immunotherapy.

Metastatic Bc: Chemotherapy

2-Deoxy-2-(18F)fluoro-D-glucose positron emission tomography/computed tomography is useful for evaluating response to systemic chemotherapy (**Figure 1C**). In one study, 18F-FDG PET/CT using the European Organization for Research and Treatment of Cancer (EORTC) criteria outperformed CT interpretation alone based on Response Criteria in Solid Tumors (RECIST) criteria for the prediction of response to first-line systemic chemotherapy (cisplatin and gemcitabine) (49). Furthermore, early response assessment using 18F-FDG PET/CT predicted progression-free survival and overall survival after two cycles of combination of methotrexate, vinblastine, doxorubicin, and cisplatin (MVAC) in first-line metastatic chemotherapy (50).

Metastatic Bc: Immunotherapy

The long-term benefit of first-line immunotherapy compared to carboplatin-based chemotherapy was recently reported in patients with metastatic or locally advanced urothelial carcinoma (5). In this new era of immuno-oncology, the treatment paradigm is shifting toward restoring tumor elimination by the immune system. This new treatment paradigm has introduced novel patterns of response and progression, such as pseudo-progression and hyperprogression, which have been observed in a wide range of cancers, particularly using 18F-FDG PET/CT (51).

Pseudo-progression is a well-described novel immune-related pattern of response. Pseudo-progression is defined as a transient increase followed by a decrease in apparent total tumor burden. Its incidence is highly variable between studies, ranging between 2 and 10%, depending on tumor type, treatment, and patients (52, 53). Pseudo-progression incidence rates have been reported as ranging between 1.5 and 17% of patients with advanced urothelial carcinoma on immunotherapy (54). Thus, the majority of apparent early progressive disease visualized on 18F-FDG PET/CT represents true progression rather than pseudo-progression (55).

Hyperprogression is defined as a rapid increase in tumor growth rate ($\geq$2-fold) compared to the expected growth rate in cancer patients treated with Anti-PD-1/PD-L1 agent. Hyperprogression occurs with an incidence of 9% in solid tumors and is associated with a poor outcome (56). The definition of hyperprogression on 18F-FDG PET/CT is currently controversial since metrics differ between institutions (57).

Immune-related adverse events (IrAEs) are well-known side effects of ICIs that can involve almost all organs. IrAEs in patients with MIBC treated with ICIs occur in up to 23% of patients, similar to patients with other solid malignancies (6). 2-Deoxy-2-(18F)fluoro-D-glucose positron emission tomography/computed tomography performed during the treatment for restaging and/or response assessment can also reveal a wide range of IrAEs, (e.g., sarcoidosis-like syndrome, thyroiditis, hypophysitis, enterocolitis, interstitial pneumonitis, pancreatitis, and arthritis) with an accuracy of 83% (58, 59). Authors recently reported that shared genetic factors impact risk for IrAEs and survival on immunotherapy in BC (60).

To date, there have been no studies investigating the specific role of 18F-FDG PET/CT in evaluating the response of advanced BC to immunotherapy.

RECURRENT DISEASE

Few published studies have evaluated the diagnostic performance of 18F-FDG PET/CT in detecting relapse of BC after systemic chemotherapy and/or RC (**Figure 1C**). Sensitivity is reported to be between 87 and 92% and specificity between 83 and 94%, with a significant change in the management of up to 40% of patients compared to conventional imaging alone (61–63). Alongi et al. reported SUVmax > 6 and TLG > 8.5 of recurrent bladder tumors as the most significant predictors of 2-year progression-free survival (61).

TECHNICAL CONSIDERATIONS, NEW TRACERS, AND DEVICES

Although several studies have demonstrated the clinical utility of 18F-FDG PET/CT throughout the management of patients with BC, technical caveats should be considered.

Pet Technology

Over the past decade the gradual spread of 3D PET equipped with time-of-flight technology and using iterative reconstruction algorithms including point-spread-function correction has significatively improved image quality. More specifically, these recently developed technologies have enhanced detectability of subcentimeter foci of disease (64) and therefore may significantly influence performances of PET/CT in regard to nodal disease (65).

Rationale for Diuretic Protocols

Although detection and characterization of primary bladder tumors is not the aim of 18F-FDG PET/CT, adapted protocols are of interest in terms of improving LN staging in pelvic malignancies by lowering artifacts caused by concentrated urinary radioactivity (66). Due to its diagnostic performance with respect to nodal disease, 18F-FDG PET/CT protocols for BC should integrate forced diuresis (with 20 mg to 40 mg of furosemide intravenously) and oral hyperhydration (1.5–2 L) in everyday clinical practice (10, 16). Supplementary delayed pelvic images may be helpful in selected patients with inconclusive standard images.

Early Dynamic Acquisitions

In an effort to improve tumor conspicuity through an increase of tumor-to-urine SUVmax ratio, few authors have investigated the utility of early dynamic acquisitions, before radioactive urine has had a chance to fill the bladder (13, 67, 68). In these proof-of-concept studies, the authors suggested that such dynamic acquisitions might improve tumor detection and staging; however, the impact on LN staging was not evaluated.

Other Pet Tracers

In an effort to overcome the limitations of excreted urinary 18F-FDG in the setting of BC and pelvic LN evaluation, additional PET tracers have been studied. These studies have not demonstrated sufficient improvement to justify implementation of these tracers in everyday practice.

11C-Acetate PET/CT has been investigated in small cohorts. It has not shown significantly different results compared to MRI and CECT for tumor detection and LN staging (69–71).

In BC patients before cystectomy, 11C-Choline PET/CT detected pelvic LN involvement with an accuracy of 81% (sensitivity of 90% and specificity of 71%) (72). In an intra-patient comparison, 11C-choline PET/CT appeared to have no significant advantage compared to 18F-FDG PET/CT (73). In two studies comparing 11C-Choline PET/CT to conventional imaging for LN staging, sensitivity was of 42–58% and specificity was 66–84% for 11C-Choline PET/CT, compared to 14–75% and 56–90%, respectively, with CECT alone (74, 75). At initial staging, 11C-Choline PET/CT was not able to significantly predict overall survival or cancer-specific death (76). Graziani et al. studied the performance of 11C-Choline PET/CT in detection BC relapse. The authors reported a sensitivity of 67% and a specificity of 85% for local relapse and a sensitivity of 90% and specificity of 92% for distant relapse (77). These results are in line with the performance of 18F-FDG PET/CT.

Regarding the investigation of bone metastases, 18F-sodium fluoride PET/CT has been shown to reveal more bone metastases than standard bone scintigraphy (78) or 18F-FDG PET/CT (79).

Rationale for Pet/Mri

Thanks to its superior contrast resolution, MRI is currently the most effective non-invasive imaging modality for local staging of BC, with an accuracy between 92 and 98% for detection of muscle invasion (18, 80). In spite of this contrast resolution, MRI's anatomical sequences (T2 weighted, T1 weighted) have not been shown to be superior to CECT in detecting LN involvement. Combining functional sequences such as DWI and dynamic contrast enhancement (DCE) still results in a relatively low sensitivity (80). Thus, MRI and PET complement one another for the staging of BC; the former is useful for local staging, the latter for distant metastasis detection, while together they may be synergistic for revealing pelvic LN involvement.

To date, only small studies have been published investigating the role of PET/MRI in MIBC. In a prospective pilot study including 24 patients, Rosenkrantz et al. highlighted the potential interest of hybrid 18F-FDG PET/MRI over MRI alone including DWI, especially to detect pelvic LN involvement with an accuracy of 95% for PET/MRI versus 76% for MRI alone, as well as non-nodal extravesical pelvic involvement with an accuracy of 100 versus 91%, respectively (81). In contrast, another recently published pilot study investigating PET/MRI in MIBC patients after NAC (N = 18, with LN involvement in only three patients) showed a sensitivity of 80% and a specificity of 56% for detection of the primary tumor, and 0% and 100%, respectively, for detection of LN involvement (82).

Rationale for Artificial Intelligence

Recent advances in artificial intelligence (AI) techniques are able to harness routine PET images into data that can be mined to extract imaging biomarkers for purposes of guiding precision medicine for BC. The term AI encompasses distinct fields such as deep learning, and radiomics, which go beyond the scope of this review. AI can be leveraged on clinical data, molecular and genetic biomarkers, and imaging for several narrow tasks in BC. While there are currently no studies published on the use of AI on 18F-FDG PET images for BC, promising results of its application using CT and MRI images have been reported in terms of predicting the depth of invasion of the primary tumor (83), grade (84), local and systemic staging (85), and assessment of treatment response (86). Additionally, AI based on PET/CT images has been used in other malignancies to predict nodal disease (87), risk stratification (88), treatment response (89), and patient outcomes (90). The main advantages of using AI is its potential reproducibility, as compared to the inherently subjective interpretation of medical images by physicians as well as the ability to harness large quantities of data that may escape the pattern recognition abilities of humans.

CONCLUSION AND FUTURE DIRECTIONS

There is strong and still evolving evidence supporting the utility of 18F-FDG PET/CT in the management of MIBC. 18F-FDG PET/CT appears to outperform and/or complement conventional imaging techniques for several tasks. Current data support the idea that 18F-FDG PET/CT may help to select the most efficacious treatment for each patient at each step of MIBC management. There is some agreement among the medical community that 18F-FDG PET/CT is relevant to guide management at initial staging in patients considered as oligometastatic by conventional imaging, such as patients with enlarged pelvic LN. However, to date, prospective studies with high level of evidence are lacking in order to allow the systematic adoption of 18F-FDG PET/CT in structured guidelines. With the advent of AI, the broad range of clinical, biological, anatomical, and metabolic data may be harnessed in order to lead to improved precision medicine.

AUTHOR CONTRIBUTIONS

AG, LD, and MR designed the review article. AG, HVR, MR, HS, and LD performed the literature review and wrote and edited the manuscript. J-FG, OD, P-YS, and HS critically revised the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2018) 68:394–424. doi: 10.3322/caac.21492
2. Kamat AM, Hahn NM, Efstathiou JA, Lerner SP, Malmström P-U, Choi W, et al. Bladder cancer. *Lancet.* (2016) 388:2796–810. doi: 10.1016/S0140-6736(16)30512-8
3. Witjes JA, Bruins HM, Cathomas R, Compérat EM, Cowan NC, Gakis G, et al. European association of urology guidelines on muscle-invasive and metastatic bladder cancer: summary of the 2020 guidelines. *Eur Urol.* (2020) 29:S0302-2838(20)30230-X. doi: 10.1016/j.eururo.2020.03.055
4. Rouanne M, Roumiguié M, Houédé N, Masson-Lecomte A, Colin P, Pignot G, et al. Development of immunotherapy in bladder cancer: present and future on targeting PD(L)1 and CTLA-4 pathways. *World J Urol.* (2018) 36:1727–40. doi: 10.1007/s00345-018-2332-5
5. Feld E, Harton J, Meropol NJ, Adamson BJS, Cohen A, Parikh RB, et al. Effectiveness of first-line immune checkpoint blockade versus carboplatin-based chemotherapy for metastatic urothelial cancer. *Eur Urol.* (2019) 76:524–32. doi: 10.1016/j.eururo.2019.07.032
6. Siefker-Radtke A, Curti B. Immunotherapy in metastatic urothelial carcinoma: focus on immune checkpoint inhibition. *Nat Rev Urol.* (2018) 15:112–24. doi: 10.1038/nrurol.2017.190
7. Whyard T, Waltzer WC, Waltzer D, Romanov V. Metabolic alterations in bladder cancer: applications for cancer imaging. *Exp Cell Res.* (2016) 341:77–83. doi: 10.1016/j.yexcr.2016.01.005
8. Wang N, Jiang P, Lu Y. Is fluorine-18 fluorodeoxyglucose positron emission tomography useful for detecting bladder lesions? a meta-analysis of the literature. *Urol Int.* (2014) 92:143–9. doi: 10.1159/000351964
9. Fang N, Wang Y, Zeng L, Wu Z, Cui X, Wang Q, et al. Feasible Method to enable clear visualization of suspected bladder cancer with 18F-FDG PET/CT. *Clinical Imaging.* (2014) 38:704–9. doi: 10.1016/j.clinimag.2014.04.018
10. Harkirat S, Anand S, Jacob M. Forced diuresis and dual-phase ^{18} F-fluorodeoxyglucose-PET/CT scan for restaging of urinary bladder cancers. *Indian J Radiol Imaging.* (2010) 20:13. doi: 10.4103/0971-3026.59746
11. Higashiyama A, Komori T, Juri H, Inada Y, Azuma H, Narumi Y. Detectability of residual invasive bladder cancer in delayed 18F-FDG PET imaging with oral hydration using 500 mL of water and voiding-refilling. *Ann Nucl Med.* (2018) 32:561–7. doi: 10.1007/s12149-018-1280-x
12. Nayak B, Dogra PN, Naswa N, Kumar R. Diuretic 18F-FDG PET/CT imaging for detection and locoregional staging of urinary bladder cancer: prospective evaluation of a novel technique. *Eur J Nucl Med Mol Imaging.* (2013) 40:386–93. doi: 10.1007/s00259-012-2294-6
13. Sharma A, Mete UK, Sood A, Kakkar N, Gorla AKR, Mittal BR. Utility of early dynamic and delayed post-diuretic ^{18} F-FDG PET/CT SUV $_{max}$ in predicting tumour grade and T-stage of urinary bladder carcinoma: results from a prospective single centre study. *BJR.* (2017) 90:20160787. doi: 10.1259/bjr.20160787
14. Yan H, Zhou X, Wang X, Li R, Shi Y, Xia Q, et al. Delayed ^{18} F FDG PET/CT imaging in the assessment of residual tumors after transurethral resection of bladder cancer. *Radiology.* (2019) 293:144–50. doi: 10.1148/radiol.2019190032
15. Yang Z, Cheng J, Pan L, Hu S, Xu J, Zhang Y, et al. Is whole-body fluorine-18 fluorodeoxyglucose PET/CT plus additional pelvic images (oral hydration–voiding–refilling) useful for detecting recurrent bladder cancer? *Ann Nucl Med.* (2012) 26:571–7. doi: 10.1007/s12149-012-0614-3
16. Yildirim-Poyraz N, Ozdemir E, Uzun B, Turkolmez S. Dual phase 18F-fluorodeoxyglucose positron emission tomography/computed tomography with forced diuresis in diagnostic imaging evaluation of bladder cancer. *Rev Españo Med Nucl Imagen Mol (Engl Ed).* (2013) 32:214–21. doi: 10.1016/j.remnie.2013.07.005
17. Mertens LS, Bruin NM, Vegt E, de Blok WM, Fioole-Bruining A, van Rhijn BW, et al. Catheter-assisted 18F-FDG-PET/CT imaging of primary bladder cancer: a prospective study. *Nucl Med Commun.* (2012) 33:1195–201. doi: 10.1097/MNM.0b013e3283567473
18. van der Pol CB, Sahni VA, Eberhardt SC, Oto A, Akin O, Alexander LF, et al. ACR appropriateness criteria ® pretreatment Staging of muscle-invasive bladder cancer. *J Am Coll Radiol.* (2018) 15:S150–9. doi: 10.1016/j.jacr.2018.03.020
19. Ha HK, Koo PJ, Kim S-J. Diagnostic accuracy of F-18 FDG PET/CT for preoperative lymph node staging in newly diagnosed bladder cancer patients: a systematic review and meta-analysis. *Oncology.* (2018) 95:31–8. doi: 10.1159/000488200
20. Soubra A, Hayward D, Dahm P, Goldfarb R, Froehlich J, Jha G, et al. The diagnostic accuracy of 18F-fluorodeoxyglucose positron emission tomography and computed tomography in staging bladder cancer: a single-institution study and a systematic review with meta-analysis. *World J Urol.* (2016) 34:1229–37. doi: 10.1007/s00345-016-1772-z
21. Dason S, Wong NC, Donahue TF, Meier A, Zheng J, Mannelli L, et al. Utility of routine preoperative 18F-fluorodeoxyglucose positron emission tomography-computed tomography (18F-FDG PET/CT) in identifying pathologic lymph node metastases at radical cystectomy. *J Urol.* (2020) 204:101097. doi: 10.1097/JU.0000000000001006
22. Girard A, Rouanne M, Taconet S, Radulescu C, Neuzillet Y, Girma A, et al. Integrated analysis of 18F-FDG PET/CT improves preoperative lymph node staging for patients with invasive bladder cancer. *Eur Radiol.* (2019) 29:4286–93. doi: 10.1007/s00330-018-5959-0
23. Vind-Kezunovic S, Bouchelouche K, Ipsen P, Høyer S, Bell C, Bjerggaard Jensen J. Detection of lymph node metastasis in patients with bladder cancer using maximum standardised uptake value and 18F-fluorodeoxyglucose positron emission tomography/computed tomography: results from a high-volume centre including long-term follow-up. *Eur Urol Focus.* (2019) 5:90–6. doi: 10.1016/j.euf.2017.06.005
24. Jensen TK, Holt P, Gerke O, Riehmann M, Svolgaard B, Marcussen N, et al. Preoperative lymph-node staging of invasive urothelial bladder cancer with ^{18} F-fluorodeoxyglucose positron emission tomography/computed axial tomography and magnetic resonance imaging: correlation with histopathology. *Scand J Urol Nephrol.* (2011) 45:122–8. doi: 10.3109/00365599.2010.544672
25. Apolo AB, Riches J, Schöder H, Akin O, Trout A, Milowsky MI, et al. Clinical value of fluorine-18 2-fluoro-2-deoxy-d-glucose positron emission tomography/computed tomography in bladder cancer. *JCO.* (2010) 28:3973–8. doi: 10.1200/JCO.2010.28.7052
26. Goodfellow H, Viney Z, Hughes P, Rankin S, Rottenberg G, Hughes S, et al. Role of fluorodeoxyglucose positron emission tomography (FDG PET)-computed tomography (CT) in the staging of bladder cancer: FDG pet in the staging of bladder cancer. *BJU Int.* (2014) 114:389–95. doi: 10.1111/bju.12608
27. Öztürk H. Detecting metastatic bladder cancer using ^{18}F-fluorodeoxyglucose positron-emission tomography/computed tomography. *Cancer Res Treat.* (2015) 47:834–43. doi: 10.4143/crt.2014.157
28. Mertens LS, Mir MC, Scott AM, Lee ST, Fioole-Bruining A, Vegt E, et al. 18F-fluorodeoxyglucose–positron emission tomography/computed tomography aids staging and predicts mortality in patients with muscle-invasive bladder cancer. *Urology.* (2014) 83:393–9. doi: 10.1016/j.urology.2013.10.032
29. Mertens LS, Fioole-Bruining A, Vegt E, Vogel WV, van Rhijn BW, Horenblas S. Impact of ^{18} F-fluorodeoxyglucose (FDG)-positron-emission tomography/computed tomography (PET/CT) on management of patients with carcinoma invading bladder muscle: FDG-PET/CT in carcinoma invading bladder muscle. *BJU Int.* (2013) 112:729–34. doi: 10.1111/bju.12109
30. Rouanne M, Alhammadi A, Vilain D, Radulescu C, Lebret T. Value of positron emission tomography in diagnosing synchronous penile metastasis from urothelial bladder cancer. *World J Surg Onc.* (2015) 13:276. doi: 10.1186/s12957-015-0696-1
31. Guney IB, Küçüker KA, Izol V, Kibar M. The role and effect of FDG-PET/CT on patient management and restaging of bladder carcinoma. *Turkish J Urol.* (2019) 45:423–30. doi: 10.5152/tud.2019.84453
32. Yang Z, Pan L, Cheng J, Hu S, Xu J, Ye D, et al. Clinical value of whole body fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography in the detection of metastatic bladder cancer: FDG-PET metastatic bladder cancer. *Int J Urol.* (2012) 19:639–44. doi: 10.1111/j.1442-2042.2012.02989.x
33. Kibel AS, Dehdashti F, Katz MD, Klim AP, Grubb RL, Humphrey PA, et al. Prospective Study of [^{18} F]fluorodeoxyglucose positron emission tomography/computed tomography for staging of muscle-invasive bladder carcinoma. *JCO.* (2009) 27:4314–20. doi: 10.1200/JCO.2008.20.6722

34. Huo J, Chu Y, Chamie K, Smaldone MC, Boorjian SA, Baillargeon JG, et al. Increased utilization of positron emission tomography/computed tomography (PET/CT) imaging and its economic impact for patients diagnosed with bladder cancer. *Clin Genitour Cancer.* (2018) 16:e99–111. doi: 10.1016/j.clgc.2017.07.018
35. Flaig TW, Spiess PE, Agarwal N, Bangs R, Boorjian SA, Buyyounouski MK, et al. NCCN guidelines insights: bladder cancer, version 5.2018. *J Natl Compr Canc Netw.* (2018) 16:1041–53. doi: 10.6004/jnccn.2018.0072
36. Salaün P-Y, Abgral R, Malard O, Querellou-Lefranc S, Quere G, Wartski M, et al. Good clinical practice recommendations for the use of PET/CT in oncology. *Eur J Nucl Med Mol Imaging.* (2020) 47:28–50. doi: 10.1007/s00259-019-04553-8
37. Witjes JA, Babjuk M, Bellmunt J, Bruins HM, De Reijke TM, De Santis M, et al. EAU-ESMO consensus statements on the management of advanced and variant bladder cancer—an international collaborative multistakeholder effort†. *Eur Urol.* (2020) 77:223–50. doi: 10.1016/j.eururo.2019.09.035
38. Pontoizeau C, Girard A, Mesbah H, Haumont L-A, Devillers A, Tempescul A, et al. Prognostic value of baseline total metabolic tumor volume measured on FDG PET in patients with richter syndrome. *Clin Nucl Med.* (2020) 45:118–22. doi: 10.1097/RLU.0000000000002879
39. Seban R-D, Mezquita L, Berenbaum A, Dercle L, Botticella A, Le Pechoux C, et al. Baseline metabolic tumor burden on FDG PET/CT scans predicts outcome in advanced NSCLC patients treated with immune checkpoint inhibitors. *Eur J Nucl Med Mol Imaging.* (2020) 47:1147–57. doi: 10.1007/s00259-019-04615-x
40. Seban R-D, Moya-Plana A, Antonios L, Yeh R, Marabelle A, Deutsch E, et al. Prognostic 18F-FDG PET biomarkers in metastatic mucosal and cutaneous melanoma treated with immune checkpoint inhibitors targeting PD-1 and CTLA-4. *Eur J Nucl Med Mol Imaging.* (2020) 47:2301–12. doi: 10.1007/s00259-020-04757-3
41. Dercle L, Seban R-D, Lazarovici J, Schwartz LH, Houot R, Ammari S, et al. ^{18} F-FDG PET and CT scans detect new imaging patterns of response and progression in patients with hodgkin lymphoma treated by anti–programmed death 1 immune checkpoint inhibitor. *J Nucl Med.* (2018) 59:15–24. doi: 10.2967/jnumed.117.193011
42. Seban R-D, Nemer JS, Marabelle A, Yeh R, Deutsch E, Ammari S, et al. Prognostic and theranostic 18F-FDG PET biomarkers for anti-PD1 immunotherapy in metastatic melanoma: association with outcome and transcriptomics. *Eur J Nucl Med Mol Imaging.* (2019) 46:2298–310. doi: 10.1007/s00259-019-04411-7
43. Chen R, Zhou X, Liu J, Huang G. Relationship between the expression of PD-1/PD-L1 and 18F-FDG uptake in bladder cancer. *Eur J Nucl Med Mol Imaging.* (2019) 46:848–54. doi: 10.1007/s00259-018-4208-8
44. Girard A, Rouanne M. Comment on: relationship between the expression of PD-1/PD-L1 and 18F-FDG uptake in bladder cancer. *Eur J Nucl Med Mol Imaging.* (2019) 46:1212–3. doi: 10.1007/s00259-019-04296-6
45. Soubra A, Gencturk M, Froelich J, Balaji P, Gupta S, Jha G, et al. FDG-PET/CT for assessing the response to neoadjuvant chemotherapy in bladder cancer patients. *Clin Genitour Cancer.* (2018) 16:360–4. doi: 10.1016/j.clgc.2018.05.008
46. Mertens LS, Fioole-Bruining A, van Rhijn BWG, Kerst JM, Bergman AM, Vogel WV, et al. FDG-positron emission tomography/computerized tomography for monitoring the response of pelvic lymph node metastasis to neoadjuvant chemotherapy for bladder cancer. *J Urol.* (2013) 189:1687–91. doi: 10.1016/j.juro.2012.11.009
47. van de Putte EEF, Vegt E, Mertens LS, Bruining A, Hendricksen K, van der Heijden MS, et al. FDG-PET/CT for response evaluation of invasive bladder cancer following neoadjuvant chemotherapy. *Int Urol Nephrol.* (2017) 49:1585–91. doi: 10.1007/s11255-017-1637-4
48. Kollberg P, Almquist H, Bläckberg M, Cwikiel M, Gudjonsson S, Lyttkens K, et al. [^{18} F]Fluorodeoxyglucose-positron emission tomography/computed tomography response evaluation can predict histological response at surgery after induction chemotherapy for oligometastatic bladder cancer. *Scand J Urol.* (2017) 51:308–13. doi: 10.1080/21681805.2017.1321579
49. Öztürk H. Comparing RECIST with EORTC criteria in metastatic bladder cancer. *J Cancer Res Clin Oncol.* (2016) 142:187–94. doi: 10.1007/s00432-015-2022-2
50. Giannatempo P, Alessi A, Miceli R, Raggi D, Farè E, Nicolai N, et al. Interim fluorine-18 fluorodeoxyglucose positron emission tomography for early metabolic assessment of therapeutic response to chemotherapy for metastatic transitional cell carcinoma. *Clin Genitourinary Cancer.* (2014) 12:433–9. doi: 10.1016/j.clgc.2014.03.007
51. Aide N, Hicks RJ, Le Tourneau C, Lheureux S, Fanti S, Lopci E. FDG PET/CT for assessing tumour response to immunotherapy : report on the EANM symposium on immune modulation and recent review of the literature. *Eur J Nucl Med Mol Imaging.* (2019) 46:238–50. doi: 10.1007/s00259-018-4171-4
52. Chiou VL, Burotto M. Pseudoprogression and Immune-related response in solid tumors. *JCO.* (2015) 33:3541–3. doi: 10.1200/JCO.2015.61.6870
53. Dromain C, Beigelman C, Pozzessere C, Duran R, Digklia A. Imaging of tumour response to immunotherapy. *Eur Radiol Exp.* (2020) 4:2. doi: 10.1186/s41747-019-0134-1
54. Soria F, Beleni AI, D'Andrea D, Resch I, Gust KM, Gontero P, et al. Pseudoprogression and hyperprogression during immune checkpoint inhibitor therapy for urothelial and kidney cancer. *World J Urol.* (2018) 36:1703–9. doi: 10.1007/s00345-018-2264-0
55. Humbert O, Cadour N, Paquet M, Schiappa R, Poudenx M, Chardin D, et al. 18FDG PET/CT in the early assessment of non-small cell lung cancer response to immunotherapy: frequency and clinical significance of atypical evolutive patterns. *Eur J Nucl Med Mol Imaging.* (2020) 47:1158–67. doi: 10.1007/s00259-019-04573-4
56. Champiat S, Dercle L, Ammari S, Massard C, Hollebecque A, Postel-Vinay S, et al. Hyperprogressive disease is a new pattern of progression in cancer patients treated by anti-PD-1/PD-L1. *Clin Cancer Res.* (2017) 23:1920–8. doi: 10.1158/1078-0432.CCR-16-1741
57. Seban R-D, Schwartz LH, Bonardel G, Dercle L. Diagnosis of hyper-progressive disease in patients treated with checkpoint inhibitors using 18F-FDG PET/CT. *J Nucl Med.* (2020) 61:1404–5. doi: 10.2967/jnumed.120.242768
58. Dercle L, Mokrane F-Z, Schiano de Colella JM, Stamatoullas A, Morschhauser F, Brice P, et al. Unconventional immune-related phenomena observed using 18F-FDG PET/CT in Hodgkin lymphoma treated with anti PD-1 monoclonal antibodies. *Eur J Nucl Med Mol Imaging.* (2019) 46:1391–2. doi: 10.1007/s00259-019-04310-x
59. Mekki A, Dercle L, Lichtenstein P, Marabelle A, Michot J-M, Lambotte O, et al. Detection of immune-related adverse events by medical imaging in patients treated with anti-programmed cell death 1. *Eur J Cancer.* (2018) 96:91–104. doi: 10.1016/j.ejca.2018.03.006
60. Khan Z, Di Nucci F, Kwan A, Hammer C, Mariathasan S, Rouilly V, et al. Polygenic risk for skin autoimmunity impacts immune checkpoint blockade in bladder cancer. *Proc Natl Acad Sci USA.* (2020) 117:12288–94. doi: 10.1073/pnas.1922867117
61. Alongi P, Caobelli F, Gentile R, Stefano A, Russo G, Albano D, et al. Recurrent bladder carcinoma: clinical and prognostic role of 18 F-FDG PET/CT. *Eur J Nucl Med Mol Imaging.* (2017) 44:224–33. doi: 10.1007/s00259-016-3500-8
62. Öztürk H, Karapolat I. Efficacy of 18F-fluorodeoxyglucose-positron emission tomography/computed tomography in restaging muscle-invasive bladder cancer following radical cystectomy. *Exp Therap Med.* (2015) 9:717–24. doi: 10.3892/etm.2015.2187
63. Zattoni F, Incerti E, Dal Moro F, Moschini M, Castellucci P, Panareo S, et al. 18F-FDG PET/CT and urothelial carcinoma: impact on management and prognosis—a multicenter retrospective study. *Cancers.* (2019) 11:700. doi: 10.3390/cancers11050700
64. Hashimoto N, Morita K, Tsutsui Y, Himuro K, Baba S, Sasaki M. Time-of-flight information improved the detectability of subcentimeter spheres using a clinical PET/CT scanner. *J Nucl Med Technol.* (2018) 46:268–73. doi: 10.2967/jnmt.117.204735
65. Akamatsu G, Mitsumoto K, Taniguchi T, Tsutsui Y, Baba S, Sasaki M. Influences of point-spread function and time-of-flight reconstructions on standardized uptake value of lymph node metastases in FDG-PET. *Eur J Radiol.* (2014) 83:226–30. doi: 10.1016/j.ejrad.2013.09.030
66. Dierickx LO, Dercle L, Chaltiel L, Caselles O, Brillouet S, Zerdoud S, et al. Evaluation of 2 diuretic 18fluorine-fluorodeoxyglucose positron emission tomography/computed tomography imaging protocols for intrapelvic cancer. *Q J Nucl Med Mol Imaging.* (2019) 63:284–91. doi: 10.23736/S1824-4785.17.02912-0

67. Belakhlef S, Church C, Jani C, Lakhanpal S. Early dynamic PET/CT and 18F-FDG blood flow imaging in bladder cancer detection: a novel approach. *Clin Nucl Med.* (2012) 37:366–8. doi: 10.1097/RLU.0b013e3182443110
68. Yoon H-J, Yoo J, Kim Y, Lee DH, Kim BS. Enhanced application of 18F-FDG PET/CT in bladder cancer by adding early dynamic acquisition to a standard delayed PET protocol. *Clin Nucl Med.* (2017) 42:749–55. doi: 10.1097/RLU.0000000000001780
69. Salminen A, Jambor I, Merisaari H, Ettala O, Virtanen J, Koskinen I, et al. 11C-acetate PET/MRI in bladder cancer staging and treatment response evaluation to neoadjuvant chemotherapy: a prospective multicenter study (ACEBIB trial). *Cancer Imaging.* (2018) 18:25. doi: 10.1186/s40644-018-0158-4
70. Schöder H, Ong SC, Reuter VE, Cai S, Burnazi E, Dalbagni G, et al. Initial Results with 11C-Acetate positron emission tomography/computed tomography (PET/CT) in the staging of urinary bladder cancer. *Mol Imaging Biol.* (2012) 14:245–51. doi: 10.1007/s11307-011-0488-0
71. Vargas HA, Akin O, Schöder H, Olgac S, Dalbagni G, Hricak H, et al. Prospective evaluation of MRI, 11C-acetate PET/CT and contrast-enhanced CT for staging of bladder cancer. *Eur J Radiol.* (2012) 81:4131–7. doi: 10.1016/j.ejrad.2012.06.010
72. Ceci F, Bianchi L, Graziani T, Castellucci P, Pultrone C, Eugenio B, et al. 11C-Choline PET/CT and bladder cancer: lymph node metastasis assessment with pathological specimens as reference standard. *Clin Nucl Med.* (2015) 40:e124–8. doi: 10.1097/RLU.0000000000000604
73. Golan S, Sopov V, Baniel J, Groshar D. Comparison of ^{11} C-choline With ^{18} F-FDG in positron emission tomography/computerized tomography for staging urothelial carcinoma: a prospective study. *J Urol.* (2011) 186:436–41. doi: 10.1016/j.juro.2011.03.121
74. Brunocilla E, Ceci F, Schiavina R, Castellucci P, Maffione AM, Cevenini M, et al. Diagnostic accuracy of 11C-choline PET/CT in preoperative lymph node staging of bladder cancer: a systematic comparison with contrast-enhanced CT and histologic findings. *Clin Nucl Med.* (2014) 39:e308–12. doi: 10.1097/RLU.0000000000000342
75. Maurer T, Souvatzoglou M, Kübler H, Opercan K, Schmidt S, Herrmann K, et al. Diagnostic efficacy of [11C]choline positron emission tomography/computed tomography compared with conventional computed tomography in lymph node staging of patients with bladder cancer prior to radical cystectomy. *Eur Urol.* (2012) 61:1031–8. doi: 10.1016/j.eururo.2011.12.009
76. Maurer T, Horn T, Souvatzoglou M, Eiber M, Beer AJ, Heck MM, et al. Prognostic value of ^{11}C-choline PET/CT and CT for predicting survival of bladder cancer patients treated with radical cystectomy. *Urol Int.* (2014) 93:207–13. doi: 10.1159/000357686
77. Graziani T, Ceci F, Lopes FL, Chichero J, Castellucci P, Schiavina R, et al. 11C-Choline PET/CT for restaging of bladder cancer. *Clin Nucl Med.* (2015) 40:e1–5. doi: 10.1097/RLU.0000000000000573
78. Araz M, Aras G, Küçük ON. The role of 18F–NaF PET/CT in metastatic bone disease. *J Bone Oncol.* (2015) 4:92–7.
79. Mena E, Lin J, Lindenberg ML, Turkbey I, Adler S, McKinney Y, et al. Value of combined 18F-FDG/18F-NaF PET/CT in tumor detection and therapy response in patients with advanced bladder cancer treated with Cabozantinib plus Nivolumab alone or in combination with Ipilimumab. *J Nucl Med.* (2017) 58:754.
80. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. The diagnostic performance of MRI for detection of lymph node metastasis in bladder and prostate cancer: an updated systematic review and diagnostic meta-analysis. *Am J Roentgenol.* (2018) 210:W95–109. doi: 10.2214/AJR.17.18481
81. Rosenkrantz AB, Friedman KP, Ponzo F, Raad RA, Jackson K, Huang WC, et al. Prospective pilot study to evaluate the incremental value of PET information in patients with bladder cancer undergoing 18F-FDG simultaneous PET/MRI. *Clin Nucl Med.* (2017) 42:e8–15. doi: 10.1097/RLU.0000000000001432
82. Eulitt PJ, Altun E, Sheikh A, Wong TZ, Woods ME, Rose TL, et al. Pilot study of [18F] fluorodeoxyglucose positron emission tomography (FDG-PET)/magnetic resonance imaging (MRI) for staging of muscle-invasive bladder cancer (MIBC). *Clin Genitour Cancer.* (2020). doi: 10.1016/j.clgc.2020.02.008
83. Wang H, Xu X, Zhang X, Liu Y, Ouyang L, Du P, et al. Elaboration of a multisequence MRI-based radiomics signature for the preoperative prediction of the muscle-invasive status of bladder cancer: a double-center study. *Eur Radiol.* (2020) 30:4816–27. doi: 10.1007/s00330-020-06796-8
84. Zhang G, Xu L, Zhao L, Mao Li, Li X, Jin Z, et al. CT-based radiomics to predict the pathological grade of bladder cancer. *Eur Radiol.* (2020). doi: 10.1007/s00330-020-06893-8. [Epub ahead of print].
85. Wu S, Zheng J, Li Y, Wu Z, Shi S, Huang M, et al. Development and validation of an MRI-based radiomics signature for the preoperative prediction of lymph node metastasis in bladder cancer. *EBioMedicine.* (2018) 34:76–84. doi: 10.1016/j.ebiom.2018.07.029
86. Cha KH, Hadjiiski LM, Cohan RH, Chan H-P, Caoili EM, Davenport MS, et al. Diagnostic accuracy of CT for prediction of bladder cancer treatment response with and without computerized decision support. *Acad Radiol.* (2019) 26:1137–45. doi: 10.1016/j.acra.2018.10.010
87. Shaish H, Mutasa S, Makkar J, Chang P, Schwartz L, Ahmed F. Prediction of Lymph node maximum standardized uptake value in patients with cancer using a 3D convolutional neural network: a proof-of-concept study. *Am J Roentgenol.* (2019) 212:238–44. doi: 10.2214/AJR.18.20094
88. Cysouw MCF, Jansen BHE, van de Brug T, Oprea-Lager DE, Pfaehler E, de Vries BM, et al. Machine learning-based analysis of [18 F]DCFPyL PET radiomics for risk stratification in primary prostate cancer. *Eur J Nucl Med Mol Imaging.* (2020). doi: 10.1007/s00259-020-04971-z. [Epub ahead of print].
89. Valentinuzzi D, Vrankar M, Boc N, Ahac V, Zupancic Z, Unk M, et al. [18F]FDG PET immunotherapy radiomics signature (iRADIOMICS) predicts response of non-small-cell lung cancer patients treated with pembrolizumab. *Radiol Oncol.* (2020) 54:285–94. doi: 10.2478/raon-2020-0042
90. Nie P, Yang G, Wang N, Yan L, Miao W, Duan Y, et al. Additional value of metabolic parameters to PET/CT-based radiomics nomogram in predicting lymphovascular invasion and outcome in lung adenocarcinoma. *Eur J Nucl Med Mol Imaging.* (2020). doi: 10.1007/s00259-020-04747-5. [Epub ahead of print].
91. Pichler R, De Zordo T, Fritz J, Kroiss A, Aigner F, Heidegger I, et al. Pelvic lymph node staging by combined 18 F-FDG-PET/CT imaging in bladder cancer prior to radical cystectomy. *Clin Genitour Cancer.* (2017) 15:e387–95. doi: 10.1016/j.clgc.2016.08.009
92. Uttam M, Pravin N, Anish B, Nandita K, Arup M. Is [F-18]-fluorodeoxyglucose FDG-PET/CT better than ct alone for the preoperative lymph node staging of muscle invasive bladder cancer? *Int braz J Urol.* (2016) 42:234–41. doi: 10.1590/S1677-5538.IBJU.2014.0579
93. Jeong IG, Hong S, You D, Hong JH, Ahn H, Kim C-S. FDG PET–CT for lymph node staging of bladder cancer: a prospective study of patients with extended pelvic lymphadenectomy. *Ann Surg Oncol.* (2015) 22:3150–6. doi: 10.1245/s10434-015-4369-7
94. Aljabery F, Lindblom G, Skoog S, Shabo I, Olsson H, Rosell J, et al. PET/CT versus conventional CT for detection of lymph node metastases in patients with locally advanced bladder cancer. *BMC Urol.* (2015) 15:87. doi: 10.1186/s12894-015-0080-z
95. Rouanne M, Girma A, Neuzillet Y, Vilain D, Radulescu C, Letang N, et al. Potential impact of 18F-FDG PET/CT on patients selection for neoadjuvant chemotherapy before radical cystectomy. *Eur J Surg Oncol (EJSO).* (2014) 40:1724–30. doi: 10.1016/j.ejso.2014.08.479
96. Hitier-Berthault M, Ansquer C, Branchereau J, Renaudin K, Bodere F, Bouchot O, et al. ^{18} F-fluorodeoxyglucose positron emission tomography-computed tomography for preoperative lymph node staging in patients undergoing radical cystectomy for bladder cancer: a prospective study: PET scan for bladder cancer. *Int J Urol.* (2013) 20:788–96. doi: 10.1111/iju.12045
97. Swinnen G, Maes A, Pottel H, Vanneste A, Billiet I, Lesage K, et al. FDG-PET/CT for the preoperative lymph node staging of invasive bladder cancer. *Eur Urol.* (2010) 57:641–7. doi: 10.1016/j.eururo.2009.05.014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 29 October 2020
doi: 10.3389/fonc.2020.580277

Dual Tracers of 16α-[18F]fluoro-17β-Estradiol and [18F] fluorodeoxyglucose for Prediction of Progression-Free Survival After Fulvestrant Therapy in Patients With HR+/HER2- Metastatic Breast Cancer

Cheng Liu[1,2,3,4,5], *Xiaoping Xu*[1,2,3,4,5], *Huiyu Yuan*[1,2,3,4,5], *Yongping Zhang*[1,2,3,4,5], *Yingjian Zhang*[1,2,3,4,5], *Shaoli Song*[1,2,3,4,5]* and *Zhongyi Yang*[1,2,3,4,5]*

[1] Department of Nuclear Medicine, Fudan University Shanghai Cancer Center, Shanghai, China, [2] Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China, [3] Center for Biomedical Imaging, Fudan University, Shanghai, China, [4] Shanghai Engineering Research Center of Molecular Imaging Probes, Shanghai, China, [5] Department of Nuclear Medicine, Shanghai Proton and Heavy Ion Center, Shanghai, China

OPEN ACCESS

Edited by:
Laurent Dercle,
Columbia University Irving Medical Center, United States

Reviewed by:
Athanasios G. Zafeirakis,
Army Share Fund Hospital (NIMTS), Greece
Lanell Peterson,
University of Washington Medical Center, United States

***Correspondence:**
Zhongyi Yang
yangzhongyi21@163.com
Shaoli Song
shaoli-song@163.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 05 July 2020
Accepted: 16 September 2020
Published: 29 October 2020

Citation:
Liu C, Xu X, Yuan H, Zhang Y, Zhang Y, Song S and Yang Z (2020) Dual Tracers of 16α-[18F]fluoro-17β-Estradiol and [18F]fluorodeoxyglucose for Prediction of Progression-Free Survival After Fulvestrant Therapy in Patients With HR+/HER2- Metastatic Breast Cancer. Front. Oncol. 10:580277. doi: 10.3389/fonc.2020.580277

Objective: The purpose of this study was to employ dual tracers 16α-[18F]fluoro-17β-estradiol (^{18}F-FES) and [18F]fluorodeoxyglucose (^{18}F-FDG) as imaging biomarkers in predicting progression-free survival (PFS) in ER-positive metastatic breast cancer (MBC) patients receiving fulvestrant therapy.

Methods: We retrospectively analyzed 35 HR+HER2- MBC patients who underwent ^{18}F-FES and ^{18}F-FDG PET/CT scans prior to fulvestrant therapy in our center. The SUVmax across all metastatic lesions on the PET/CT were assessed. The heterogeneity of ER expression was assigned by the presence of any ^{18}F-FES negative lesions for patients with entirely ^{18}F-FES positive lesions categorized into two groups by the median ratio of FES/FDG SUVmax, low FES/FDG, and high FES/FDG. PFS were estimated by the Kaplan-Meier method and compared by the log-rank test. Univariate and multivariate analyses were performed using the Cox proportional hazard model.

Results: In total, 12 patients had both ^{18}F-FES negative and positive lesions, indicating the heterogeneity of ER expression in metastatic lesions. These patients had a low median PFS of 5.5 months (95% CI 2.3–8.7). Of patients with entirely ^{18}F-FES positive lesions, 11 had a low FES/FDG, and 12 had a high FES/FDG. These groups had a median PFS of 29.4 months (95% CI 2.3–56.5) and 14.7 months (95% CI 10.9–18.5), respectively. The patients were stratified in three categories based on incorporating both ^{18}F-FES and ^{18}F-FDG imaging results that were significantly correlated with PFS by univariate analysis ($P < 0.001$) and multivariate analysis ($P = 0.006$).

Conclusion: ^{18}F-FES and ^{18}F-FDG PET could serve as prognostic imaging biomarkers for ER-positive MBC patients treated with fulvestrant therapy.

Keywords: heterogeneity, ER expression, breast cancer, FES/FDG, fulvestrant

INTRODUCTION

Breast cancer is the most common cancer in women worldwide. According to U.S. cancer statistics, about 276,480 newly diagnosed cases are estimated in 2020, resulting in approximately 42,170 deaths (1). It is the second most common cause of cancer death in women. Approximately 70%–80% of breast cancers are hormone receptor (HR)-positive, and endocrine therapy plays a vital role in the management of such cancers (2).

Fulvestrant, a pure anti-estrogen drug that exerts no partial agonist effects, is approved for postmenopausal women with HR+ metastatic breast cancer (MBC) and disease progression following the prior failure of other endocrine therapy (3, 4). Although many patients have a prolonged clinical response to fulvestrant, there are still some patients who are unable to benefit or develop resistance. Therefore, the identification of clinical or molecular markers that predict which patients with MBC might benefit from Fulvestrant is vitally important because it helps to individualize treatment and could significantly improve the management of breast cancer. The level of ER expression has been shown to provide important prognostic information, and in most cases, higher levels of tumor ER expression are associated with more noteworthy clinical benefit from conventional endocrine therapy (5). A biopsy was routinely utilized to discriminate between ER-positive and ER-negative lesions. However, this gold standard method may not be representative of ER heterogeneity. Furthermore, collecting a biopsy sample from metastatic tissue is not always feasible in daily practice because of the characteristics of lesion location and the risk associated with biopsy.

Positron emission tomography (PET) with 16α-[18F]fluoro-17β-oestradiol (^{18}F-FES) has been proposed as a noninvasive method to visualize and quantify ER expression in recurrent or metastatic lesions (6, 7). Early clinical studies focused on sole ^{18}F-FES PET imaging to predict clinical response to endocrine therapy, rarely performed in combination with ^{18}F-FDG imaging (8–10). Kurland and colleagues evaluated the ability of ^{18}F-FDG and ^{18}F-FES to predict progression-free survival (PFS) in 84 patients treated by salvage endocrine therapy for ER-positive MBC (11). They summarize that, although ^{18}F-FES PET is not predictive of the patient's PFS in the whole population, it is meaningful that this imaging could stratify the patients with high FDG uptake. However, this study fails to discuss the response to different endocrine therapies because of differences in the pharmacodynamics of different ER antagonists (12). We have previously reported that early change in SUVmax of ^{18}F-FES PET/CT could be used to predict response to fulvestrant (13). Nevertheless, this method still requires a 28-day period of fulvestrant treatment before the effect can be observed. Therefore, the purpose of our study was to evaluate the clinical value of dual tracers ^{18}F-FDG and ^{18}F-FES at baseline in predicting the response of fulvestrant in HR-positive MBC patients.

METHODS

Patients

In this retrospective analysis, we evaluated 35 HR+/HER2- MBC patients who were treated with 500 mg fulvestrant and underwent both ^{18}F-FES PET/CT and ^{18}F-FDG PET/CT scans within 4 weeks before initiating treatment between May 2016 and March 2019. The lag time between the two scans was within 1 week. All data were retrospectively collected from the medical records. To ensure the sensitivity and specificity of ^{18}F-FDG and ^{18}F-FES imaging, patients with ER antagonist discontinuation for less than 5 weeks and medical comorbidities (diabetes, a chronic infection, or chronic inflammatory conditions) were not enrolled in this study (12, 14). The enrolled patients had performed ^{18}F-FES scans for one of the following purposes: 1) predicting response to fulvestrant, a phase II study (NCT03507088, n = 23) or 2) identifying the ER status of metastatic lesions for clinical practice (n = 12). The study was approved by the Fudan University Shanghai Cancer Center Ethic Committee and Institutional Review Boards for clinical investigation, and the need for informed consent was waived as it is a retrospective study.

Fulvestrant and Clinical Follow-Up

Fulvestrant 500 mg was administered by intramuscular injection on days 1, 15, and 29 and every 28 days after that. For premenopausal women, patients received concurrent luteinising hormone-releasing hormone analogues (LHRHa). Treatment continued until progressive disease (PD) or other criteria for discontinuation were met in terms of adverse events or a patient's decision to withdraw.

Clinical follow-up was performed every 3 months by radiologic imaging (e.g., diagnostic CT, MRI, bone scan, ^{18}F-FDG), serum tumor markers, and evaluation of symptoms until disease progression or death. PFS was defined as the time from fulvestrant treatment to disease progression or death from any cause. For patients with measurable disease, tumor response was determined by an experienced radiologist according to the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 and was blinded to the results of baseline ^{18}F FES and ^{18}F FDG PET/CT. Patients with only non-measurable lesions were considered to have disease progression when there was a definite progression of existing lesions or when new lesions were detected at follow-up.

PET/CT Procedure

The synthesis and quality control of ^{18}F-FDG and ^{18}F-FES were performed as reported in our previous study (15).

^{18}F-FDG PET/CT imaging was done according to standard clinical procedures. All patients fasted for at least 6 h and had serum glucose levels less than 10 mmol/L before the intravenous injection of ^{18}F-FDG (3.7–7.4 MBq/kg). The patients were kept lying comfortably in a quiet, dimly lit room before and after the tracer injection. About 1 h after tracer injection, the patients were administered 1 L of plain water orally and then scanned in the PET/CT (Siemens Biograph 16HR PET/CT or mCT Flow PET/CT scanner). About 222 MBq of ^{18}F-FES was injected intravenously over 1–2 min. The scanning was initiated 1 h after administration of the tracer on the same PET/CT scanner as the ^{18}F-FDG. The detail of PET/CT acquisition parameters were described as reported in prior studies (16).

Image Analysis

PET images were reviewed and analyzed by two board-certified nuclear medicine physicians using a multimodality computer platform (Syngo, Siemens, Knoxville, TN, USA). All parameters were assessed in 3-dimensional volumes. Regions of interest (ROI) were manually drawn over lesions by an experienced nuclear medicine physician using the PET images with the corresponding noncontrast CT serving as a guide, and the contours of lesions were checked for concurrence by a second experienced nuclear medicine physician. In case of a discrepancy between the two physicians, consensus was reached on a final reading for the statistical analyses. Semiquantitative analysis of tumor metabolic activity was obtained using standardized uptake value (SUV) normalized to body weight. A lesion showing uptake intensity higher than with adjacent normal tissue background was defined as positive for ^{18}F-FDG and ^{18}F-FES, and hypermetabolic foci estimated by inflammatory or physiologic activity were not considered. We used the cutoff value of SUVmax ≥ 1.8 to define ^{18}F-FES positivity and quantify the ER expression based on our previous study (17). Lesions seen on ^{18}F-FES and ^{18}F-FDG PET/CT images were also identified and localized by other conventional imaging techniques (bone scan, diagnostic CT, MRI, or ultrasound). In patients with extensive metastatic lesions, an arbitrary maximum of 20 randomly chosen lesions of ^{18}F-FDG PET correspond to the ^{18}F-FES avid lesions according to the guidelines of the European Association of Nuclear Medicine (EANM) (18). Due to high physiological ^{18}F-FES uptake, patients with liver lesions were not included in the ^{18}F-FES analysis (19).

Statistical Analysis

All PET imaging parameters were dichotomized using the median as a threshold. For patients with entirely FES positive lesions, the FES/FDG ratio of each tumor was calculated, and the median value was selected as the cutoff to distinguish between high and low FES/FDG.

The survival analyses were estimated by the Kaplan–Meier method and compared by the log-rank test (image parameters and demographic factors). Univariate and multivariate analyses were estimated using the COX proportional hazards model and expressed as a hazard ratio with corresponding 95% confidence intervals and *P* values. Multivariate analysis with the stepwise model by forward selection was performed with those variables that had proven significant on univariate analysis to explore independent predictors of PFS. All data analyses were performed using IBM SPSS Statistics software, version 20.0 (IBM Corporation, Armonk, NY, USA). Two-sided *P* values of less than 0.05 were considered to indicate statistically significant differences.

RESULTS

Patient Characteristics and Treatment Outcome

The characteristics of the 35 enrolled MBC patients are listed in **Table 1**. At the time of analysis (Jan. 2020), 26 patients (74.3%) experienced progression, and all of them were radiologic PD. The median follow-up period was 9.5 months (range: 2.1–30.0), and the median PFS was 12.2 months (95% CI: 4.7–19.7). Twenty-six patients had measurable lesions according to RECIST version 1.1, four patients had non-measurable visceral lesions, and five patients had only bone metastases. Twenty-four of the 35 patients (68.6%) experienced clinical benefit from fulvestrant treatment as indicated by PFS ≥ 24 weeks. Furthermore, fulvestrant was well tolerated in all patients and no patients who discontinued treatment due to adverse events.

PET/CT Analysis

In total, 235 metastatic lesions were identified in 35 patients. The number of lesions found per patient ranged from 1 to 20 with a median of 6 lesions per patient. Lesions were located in lymph nodes (n = 78), bones (n = 117), lungs (n = 15), pleural (n = 9), soft tissue (n = 15), and the liver (n = 1). All these metastatic lesions were ^{18}F-FDG avid. In addition, using a cutoff value of SUVmax ≥ 1.82 to define ^{18}F-FES positivity, 17 lesions were ^{18}F-FES negative (nine lymph nodes, six bone lesions, one lung metastatic, and one soft tissue) in 12 (34.3%) of 35 patients, showing remarkable heterogeneity of ER expression in these metastatic breast cancer patients. Interestingly, one patient had liver metastases and also had FES-negative metastases elsewhere, so this patient was included in the 12 patients with heterogeneous ER expression.

On the ^{18}F-FDG scan, the median SUVmax values among all lesions were 4.92 (range 1.68–40.74). On the ^{18}F-FES scan, the median SUVmax values among 217 ^{18}F-FES positive lesions (excluding 17 ^{18}F-FES negative lesions and one liver metastatic) was 4.7 (range 1.8–22.8). On a per-patient level, the median SUVmax of ^{18}F-FDG and ^{18}F-FES was 4.4 (range 2.1–15.5) and 4.5 (range 2.0–13.5), respectively. For patients with entirely ^{18}F-FES positive lesions, the median ratio of FES/FDG SUVmax was 0.96 (range 0.2–3.2). The detailed PET parameters of each patient are shown in **Supplemental Table 1**.

Prediction of Response to Fulvestrant

We first examined the significance of conventional clinical parameters. Patients with disease-free interval (DFI) ≥ 5 years

TABLE 1 | Patient demographics and disease characteristics.

Characteristics	*N* = 35	%
Median age, years	56.0 (40-78)	
Menopausal status		
Premenopause[a]	7	20.0
Postmenopause	28	80.0
Histology of primary breast cancer		
Ductal	29	82.8
Lobular	4	11.4
Mucinous	1	2.9
Tubular	1	2.9
DFI[b]		
≤5 y	10	28.6
>5 y	16	45.7
PgR status		
Positive	31	88.6
Negative	4	11.4
Metastatic sites		
Non-visceral	25	71.4
Bone	20	57.1
Bone-only	5	14.3
Visceral disease	10	28.6
Any lung	7	20.0
Pleural	5	14.3
Liver	1	2.9
No. of disease sites		
1	16	45.7
2	13	27.1
≥ 3	6	17.1
De novo metastatic disease	9	25.7
Prior line of therapies for metastatic disease		
0	28	80.0
1	5	14.3
2	2	5.7
Prior ET for metastatic disease		
None	30	85.7
Yes	5	14.3
Prior chemotherapy for metastatic disease		
None	31	88.6
Yes	4	11.4
PFS		
Events	26	74.3
Censored	9	25.7

ER, estrogen receptor; PgR, progesterone receptor; ET, endocrine therapy; PFS, Progression-Free Survival; DFI, Disease-free interval.

[a]*For premenopausal women, fulvestrant was given upon on the administration of gonadotropin-releasing hormone agonist.*

[b]*Patients with stage IV breast cancer at initial diagnosis were excluded (N = 9).*

had a longer PFS compared to those with less time of DFI (median PFS 12.2 months vs. 3.1 months, $P = 0.047$). However, this was of borderline significance in univariate analysis ($P = 0.054$). Other clinical risk factors (age, menopausal status, presence of visceral disease, *de novo* metastatic disease, histology of primary breast cancer, number of disease sites, bone-only disease, prior palliative chemotherapy, and lines of endocrine therapy for MBC) were not significantly related to PFS (**Table 2**).

Next, we tested whether the PET parameters correlate with survival in patients treated with Fulvestrant. The cutoff value of SUVmax of determined by the median value of ^{18}F-FDG and ^{18}F-FES was 4.4 and 4.5, respectively. It is regrettable that neither of the single parameters of the two scans was significantly associated with PFS ($P > 0.05$) (**Table 2**). We also analyzed the data stratified by low/high FDG and FES SUVmax, and there is no predictive value of PFS (**Supplemental Figure 1**).

Given the significant heterogeneity of ER expression in these patients with metastatic breast cancer, they may fail to respond to endocrine therapy. The population was stratified in three categories: 1) The heterogeneous group ($n = 12$) had both ^{18}F-FES negative and positive sites (**Figure 1**); 23 patients with entirely ^{18}F-FES positive lesions further divided into two groups by the median FES/FDG SUVmax ratio (the median value was 0.96). 2) The other groups are the low FES/FDG group (FES/FDG < 0.96, n= 11, **Figure 2A**) and 3) the high FES/FDG group (FES/FDG ≥ 0.96, $n = 12$, **Figure 2B**). Patients with the heterogeneity of ER expression were significantly associated with shorter PFS compared to those without ^{18}F-FES negative lesions in univariate analysis ($P < 0.001$, **Figure 3**). Median PFS was 5.5 months (95% CI 2.3–8.7) for the heterogeneous group, 29.4 months (95% CI 2.3–56.5) for the low FES/FDG group, and 14.7 months (95% CI 10.918.5) for the High FES/FDG (**Table 2**).

In multivariate analysis, a three-way PET classifier (FES heterogeneous, low FES/FDG, and high FES/FDG groups) remained the only independent, statistically significant prognostic factor for PFS ($P = 0.006$). Although DFI was a trend in the log-rank test, they were not considered as statistically independent prognostic factors ($P = 0.052$).

DISCUSSION

Our results have demonstrate that an integrated parameter derived from ^{18}F-FDG and ^{18}F-FES PET may have prognostic value for fulvestrant therapy in patients with ER-positive metastatic breast cancer. In our relatively small cohort, all clinical risk factors and single PET parameters were not significantly associated with PFS on multivariate analyses, whereas the PET classifier of ^{18}F-FDG and ^{18}F-FES remained significant.

Other scholars and our previous studies have confirmed that ^{18}F-FES PET can noninvasively and systematically assess ER status in patients with recurrent or metastatic breast cancer, and as an imaging biomarker for predicting response to endocrine therapy (13, 17, 20, 21). Nevertheless, ^{18}F-FES PET is challenging to monitor nonfunctional ER lesions, which might potentially lead to losing sight of ER-negative lesions. Some studies have used ^{18}F-FDG PET/CT together with ^{18}F FES-PET for the identification of ^{18}F FES negative lesions (22).

To our knowledge, this is the first dual-tracer PET study evaluating the effect of fulvestrant on ^{18}F-FDG and ^{18}F-FES in patients with ER-positive MBC. Several previous studies have described the prognostic value of single ^{18}F-FDG or ^{18}F-FES PET in ER-positive MBC (9, 12, 13, 23). However, these studies have certain limitations, such as under a specific population or needng a period of treatment to play a predictive role. Another study investigates the utility of ^{18}F-FDG and ^{18}F-FES PET on variety endocrine therapy in patients with ER-positive MBC but did not attempt to predict the efficacy of fulvestrant precisely (11). Consistent with other ^{18}F-FES PET studies, our results indicated that baseline ^{18}F-FES SUVmax was not correlated

TABLE 2 | Univariate and multivariate Cox regression analyses for prediction of progression-free survival (PFS).

Parameters	No.	Event	Median PFS	Log-rank	Univariate analysis		Multivariate analysis	
			(95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age								
<65	25	19	7.0(5.9-8.2)	0.318	0.64(0.27-1.54)	0.322	NA	
≥65	10	6	15.5(11.0-20.0)					
Menopausal status								
Pre-menopause	7	6	5.6(5.3-5.9)	0.416	0.68(0.27-1.73)	0.419	NA	
Post-menopause	28	20	13.1(3.8-22.4)					
Disease-free interval								
≤5 y	10	10	3.1(0.0-6.4)	0.047*	0.42(0.17-1.01)	0.054	/	0.052
>5 y	16	11	12.2(1.5-22.9)					
Histology of primary breast cancer								
Ductal	29	22	9.5(0.1-18.9)	0.593	0.72(0.21-2.43)	0.595	NA	
Lobular	4	3	14.7(4.0-20.4)					
No. of disease sites								
1	16	11	12.2(2.7-21.6)	0.202				
2	13	9	13.1(3.7-22.5)		1.13(0.46-2.78)	0.267	NA	
≥3	6	6	5.6(4.7-19.7)		2.42(0.86-6.82)			
Visceral disease								
No	25	19	9.5(1.6-17.5)	0.440	0.71(0.29-1.71)	0.443	NA	
Yes	10	7	13.8(0.0-27.7)					
Bone only disease								
No	30	23	12.2(3.8-20.7)	0.709	1.26(0.24-2.68)	0.710	NA	
Yes	5	3	2.4(2.0-2.8)					
De novo metastatic disease								
No	26	21	7.0(0.0-14.3)	0.167	0.51(0.19-1.36)	0.176	NA	
Yes	9	5	18.4(10.9-25.9)					
Prior palliative chemotherapy								
No	31	22	12.2(3.4-21.0)	0.516	1.43(0.48-4.22)	0.518	NA	
Yes	4	4	6.6(0.0-16.8)					
Lines of endocrine therapy for MBC								
1	28	20	12.2(3.1-21.3)	0.479	1.39(0.55-3.51)	0.482	NA	
≥2	7	6	6.6(5.3-7.9)					
FDG SUVmax								
<4.4	17	12	15.5(9.2-21.8)	0.186	1.72(0.76-3.91)	0.192	NA	
≥4.4	18	14	6.6(5.6-7.6)					
FES SUVmax								
<4.5	17	13	12.2(2.9-21.5)	0.995	0.99(0.45-2.19)	0.995	NA	
≥4.5	18	13	7.0(0.0-20.5)					
FES/FDG ratio								
With FES negative	12	12	5.5(2.3-8.7)	<0.001*		<0.001*		0.006*
<0.96	11	6	29.4 (2.3-56.5)		0.09(0.03-0.32)		0.10(0.02-0.49)	
≥0.96	12	8	14.7 (10.9-18.5)		0.22(0.08-0.59)		0.27(0.09-0.78)	

PFS, progression-free survival; CI, confidence interval; HR, hazard ratio; MBC, metastatic breast cancer; SUVmax, maximum standard uptake value.
**P ≤ 0.05; N/A: Analysis not performed as univariate analysis not significant.*

with treatment outcome (9). The predictive value of ^{18}F-FDG PET in patients with ER-positive MBC for fulvestrant therapy was proved by our previous study (23). In the current study, however, we did not find that sole ^{18}F-FDG SUVmax provides independent prognostic information for fulvestrant. An explanation for this could be that the populations of the two studies were different because patients with only bone metastasis were excluded from the previous research.

Kurland and colleagues' study demonstrated that the FES/FDG ratio appears to provide a reasonable summary of synchronous ER expression for patients with highly discordant ^{18}F-FES uptake across tumor sites in predicting clinical response to endocrine therapy (22). Furthermore, based on our previous study it was proposed that the ratio of SUVmax-FES/FDG showing potential in predicting neoadjuvant chemotherapy response of breast cancer (24). Besides, one study has also suggested that patients with low or absent ^{18}F-FES uptake in metastases may be unlikely to benefit from endocrine therapy (25). Our recent research showed that patients with entirely ^{18}F-FES positive lesions have a median PFS that is nearly twice of patients with negative ^{18}F-FES (14.6 months vs. 7.2 months), and the difference was of borderline significance (P = .081) (13). Hence, we hypothesized that the ratio of FES/FDG and the heterogeneous uptake of FES would predict response to fulvestrant therapy. Therefore, the current study combines the above two concepts into one classification scheme by sorting patients into three groups (heterogeneous disease, low FES/FDG, and high FES/FDG) based on ^{18}F-FDG and ^{18}F-FES PET scans. Our results suggest that, for ER-positive MBC, patients with the heterogeneity of ER expression by ^{18}F-FES PET were unlikely to

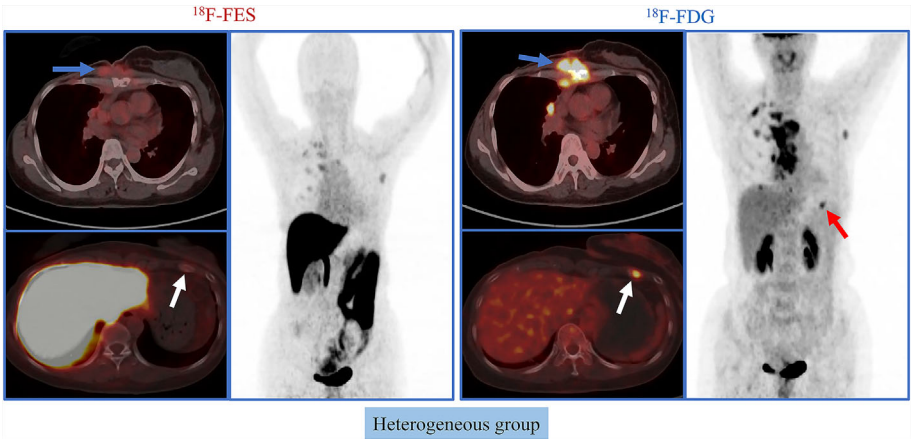


FIGURE 1 | Representative cases of heterogeneous group. A 50-year-old female patient has both ^{18}F-FES positive and negative lesions. The left rib shows significant uptake on FDG but not on FES. For this patient, the PFS was 3.7 months, and she did not receive clinical benefit from fulvestrant treatment.

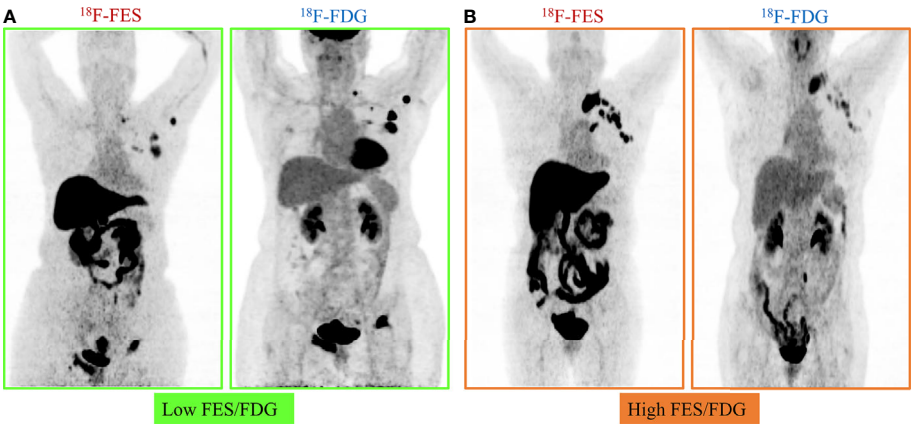


FIGURE 2 | Representative cases of FES/FDG group. Patients with 100% of the ^{18}F-FES positive metastatic lesions were divided into two groups by the median ratio of FES/FDG SUVmax (0.96). **(A)** Low FES/FDG. A 59-year-old female patient with the range of ^{18}F-FDG and ^{18}F-FES SUVmax was 5.3–40.7 and 4.1–15.5, respectively. This patient's median FES/FDG was 0.52, which was lower than the median FES/FDG of all patients. She has received fulvestrant treatment for 27.6 months until progress. **(B)** High FES/FDG. A 67-year-old female patient with the range of ^{18}F-FDG and ^{18}F-FES SUVmax was 3.0–8.1 and 8.8–16.0, respectively. This patient's median FES/FDG was 2.32, which was higher than the median FES/FDG of all patients, and the PFS was 14.7 months.

benefit from fulvestrant, and it may indicate that potential changes in ER expression of tumor in explaining endocrine therapy resistance, whereas patients with totally ^{18}F-FES positive in metastases are potential candidates for fulvestrant, particularly those with low FES/FDG.

Our study reports an incremental refinement of the classifier by integrating both ^{18}F-FES and ^{18}F-FDG imaging results based upon a smaller cohort but a more uniformly treated patient population compared with previous studies. Kurland et al. demonstrate that patients with low FDG uptake (indolent tumors) had a longer median PFS with high FDG uptake and high average FES uptake had a moderate median PFS and with high FDG uptake and low average FES uptake had a shorter median PFS (11). Nevertheless, this study differed from our current research in several respects. Patients had received different kinds of endocrine therapy, including aromatase inhibitor combined with or without fulvestrant, tamoxifen, and fulvestrant. Moreover, patients with the heterogeneity of ER expression were not individually analyzed, and those patients tended to develop resistance to endocrine therapy. We, therefore, analyzed those patients with the heterogeneity of ER expression independently, which may not benefit from sole fulvestrant, and it might be better to change management by adding complementary treatments, such as chemotherapy, everolimus, or cyclin-dependent kinases 4 and 6 (CDK4/6) inhibitors. In the current study, we report that patients with 100% ^{18}F-FES positive and low FES/FDG had a longer median PFS (29.4 months, 95% CI 2.3–56.5) compared with high FES/FDG (14.7 months, 95% CI 10.9–18.5). Consistent with our previous study, patients with high baseline ^{18}F-FDG tumor

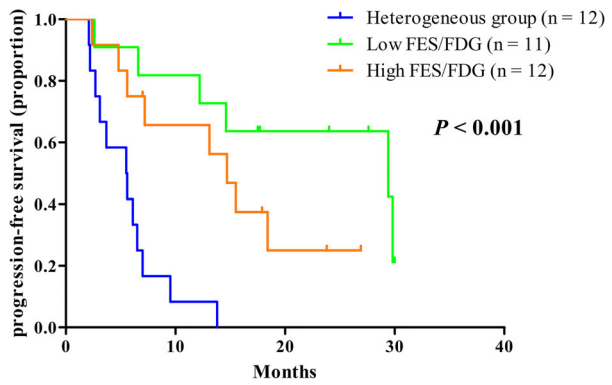


FIGURE 3 | Kaplan-Meier curves of PFS stratified by the three classification groups. Heterogeneous group (n = 12, median PFS 5.5 months), low ratio of FES/FDG (n = 11, median PFS 29.4 months), high ratio of FES/FDG (n = 12, median PFS 14.7 months).

uptake had a longer PFS (23); one of the possible reasons is that 17β-estradiol (E2) increases ER-dependent PI3K/Akt activation-mediated Glucose uptake signaling pathway in HR-positive breast cancer cell lines (26).

Increasing evidence suggests that, in addition to ER expression, progesterone receptor (PR) expression may also be related to the prognosis of fulvestrant therapy (27). The Zhao et al. study had reported that the ^{18}F-FDG/^{18}F-FES SUV ratio was correlated with ERα, PR expression (28). Therefore, the FES/FDG ratio may be more representative of comprehensive ER, PR expression, and could be a potential imaging biomarker to predict survival on fulvestrant therapy in patients with HR-positive breast cancer.

There were several limitations to our study. First, the retrospective nature of this study and the heterogeneous patient population, perhaps with inherently different prognostic factors, are a major limitation. Our study shows that patients with a high FES/FDG ratio have shorter survival than those with a low FES/FDG ratio. This is the opposite of what we expected: that patients with greater FES-avidity and lower FDG-avidity would be expected to lead to longer survival. However, most of the current studies indicate that patients with FES negative or positive lesions are related to prognosis; there is no direct linear relationship between the level of FES uptake and clinical outcomes. The FES/FDG ratio may reflect the two biological functions of hormone receptors and glucose metabolism in metastases and may be more valuable for predicting fulvestrant treatment. Third, the sample size was relatively modest. Despite the small cohort, the results were statistically significant.

CONCLUSIONS

Our data suggest that dual ^{18}F-FDG and ^{18}F-FES PET imaging could be a potential predictor of efficacy to fulvestrant therapy among HR+HER2- MBC patients. These findings indicate that endocrine therapy should be individualized for patients with ER-positive MBC, particularly the presence of ^{18}F-FES negative lesions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Fudan University Shanghai Cancer Center Ethic Committee and Institutional Review Boards. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conception and design: CL, ZY. Acquiring data, or analyzing and interpreting data: CL, XX, HY, YPZ. Drafting the manuscript: CL. Critically contributing to or revising the manuscript: SS, ZY. Enhancing its intellectual: ZY, SS, YJZ. All authors contributed to the article and approved the submitted version.

FUNDING

This research is sponsored by Shanghai Sailing Program (20YF1408500), the Shanghai Committee of Science and Technology Fund (No.19ZR1411300), the Shanghai Engineering Research Center of Molecular Imaging Probes Program (No. 19DZ2282200) and Shanghai Municipal Health Commission (No.202040267).

ACKNOWLEDGMENTS

The authors would like to thank my colleagues in the Multiple Disciplinary Team of Breast Cancer and the Nuclear Medicine team at Fudan University Shanghai Cancer Center.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.580277/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70:7–30. doi: 10.3322/caac.21590
2. Lobbezoo DJ, van Kampen RJ, Voogd AC, Dercksen MW, van den Berkmortel F, Smilde TJ, et al. Prognosis of metastatic breast cancer subtypes: the hormone receptor/HER2-positive subtype is associated with the most favorable outcome. *Breast Cancer Res Treat* (2013) 141:507–14. doi: 10.1007/s10549-013-2711-y
3. Lei W, Li H, Song G, Zhang R, Ran R, Yan Y. Efficacy and Safety of Fulvestrant 500mg in Hormone-receptor Positive Human Epidermal Receptor 2 Negative Advanced Breast Cancer: A Real-world Study in China. *J Cancer* (2020) 11(22):6612–22. doi: 10.7150/jca.47960
4. Wang J, Xu B, Wang W, Zhai X, Chen X. Efficacy and safety of fulvestrant in postmenopausal patients with hormone receptor-positive advanced breast cancer: a systematic literature review and meta-analysis. *Breast Cancer Res Treat* (2018) 171:535–44. doi: 10.1007/s10549-018-4867-y
5. Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* (2011) 378:771–84. doi: 10.1016/S0140-6736(11)60993-8
6. Liao GJ, Clark AS, Schubert EK, Mankoff DA. 18F-Fluoroestradiol PET: Current Status and Potential Future Clinical Applications. *J Nucl Med* (2016) 57:1269–75. doi: 10.2967/jnumed.116.175596
7. van Kruchten M, Glaudemans AW, de Vries EF, Beets-Tan RG, Schroder CP, Dierckx RA, et al. PET imaging of estrogen receptors as a diagnostic tool for breast cancer patients presenting with a clinical dilemma. *J Nucl Med* (2012) 53:182–90. doi: 10.2967/jnumed.111.092734
8. Linden HM, Stekhova SA, Link JM, Gralow JR, Livingston RB, Ellis GK, et al. Quantitative fluoroestradiol positron emission tomography imaging predicts response to endocrine treatment in breast cancer. *J Clin Oncol* (2006) 24:2793–9. doi: 10.1200/JCO.2005.04.3810
9. van Kruchten M, de Vries EG, Glaudemans AW, van Lanschot MC, van Faassen M, Kema IP, et al. Measuring residual estrogen receptor availability during fulvestrant therapy in patients with metastatic breast cancer. *Cancer Discovery* (2015) 5:72–81. doi: 10.1158/2159-8290.CD-14-0697
10. van Kruchten M, Glaudemans A, de Vries EFJ, Schroder CP, de Vries EGE, Hospers GAP. Positron emission tomography of tumour [(18)F] fluoroestradiol uptake in patients with acquired hormone-resistant metastatic breast cancer prior to oestradiol therapy. *Eur J Nucl Med Mol Imaging* (2015) 42:1674–81. doi: 10.1007/s00259-015-3107-5
11. Kurland BF, Peterson LM, Lee JH, Schubert EK, Currin ER, Link JM, et al. Estrogen Receptor Binding (18F-FES PET) and Glycolytic Activity (18F-FDG PET) Predict Progression-Free Survival on Endocrine Therapy in Patients with ER+ Breast Cancer. *Clin Cancer Res* (2017) 23:407–15. doi: 10.1158/1078-0432.CCR-16-0362
12. Linden HM, Kurland BF, Peterson LM, Schubert EK, Gralow JR, Specht JM, et al. Fluoroestradiol positron emission tomography reveals differences in pharmacodynamics of aromatase inhibitors, tamoxifen, and fulvestrant in patients with metastatic breast cancer. *Clin Cancer Res* (2011) 17:4799–805. doi: 10.1158/1078-0432.CCR-10-3321
13. He M, Liu C, Shi Q, Sun Y, Zhang Y, Xu X, et al. The Predictive Value of Early Changes in (18) F-Fluoroestradiol Positron Emission Tomography/Computed Tomography During Fulvestrant 500 mg Therapy in Patients with Estrogen Receptor-Positive Metastatic Breast Cancer. *Oncologist* (2020). doi: 10.1634/theoncologist.2019-0561
14. Arnon-Sheleg E, Israel O, Keidar Z. PET/CT Imaging in Soft Tissue Infection and Inflammation-An Update. *Semin Nucl Med* (2020) 50:35–49. doi: 10.1053/j.semnuclmed.2019.07.005
15. Liu C, Gong C, Liu S, Zhang Y, Zhang Y, Xu X, et al. (18)F-FES PET/CT Influences the Staging and Management of Patients with Newly Diagnosed Estrogen Receptor-Positive Breast Cancer: A Retrospective Comparative Study with (18)F-FDG PET/CT. *Oncologist* (2019) 24:e1277–85. doi: 10.1634/theoncologist.2019-0096
16. Gong C, Yang Z, Sun Y, Zhang J, Zheng C, Wang L, et al. A preliminary study of (18)F-FES PET/CT in predicting metastatic breast cancer in patients receiving docetaxel or fulvestrant with docetaxel. *Sci Rep* (2017) 7:6584. doi: 10.1038/s41598-017-06903-8
17. Yang Z, Sun Y, Xu X, Zhang Y, Zhang J, Xue J, et al. The Assessment of Estrogen Receptor Status and Its Intratumoral Heterogeneity in Patients With Breast Cancer by Using 18F-Fluoroestradiol PET/CT. *Clin Nucl Med* (2017) 42:421–7. doi: 10.1097/RLU.0000000000001587
18. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging* (2015) 42:328–54. doi: 10.1007/s00259-014-2961-x
19. Nienhuis HH, van Kruchten M, Elias SG, Glaudemans A, de Vries EFJ, Bongaerts AHH, et al. (18)F-Fluoroestradiol Tumor Uptake Is Heterogeneous and Influenced by Site of Metastasis in Breast Cancer Patients. *J Nucl Med* (2018) 59:1212–8. doi: 10.2967/jnumed.117.198846
20. Lin FI, Gonzalez EM, Kummar S, Do K, Shih J, Adler S, et al. Utility of (18)F-fluoroestradiol ((18)F-FES) PET/CT imaging as a pharmacodynamic marker in patients with refractory estrogen receptor-positive solid tumors receiving Z-endoxifen therapy. *Eur J Nucl Med Mol Imaging* (2017) 44:500–8. doi: 10.1007/s00259-016-3561-8
21. Chae SY, Ahn SH, Kim SB, Han S, Lee SH, Oh SJ, et al. Diagnostic accuracy and safety of 16alpha-[(18)F]fluoro-17beta-oestradiol PET-CT for the assessment of oestrogen receptor status in recurrent or metastatic lesions in patients with breast cancer: a prospective cohort study. *Lancet Oncol* (2019) 20:546–55. doi: 10.1016/S1470-2045(18)30936-7
22. Kurland BF, Peterson LM, Lee JH, Linden HM, Schubert EK, Dunnwald LK, et al. Between-patient and within-patient (site-to-site) variability in estrogen receptor binding, measured in vivo by 18F-fluoroestradiol PET. *J Nucl Med* (2011) 52:1541–9. doi: 10.2967/jnumed.111.091439
23. Zhao Y, Liu C, Zhang Y, Gong C, Li Y, Xie Y, et al. Prognostic Value of Tumor Heterogeneity on 18F-FDG PET/CT in HR+HER2- Metastatic Breast Cancer Patients receiving 500 mg Fulvestrant: a retrospective study. *Sci Rep* (2018) 8:14458. doi: 10.1038/s41598-018-32745-z
24. Yang Z, Sun Y, Xue J, Yao Z, Xu J, Cheng J, et al. Can positron emission tomography/computed tomography with the dual tracers fluorine-18 fluoroestradiol and fluorodeoxyglucose predict neoadjuvant chemotherapy response of breast cancer?–A pilot study. *PloS One* (2013) 8:e78192. doi: 10.1371/journal.pone.0078192
25. Peterson LM, Kurland BF, Schubert EK, Link JM, Gadi VK, Specht JM, et al. A phase 2 study of 16alpha-[18F]-fluoro-17beta-estradiol positron emission tomography (FES-PET) as a marker of hormone sensitivity in metastatic breast cancer (MBC). *Mol Imaging Biol* (2014) 16:431–40. doi: 10.1007/s11307-013-0699-7
26. Garrido P, Moran J, Alonso A, Gonzalez S, Gonzalez C. 17beta-estradiol activates glucose uptake via GLUT4 translocation and PI3K/Akt signaling pathway in MCF-7 cells. *Endocrinology* (2013) 154:1979–89. doi: 10.1210/en.2012-1558
27. Fowler AM, Chan SR, Sharp TL, Fettig NM, Zhou D, Dence CS, et al. Small-animal PET of steroid hormone receptors predicts tumor response to endocrine therapy using a preclinical model of breast cancer. *J Nucl Med* (2012) 53:1119–26. doi: 10.2967/jnumed.112.103465
28. Zhao Z, Yoshida Y, Kurokawa T, Kiyono Y, Mori T, Okazawa H. 18F-FES and 18F-FDG PET for differential diagnosis and quantitative evaluation of mesenchymal uterine tumors: correlation with immunohistochemical analysis. *J Nucl Med* (2013) 54:499–506. doi: 10.2967/jnumed.112.113472

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 11 November 2020
doi: 10.3389/fonc.2020.567160

Value of ^{18}F-FDG PET/CT-Based Radiomics Nomogram to Predict Survival Outcomes and Guide Personalized Targeted Therapy in Lung Adenocarcinoma With EGFR Mutations

Bin Yang[1†], Hengshan Ji[2†], Jing Zhong[1†], Lu Ma[1†], Jian Zhong[1†], Hao Dong[3†], Changsheng Zhou[1], Shaofeng Duan[4], Chaohui Zhu[5], Jiahe Tian[6], Longjiang Zhang[1], Feng Wang[7], Hong Zhu[2] and Guangming Lu[1*]*

[1] Department of Medical Imaging, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing, China, [2] Department of Nuclear Medicine, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing, China, [3] College of Medical Imaging, Xuzhou Medical University, Xuzhou, China, [4] Institute of Precision Medicine, GE Healthcare China, Shanghai, China, [5] Department of Nuclear Medicine, Peking Union Medical College Hospital, Beijing, China, [6] Department of Nuclear Medicine, The Chinese People's Liberation Army (PLA) General Hospital, Beijing, China, [7] Department of Nuclear Medicine, First People's Hospital of Nanjing, Nanjing, China

Objectives: To investigate the development and validation of a radiomics nomogram based on PET/CT for guiding personalized targeted therapy in patients with lung adenocarcinoma mutation(s) in the *EGFR* gene.

Methods: A cohort of 109 (77/32 in training/validation cohort) consecutive lung adenocarcinoma patients with an *EGFR* mutation was enrolled in this study. A total of 1672 radiomic features were extracted from PET and CT images, respectively. The least absolute shrinkage and selection operator (LASSO) Cox regression was used to select the radiomic features and construct the radiomics nomogram for the estimation of overall survival (OS), which was then assessed with respect to calibration and clinical usefulness. Patients with an *EGFR* mutation were divided into high- and low- risk groups according to their nomogram score. The treatment strategy for high- and low-risk groups was analyzed using Kaplan–Meier analysis and a log-rank test.

Results: The C-index of the radiomics nomogram for the prediction of OS in lung adenocarcinoma in patients with an *EGFR* mutation was 0.840 and 0.803 in the training and validation cohorts, respectively. Distant metastasis [(Hazard ratio, HR),1.80], metabolic tumor volume (MTV, HR, 1.62), and rad score (HR, 17.23) were the independent risk factors for patients with an *EGFR* mutation. The calibration curve showed that the predicted survival time was remarkably close to the actual time. Decision curve analysis demonstrated that the radiomics nomogram was clinically useful. Targeted therapy for patients with high-risk *EGFR* mutations attained a greater

OPEN ACCESS

Edited by:
Laurent Dercle, Columbia University Irving Medical Center, United States

Reviewed by:
Yuming Jiang, Stanford University, United States
Liyu Huang, Xidian University, China
Ying Liu, Tianjin Medical University Cancer Institute and Hospital, China

***Correspondence:**
Guangming Lu
cjr.luguangming@vip.163.com
Hong Zhu
zh_zy@163.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *29 May 2020*
Accepted: *05 October 2020*
Published: *11 November 2020*

Citation:
Yang B, Ji H, Zhong J, Ma L, Zhong J, Dong H, Zhou C, Duan S, Zhu C, Tian J, Zhang L, Wang F, Zhu H and Lu G (2020) Value of ^{18}F-FDG PET/CT-Based Radiomics Nomogram to Predict Survival Outcomes and Guide Personalized Targeted Therapy in Lung Adenocarcinoma With EGFR Mutations. Front. Oncol. 10:567160. doi: 10.3389/fonc.2020.567160

benefit than other therapies ($p < 0.0001$), whereas the prognoses of the two therapies were similar in the low-risk group ($p = 0.85$).

Conclusions: Development and validation of a radiomics nomogram based on PET/CT radiomic features combined with clinicopathological factors may guide targeted therapy for patients with lung adenocarcinoma with *EGFR* mutations. This is conducive to the advancement of precision medicine.

Keywords: lung adenocarcinoma, positron emission tomography/computed tomography, radiomics, nomogram, targeted therapy

INTRODUCTION

Lung cancer is the leading cause of cancer deaths in the world and has the highest morbidity and mortality rates among all malignant tumors (1, 2). Non-small cell lung cancer (NSCLC) accounts for 85% of all lung cancers (3, 4). Due to the lack of early clinical symptoms, lymph node metastasis or distant metastasis has already occurred by the time of diagnosis, and it is usually too late for surgical intervention (5, 6). Although the prognosis of lung cancer has improved significantly with improvements in treatment methods, the 5-year survival rate for lung cancer patients remains at 17–18% (7, 8).

The tumor, node, and metastasis (TNM) staging system is currently the most valuable and commonly used tumor staging system for assessing the prognosis of malignant tumors (9–12). However, in clinical practice, it is found that the TNM staging system continues to have many shortcomings in the prognostic evaluation of lung cancer. The survival time of patients at the same stage may differ. Therefore, a TNM-based one-size-fits-all strategy might not be suitable for all patients. In addition, it is not currently possible to fully predict the progression and outcome of disease in patients with NSCLC. Therefore, identification of patients at high risk of death would be valuable for guiding therapy (13–15). New methods of prognostic assessment are urgently needed to achieve personalized treatment. A nomogram is an intuitive chart prepared by establishing a statistical prediction model, which includes important tumor prognosis factors. A nomogram is regarded as a tool for quantifying risks and has become the focus of cancer research (16–18).

The ^{18}F-fluordeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) can provide functional, metabolic, anatomical, and morphological imaging. Its' metabolic parameters can reflect the metabolism of tumor tissue. Studies have shown that FDG uptake in primary tumors is an independent risk factor for patients with early NSCLC (19, 20), although the value of the prognosis in evaluation of advanced NSCLC patients remains controversial (21, 22). Moreover, the ^{18}F-FDG PET/CT features of lung cancer are significantly correlated with T stages, N status, pathological stages, and tumor grades (23–25). Therefore, it has been widely used in the diagnosis, staging, and monitoring of the therapeutic effects and prognostic evaluation of NSCLC (26). Radiomics is the high-throughput extraction and analysis of quantitative features from images. Consequently, the prognostic evaluation of NSCLC by PET/CT can be improved (27). Currently, several attempts have been made to improve the performance of predictive models. However, the prognostic prediction performance of radiomics models in these studies was generally poor. Thus the prognostic performance of radiomics has room for further improvement (15, 28). A few studies have evaluated the use of ^{18}F-FDG PET/CT radiomics features to predict the NSCLC prognosis; nevertheless the effect of the driver gene mutation status and treatment methods was ignored. The prognosis of patients with NSCLC is closely related to the driving gene mutation status and treatment. So, it is necessary to conduct independent research with these patients to achieve individualized treatment.

The main purpose of this study was to develop a radiomics nomogram based on ^{18}F-FDG PET/CT radiomic features combined with clinicopathological factors to predict the survival outcomes of patients diagnosed with lung adenocarcinoma with an epidermal growth factor receptor (*EGFR*) mutation. We also endeavored to provide guidance for treatment strategies and prognostic evaluation of patients with an *EGFR* mutation.

MATERIALS AND METHODS

Patients

The institutional review board of Affiliated Jinling Hospital, Medical School of Nanjing University approved this retrospective study and waived the requirement to obtain informed consent from the patients. In our retrospective investigation, the following inclusion criteria were applied to select patients from the medical database: a) an ^{18}F-FDG PET/CT examination within 1 month prior to surgery or biopsy, b) no anti-tumor treatment received before the ^{18}F-FDG PET/CT examination, c) with surgical or biopsy specimens confirmed by pathology, and d) with *EGFR* mutation detection results. The exclusion criteria were as follows: a) patients with partial loss of PET or CT images, b) patients with metastases in the lung, and

Abbreviations: CEA, carcinoembryonic antigen; EGFR, epidermal growth factor receptor; MTV, metabolic tumor volume; NSCLC, non-small cell lung cancer; PET/CT, positron emission tomography/computed tomography; SUVmax, maximal standard uptake value; SUV_{mean}, mean standard uptake value; TKIs, tyrosine kinase inhibitors; TLG, total lesion glycolysis.

c) images with unclear boundaries of the tumor that could not be accurately delineated.

Altogether, 174 consecutive lung adenocarcinoma patients were identified by applying the above-mentioned inclusion/exclusion criteria from the institutional database between July 2009 and August 2016, and 109 cases were patients with an EGFR mutation. Among those with EGFR mutations, 44 had the 19DEL, 61 had the 21L858R-mutation and four had other EGFR mutations sites. We randomly divided patients with the EGFR mutation into training (n = 77) and validation (n = 32) cohorts following a 7:3 ratio. The clinicopathological data obtained from medical records included age, sex, family history, smoking history, histological grade, lymph node metastasis, distant metastasis, TNM stage (defined according to the eighth edition of the TNM classification and staging system by the American Joint Committee on Cancer), thyroid transcription factor-1 (TTF-1) (− or one + was defined as negative, ≥two + was defined as positive), Ki-67 (≤25% was defined as low expression and >25% as high expression), carcinoembryonic antigen (CEA), and PET/CT metabolic parameters (**Table 1**). The follow-up time was from July 2009 to January 2019. The endpoint of this study was overall survival (OS), which was defined as the time from the date of the ^{18}F-FDG PET/CT examination to the date of telephone follow-up or the date of the patient's death.

TABLE 1 | Characteristics of the training and validation cohorts.

Characteristics	Training cohort (n = 77)	Validation cohort (n = 32)	Total(n = 109)	*p*-value
Gender-no.(%)				1.000
Female	43 (55.844)	18 (56.250)	61 (55.963)	
Male	34 (44.156)	14 (43.750)	48 (44.037)	
Age, mean(SD)	60.078 (9.373)	60.625 (8.051)	60.239 (8.972)	0.773
Family history-no.(%)				0.036
No	75 (97.403)	27 (84.375)	102 (93.578)	
Yes	2 (2.597)	5 (15.625)	7 (6.422)	
Smoking status-no.(%)				0.835
Non-smokers	57 (74.026)	25 (78.125)	82 (75.229)	
Smokers	20 (25.974)	7 (21.875)	27 (24.771)	
Histologic grade-no.(%)				0.376
Poorly differentiated	33 (42.857)	13 (40.625)	46 (42.202)	
Moderately differentiated	35 (45.455)	12 (37.500)	47 (43.119)	
Well differentiated	9 (11.688)	7 (21.875)	16 (14.679)	
Lymph node metastasis-no.(%)				0.157
Yes	62 (80.519)	21 (65.625)	83 (76.147)	
No	15 (19.481)	11 (34.375)	26 (23.853)	
Distant metastasis-no.(%)				0.455
Yes	53 (68.831)	25 (78.125)	78 (71.560)	
No	24 (31.169)	7 (21.875)	31 (28.440)	
Stage-no.(%)				0.988
I/II	9 (11.688)	3 (9.375)	12 (11.009)	
III/IV	68 (88.312)	29 (90.625)	97 (88.991)	
TTF-1-no.(%)				0.610
Positive	56 (72.727)	21 (65.625)	77 (70.642)	
Negative	21 (27.273)	11 (34.375)	32 (29.358)	
Ki-67-no.(%)				1.000
≤25%	49 (63.636)	20 (62.500)	69 (63.303)	
>25%	28 (36.364)	12 (37.500)	40 (36.697)	
CEA-no.(%)				0.344
≤2.60	18 (23.377)	11 (34.375)	29 (26.606)	
>2.60	59 (76.623)	21 (65.625)	80 (73.394)	
SUVmax-no.(%)				0.198
≤5.33	20 (25.974)	13 (40.625)	33 (30.275)	
>5.33	57 (74.026)	19 (59.375)	76 (69.725)	
SUVmean-no.(%)				0.511
≤1.74	7 (9.091)	5 (15.625)	12 (11.009)	
>1.74	70 (90.909)	27 (84.375)	97 (88.991)	
TLG(g)				0.880
≤54.02	51 (66.234)	20 (62.500)	71 (65.138)	
>54.02	26 (33.766)	12 (37.500)	38 (34.862)	
MTV(cm^3)				0.824
≤7.32	35 (45.455)	16 (50.000)	51 (46.789)	
>7.32	42 (54.545)	16 (50.000)	58 (53.211)	

CEA, carcinoembryonic antigen; MTV, metabolic tumor volume; SUV_{max}, maximal standard uptake value; SUV_{mean}, mean standard uptake value; TLG, total lesion glycolysis; TTF-1, thyroid transcription factor-1; EGFR, epidermal growth factor receptor.

PET/CT Imaging Method, Image Acquisition, and Measurement of Metabolic Parameters

Patients underwent PET/CT imaging (Biography 16, Siemens, Erlangen, Germany) using ^{18}F-FDG synthesized by the Canadian EBCO TR19 medical cyclotron and chemical synthesis system. The radiochemical purity was >95%. The patients fasted for 6–8 h before undergoing the scan. Patients were intravenously injected with ^{18}F-FDG (3.7–6.66 MBq/kg) and underwent a whole-body PET/CT scan from the skull base to the upper section of the thigh. CT scan parameters were as follows: tube voltage120 kV, Tube current 140 mAs, and layer thickness and layer spacing 5 mm, matrix 512 × 512, and tube rotation speed 0.8 s/r. The PET acquisition parameters were as follows: three-dimensional at 3 min/bed, iterative algorithm, iterations four subsets, eight resolution, 4.1 mm lateral, 4.6 mm axial, matrix 128 × 128, voxel size 5.3 × 5.3 × 5.3 mm^3. The images were reconstructed using an iterative reconstruction method resulting in CT, PET, and PET/CT fusion images that were transferred to a post-processing workstation. We used Microsoft Viewer software (version VB10, Siemens) to calculate the metabolic parameters on the PET images. PET images were first converted to SUV images in the software without other processing methods. Then, the three-dimensional region of interest (ROI) was manually delineated by a radiologist (YB) to calculate the maximum standard uptake value (SUVmax, with a threshold set to 40%), mean standard uptake value (SUVmean), and metabolic tumor volume (MTV). Subsequently, the total lesion glycolysis (TLG) (TLG = SUVmean × MTV) was calculated.

EGFR Gene Detection

EGFR genetic mutations were tested from the affected tumor tissue sample obtained by surgical resection or biopsy. The amplification refractory mutation system polymerase chain reaction method was used to detect mutation sites in four exons (exons 18–21) in the coding region of the EGFR gene, the results of which were acquired according to the interpretation principle provided by the reference test kit. If any exon mutation was detected, the tumor was identified as an EGFR mutant; otherwise, the tumor was identified as EGFR wild type.

Tumor Segmentation

A volume of interest (VOI) segmentation was semiautomatically produced by drawing a line across the boundary of the tumor and manually adjusted by a chest radiologist (YB, 9 years of experience in the lung diagnosis) in a three-dimensional domain using the radiomics prototype (Radiomics, Frontier, Siemens; **Figure 1**) and confirmed by another chest radiologist (JS, 15 years of experience). Then, the tool automatically found the neighboring voxels in 3D space with the same gray level through an automatic algorithm. This is the Random Walker-based lesion segmentation for solid and subsolid lung lesions (29). Both radiologists were blinded to the patients' clinical information. The details of the tumor segmentation are described in **Appendix 1**.

Feature Extraction, Feature Selection, and Radiomics Signature Construction

Our study followed and adhered to the Image Biomarker Standardization Initiative (IBSI) guidelines (30), and the software used was IBSI-compliant. The medical images were resampled to the 1 mm × 1 mm × 1 mm voxel size in millimeters before the subsequent feature extraction steps. The interpolator used for resampling was B-spline interpolation. For discretization of the image gray levels, the bin width was set as 25 for CT and 0.1 for PET-SUV. After preprocessing, a total of 1,672 × 2 radiomics features were extracted from the CT and PET images by the radiomics prototype after imaging preprocessing. The extracted radiomics feature groups were as

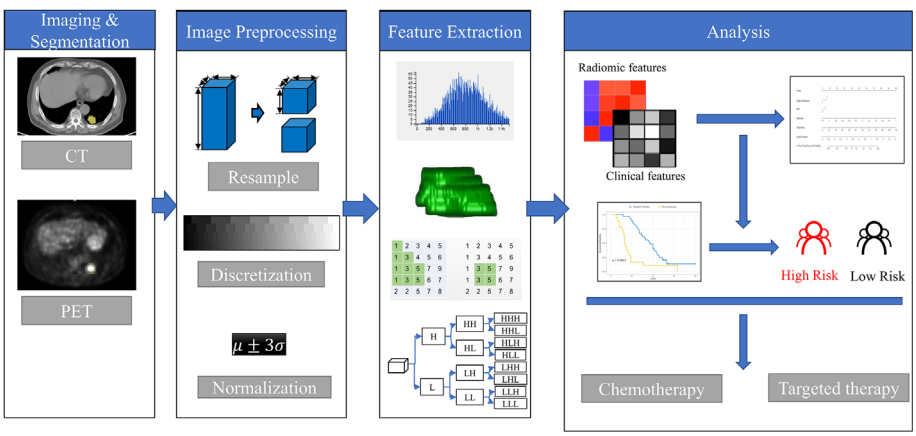


FIGURE 1 | The framework for developing the radiomics nomogram and treatment strategy decisions. The lesions were segmented on Siemens Radiomics prototype semiautomatically, and 1,672 radiomics features, including first order features, shape related features, and texture features were extracted using the software after image pre-processing. The least absolute shrinkage and selection operator (LASSO) Cox regression was used to select radiomics features and clinicopathological factors to construct the radiomics nomogram. Patients with *EGFR* mutations were divided into high- and low-risk groups according to the rad score. The treatment strategy was analyzed in the high- and low-risk groups.

follows: a) 18 first-order features, b) 16 size and shape features, and c) 74 texture encoding features. In total 1,672 radiomics features were extracted from each per lesion, including 108 from the original image, 460 [92 × 5] from the LoG-filtered images, 736 [92 × 8] from the wavelet-transformed images, and 368 [92 × 4] from non-linear intensity transforms (For detailed feature calculation formulas, please refer to the website: https://pyradiomics.readthedocs.io/en/latest/features.html#). A Spearman's correlation test was performed using the 'findCorrelation' function in the caret package (cutoff, 0.9) to reduce feature redundancy. The least absolute shrinkage and selection operator (LASSO) Cox regression method, which is suitable for the regression of high dimensional data in survival analyses, was conducted to select the most useful predictive features from the training cohort (31). A radiomics score (rad score) was calculated for each patient *via* a linear combination of selected features that were weighted by their respective coefficients (32).

Prognostic Model Establishment

The clinicopathological factors were analyzed using univariate Cox proportional hazards (CPH) regression analysis to identify significant risk factors. Significant risk factors with $p < 0.05$ were analyzed using the Kaplan–Meier curve and log-rank test. Significant risk factors were analyzed using multivariate Cox proportional hazards (CPH) regression analysis to identify independent risk factors. A clinical model was constructed based on the independent risk factors. Rad score and independent risk factors were fused into a single predictive model based on a multivariate CPH model. The performance of models was evaluated with the concordance index (C-index).

Construction of the Radiomics Nomogram and Its Performance

The rad score and independent risk factors were based on multivariate Cox regression analysis to construct the radiomics nomogram. The prediction performance of the radiomics nomogram was assessed using the Harrell's C-index in the training and validation cohorts. The C-index ranges from 0.5 to 1.0, where 0.5 indicates random data distribution and 1.0 suggests that the outcome of the model predicted the observed survival information perfectly. Calibration curves of the radiomics nomogram were then drawn for 5-year OS of the patients (33). The calibration curves illustrated both survival probabilities predicted by nomogram and the observed probabilities. A decision curve analysis determined the clinical usefulness of the radiomics nomogram by quantifying the net benefits at various threshold probabilities.

To Guide the Individualized Targeted Therapy for Patients With Lung Adenocarcinoma

Patients with an *EGFR* mutation were divided into high- and low-risk groups according to their nomogram score. The treatment strategy was explored separately in the high- and low-risk cohorts using Kaplan–Meier analysis and a log-rank test, to find the cohort that would benefit from the targeted treatment. Additionally, the various treatment strategies were explored in patients with different EGFR-mutation sites, to identify which patients could actually benefit from adjuvant therapy.

Statistical Analysis

The R software (version 3.5.0, www.Rproject.org) was used for all statistical analyses in this study. LASSO was conducted using the 'glmnet' package, while 'hdnom' was used for survival analysis. All statistical tests were two-sided and the significance level was set at $p = 0.05$.

RESULTS

Clinical Characteristics

Patient characteristics of the training and validation cohorts were summarized in **Table 1**. There were no significant differences in age, sex, smoking status, lymph node metastasis, or distant metastasis, *etc.*, between the two cohorts ($p > 0.05$).

Important Radiomics Feature Selection and Radiomics Signature Construction

In total, 1,672 radiomics features were extracted from the CT and PET images, respectively. We performed feature selection using the LASSO regression model with the PET/CT features (**Figures 2A, B**). The following ten important features were selected from 1,672 radiomics features (**Figure 2C**):

CT_wavelet−LLH_glcm_ClusterShade, CT_log−sigma−0−5−mm−3D_glcm_MaximumProbability,

CT_wavelet−LLH_firstorder_Skewness,PET_wavelet−HHL_firstorder_Mean,CT_wavelet−HLH_glcm_ClusterShade, PET_wavelet−HHL_glszm_SmallAreaLowGrayLevelEmphasis,CT_wavelet−LHL_glszm_SmallAreaHighGrayLevelEmphasis,PET_wavelet−HLL_firstorder_Kurtosis, PET_wavelet−LHL_glcm_Imc2, and CT_wavelet−LHL_firstorder_Mean.

Then the rad score was calculated using these ten radiomics features as follows: rad score = 0.051*PET_wavelet-HLL_firstorder_Kurtosis+0.006*PET_wavelet-HLL_glcm_Idn+-0.011*PET_wavelet-LHH_glcm_Imc1+0.047*PET_wavelet-LHL_glcm_Imc2+-0.011*PET_log-sigma-0-5-mm-3D_glszm_SmallAreaLowGrayLevelEmphasis + 0.093.

Prognostic Model Establishment and Performance of the Multimodality Prediction Model

We used a univariate Cox regression analysis to test the hazard ratio (HR) of each factor and to determine its' significance in the probability of death. The results were as follows: distant metastasis (HR, 2.68), metabolic tumor volume (MTV, HR, 2.02), maximal standard uptake value (SUVmax, HR, 2.48), stage (HR, 4.29), and carcinoembryonic antigen (CEA, HR, 3.16) were the significant risk factors for patients with an

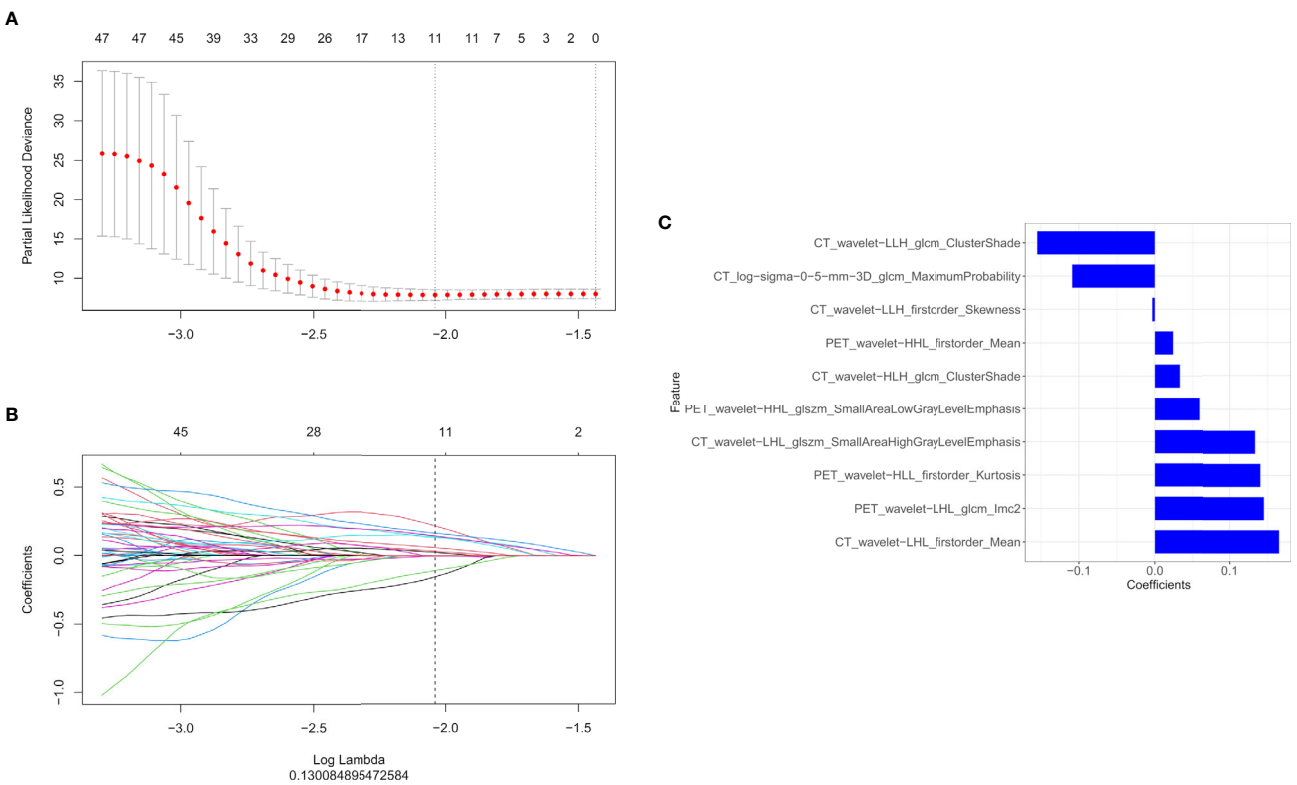


FIGURE 2 | The LASSO and ten-fold cross-validation were used to extract the optimal subset of radiomics features. The following two steps were included: determining the hyperparameter/lambda with a partial likelihood deviance as the criterion **(A)** and using the optimized/lambda (the vertical dashed line) to select features with non-zero coefficients **(B)**. **(C)** LASSO algorithm was used to select the ten radiomics features that contributed the most to the prognostic prediction model.

EGFR mutation ($P < 0.05$).The significant risk factors with $p < 0.05$ were calculated using a log-rank test, and Kaplan–Meier curves were plotted. **Figures 3A–E** illustrate the survival probability of patients in the high-risk or low-risk cohorts. The results of the log-rank test indicate significant discrimination between the two groups. A clinical model was constructed based on multivariate Cox proportional hazards (CPH) regression analysis of significant risk factors. Distant metastasis [HR,2.97 (95%CI, 1.36–6.51)] and metabolic tumor volume [MTV, HR,2.26(95%CI, 1.19–4.28)] were the independent risk factors in the training cohort. Rad score and independent risk factors were fused into a single predictive model based on the multivariate CPH regression analysis. Distant metastasis [HR,1.80(95%CI, 0.80–4.04)], metabolic tumor volume [MTV, HR,1.62(95%CI, 0.82–3.17)] and rad score [HR,17.23 (95%CI, 6.62–44.81)] were the independent risk factors in the training cohort. The C-index of the clinical model was 0.694 and 0.729 in the training and validation cohorts, respectively. The C-index of the rad score (radiomics model) was 0.819 and 0.737 in the training and validation cohorts, respectively. A rad score was combined with the independent risk factors to construct a combined model (radiomics nomogram) based on multivariate Cox regression analysis, and the C-index of the combined model (radiomics nomogram) was 0.840 and 0.803 in the training and validation cohorts, respectively (**Table 2**).

Development of the Radiomics Nomogram and Its Performance

The rad score was combined with the independent risk factors to construct a radiomics nomogram based on multivariate Cox regression analysis (**Figure 4A**). The C-index of the radiomics nomogram was 0.840 and 0.803 in the training and validation cohorts, respectively. The calibration curve result showed that the predicted probability was remarkably close to the actual survival time of patients (**Figures 4B, C**). Kaplan–Meier survival analysis of patients in the high-risk and low-risk groups in the training cohort (log-rank test $p = 0.001$; **Figure 4D**). A decision curve analysis showed that the radiomics nomogram had a higher overall net benefit than the clinical model and the radiomics model, and had a higher overall net benefit across the majority of the range of reasonable threshold probabilities (**Figure 4E**).

To Guide the Targeted Treatment for Lung Adenocarcinoma in Patients With EGFR Mutations

According to the cut-off value of nomogram score at 0.369, the corresponding 5-year overall survival probability was 0.58. Patients with an *EGFR*-mutation were divided into high- and low-risk groups, and the sensitivity of high- and low-risk patients to chemotherapy and targeted therapy was analyzed. The results showed that high-risk patients had a higher sensitivity to targeted

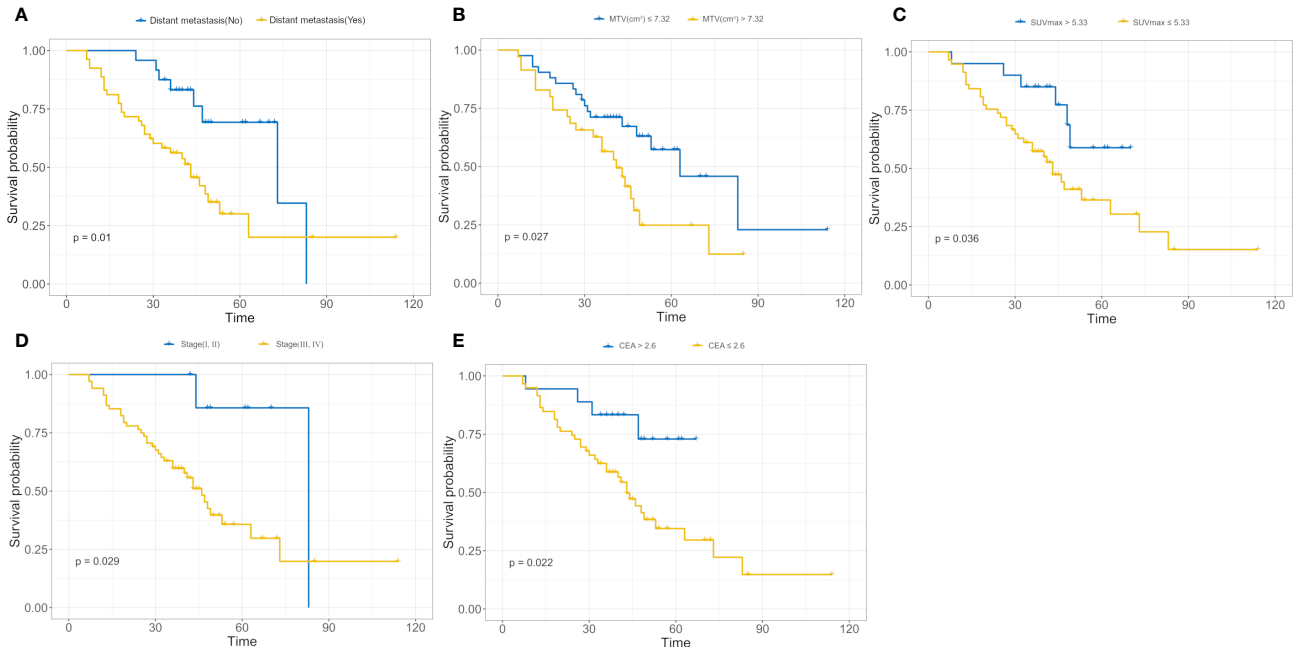


FIGURE 3 | **(A–E)** Kaplan–Meier analysis for distant metastasis **(A)**, metabolic tumor volume (MTV) **(B)**, maximal standard uptake value ($SUVmax$); **(C)**, stage **(D)**, carcinoembryonic antigen (CEA); **(E)**. The patients were stratified into high- and low-risk groups based on distant metastasis (**A**, $p = 0.01$, log-rank test), MTV (**B**, $p = 0.027$, log-rank test), SUVmax (**C**, $p = 0.036$, log-rank test), stage (**D**, $p = 0.029$, log-rank test), and CEA (**E**, $p = 0.022$, log-rank test).

TABLE 2 | The comparison of prognostic accuracy between the radiomics model and two other prognostic models.

models	Training cohort	Validation cohort
	C-index 95% CI	C-index 95% CI
Radiomics model	0.819(0.764–0.874)	0.737(0.606–0.868)
Clinical model	0.694(0.618–0.770)	0.729(0.599–0.858)
Radiomics nomogram	0.840(0.787–0.893)	0.803(0.689–0.917)

therapy ($p < 0.0001$), indicating that targeted therapy is the main treatment method for patients with high-risk *EGFR* mutations, while the prognoses of the two therapies were similar in the low-risk group ($p = 0.85$, **Figures 5A, B**). In patients with an 19DEL mutation, there was no significant difference in the sensitivity to chemotherapy and targeted therapy ($p = 0.45$). The patients with a 21L858R-mutation had significant differences in sensitivity to chemotherapy and targeted therapy, and the patients with a 21L858R-mutation were more likely to benefit from targeted therapy ($p = 0.042$; **Figures 5C, D**). In addition, there was no significant difference between patients with a 19DEL-mutation and patients with a 21L858R-mutation in their benefit from chemotherapy ($p = 0.29$; **Figure 5E**).

DISCUSSION

In our study, we developed a radiomics nomogram based on ^{18}F-FDG PET/CT radiomics features combined with clinicopathological factors to predict survival outcomes in patients with lung adenocarcinoma of EGFR mutations, with the aim of providing guidance for personalized targeted treatment of patients with lung adenocarcinoma with EGFR mutations.

In the CPH model for evaluating the prognosis of patients with *EGFR* mutations, distant metastasis, MTV, stage, CEA, and SUVmax were the significant prognostic risk factors. Among them, the patient's risk of death was higher when the patient had MTV (>7.32). MTV is a parameter that reflected the metabolic burden of the whole-body tumor compared with other PET/CT semiquantitative parameters and related clinicopathological factors. It can more effectively stratify the risk of patients and identify high-risk groups. In particular, it can effectively evaluate the prognosis of patients with advanced lung cancer. This was consistent with our findings (34, 35). SUVmax is the most used metabolic parameter of PET/CT in clinical work and only represents a single pixel value of the tumor metabolism that is most active in the outlined area. Whether SUVmax is an independent risk factor for lung cancer remains controversial (22). Some studies believe that SUVmax can effectively indicate the degree of tumor differentiation and provide evidence for the prognosis of patients (36). Our study demonstrated that when SUVmax (>5.33), the patient's risk of death increased. This was consistent with our findings.

In addition, we combined the rad score with independent risk factors (Distant metastasis and MTV) based on multivariate Cox regression analysis to construct a radiomics nomogram that predicted survival outcomes of patients with EGFR mutations. The results showed that a radiomics nomogram can predict

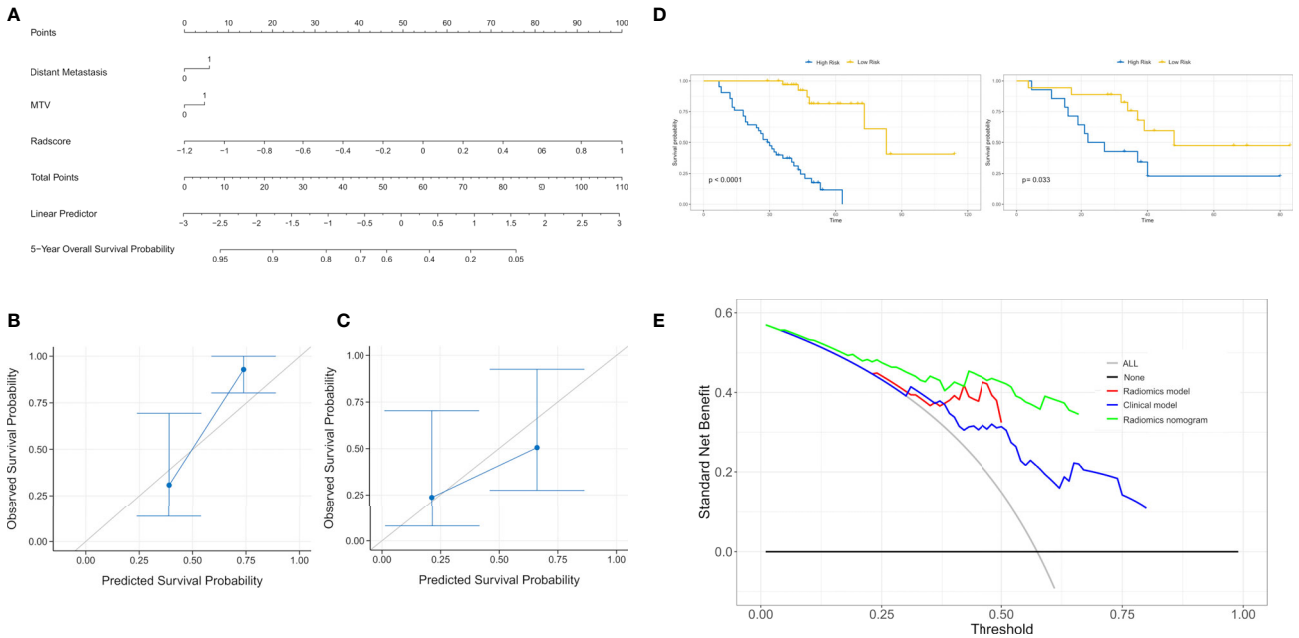


FIGURE 4 | **(A)** A radiomics nomogram for prediction of 5-year overall survival for patients with lung adenocarcinoma of EGFR mutations. **(B)** Calibration curve of the radiomics nomogram in the training cohort. **(C)** Calibration curve of the radiomics nomogram in the validation cohort. Calibration curve for the estimation of 5-year overall survival as predicted by the nomogram. The nomogram-estimated overall survival is plotted on the x-axis, and the actual overall survival is plotted on the y-axis. Dash line represents an ideal agreement. **(D)** The Kaplan–Meier curve showed that this nomogram score could effectively discriminate high-risk patients from low-risk patients. **(E)** The decision curve analysis for each model. The y-axis denotes the net benefit, which was calculated using true-positive and false-positive results. The radiomics nomogram model has the highest net benefit at the threshold from 0.1 to 0.9 among all positive predictions (line labeled "All"); all negative predictions (line labeled "None") and two other clinical models (line labeled "Radiomics model and clinical model").

survival outcomes very well. Its' C-index was 0.840 and 0.803 in the training and validation cohorts, respectively, which could stratify high- and low-risk groups quite well. At present, few studies based on PET/CT radiomics have predicted the survival of lung cancer patients with *EGFR* mutations, and their predictive performances were generally poor (37, 38). Kirienko et al. (28) used radiomics signatures based on PET/CT to predict disease-free survival (DFS) of patients with NSCLC after surgery. The results showed that the AUC of the Cox model based on the radiomics signature was 0.68, and the AUC was 0.65 after combining it with clinical predictors. Moreover, the current study focused mainly on a CT modality while predicting survival, and the value of the C-index was usually not well *i.e.*, did not exceed 0.70. The performance improved after combining it with clinicopathological factors (39, 40). Our results showed that the C-index reached 0.803, and our result was a small breakthrough in the results of previous studies. To guide the treatment of patients with *EGFR* mutations, our study analyzed the effects of different treatment strategies on the prognosis of patients with *EGFR* mutations. Our results showed that the rad score could stratify patients with *EGFR* mutations into high- and low-risk groups. For patients who were at high risk, targeted therapy is recommended to improve survival. For patients at low risk, there was no significant difference in survival regardless of whether targeted therapy or chemotherapy was chosen. The patients with a 21L858R-mutation had significant differences in sensitivity to chemotherapy and targeted therapy, and the patients with a 21L858R-mutation were more likely to benefit from targeted therapy. However, in patients with a 19DEL mutation, there was no significant difference in the sensitivity to chemotherapy and targeted therapy. It may be due to the small sample size and the bias caused by retrospective study. In addition, there was no significant difference between patients with a 19DEL mutation and patients with a 21L858R-mutation in their benefit from chemotherapy. It illustrated that patients with EGFR mutations may not benefit from chemotherapy. Our results indicated that radiomics features could identify patients who are more likely to benefit from targeted therapy among patients with EGFR mutations, and would benefit from treatment guidance.

Our study had many strengths. First, our study not only predicted survival outcomes in lung adenocarcinoma patients with *EGFR* mutations, but also identified patients with EGFR mutations who were likely to benefit from targeted therapy through rad score. We provided guidance for the selection of treatment methods in patients with *EGFR* mutations, which was rarely reported in previous studies. Second, patients in this study were scanned using the same PET/CT device used in a standard protocol, which avoided the heterogeneity of image impressions caused by the use of different scans and reconstruction parameters. This led to more stable and reliable results.

Our study had some limitations. First, this was a retrospective study with a small data set and no external validation, which may

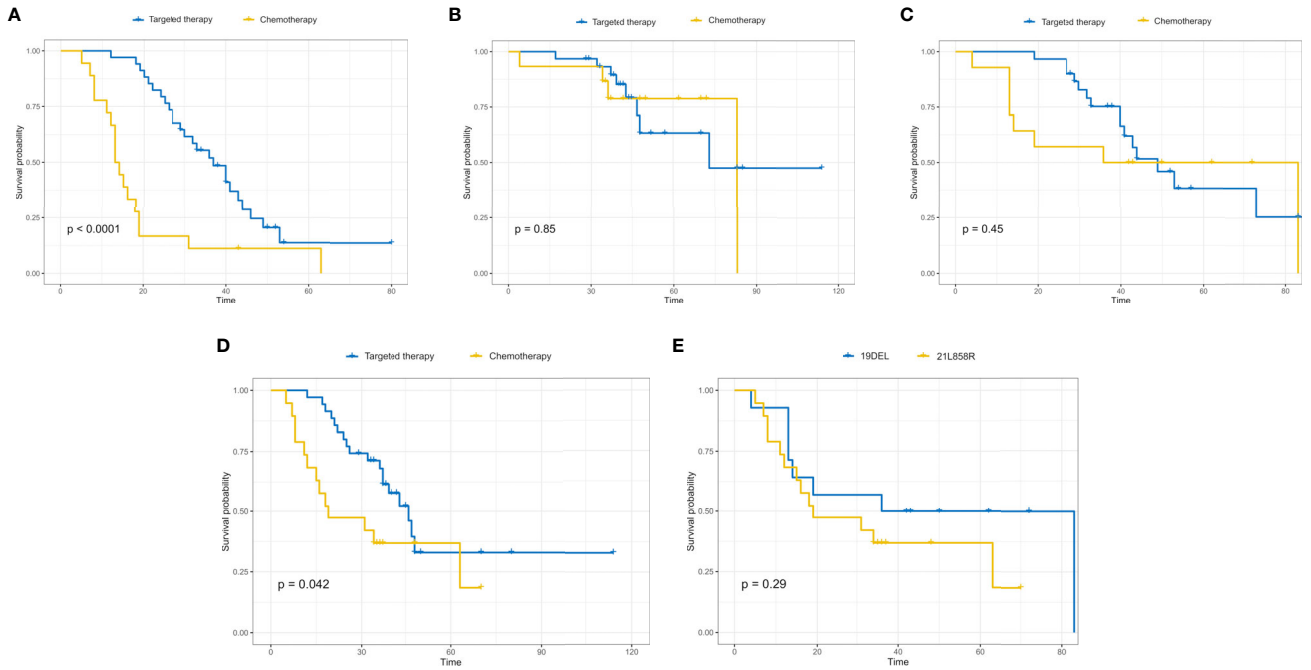


FIGURE 5 | High-risk patients with *EGFR* mutations had a higher sensitivity to targeted therapy ($p < 0.0001$) **(A)**, while the sensitivity of low-risk patients to targeted therapy and chemotherapy was not significantly different ($p = 0.85$) **(B)**. In high-risk patients, the sensitivity of patients with the 19DEL mutation to chemotherapy and targeted therapy was not significantly different ($p = 0.45$) **(C)**, while the sensitivity of patients with the 21L858R-mutation to chemotherapy and targeted therapy was significantly different ($p = 0.042$) **(D)**. **(E)** The sensitivity of patients with the 19DEL- and 21L858R-mutations to chemotherapy was not significantly different ($p = 0.29$).

have introduced selection bias. Second, we only studied the effect of treatment on the prognosis of patients with lung adenocarcinoma and an *EGFR* mutation status and did not consider the influence of other genes. Further studies are essential to evaluate other genes comprehensively.

In conclusion, a ^{18}F-FDG PET/CT rad score combined with clinicopathological factors can predict the survival outcomes of patients with lung adenocarcinoma with an *EGFR* mutation. This novel and non-invasive approach can be provide with a more precise imaging diagnosis and personalized treatment guidance for patients with an EGFR mutant and have a significant clinical application value.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The institutional review board of Affiliated Jinling Hospital, Medical School of Nanjing University approved this retrospective study and waived the need to obtain informed consent from the patients.

AUTHOR CONTRIBUTIONS

BY conceived the idea of the study. BY, HJ, JinZ, LM, JiaZ, HD, and CZ collected the data. HZ and GL performed image analysis. BY wrote the manuscript. SD performed the statistical analysis. CZ, JT, LZ, FW, and GL edited and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2017YFC0113400 for LZ) and Natural Science Foundation of Jiangsu Province (BK2011665).

ACKNOWLEDGMENTS

We would like to thank all the personnel from the Department of Nuclear Medicine and Department of Medical Imaging, Affiliated Jinling Hospital, Medical school of Nanjing University for their collective efforts on the image post-processing, clinical management, and follow-up.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.567160/full#supplementary-material

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* (2015) 65:87–108. doi: 10.3322/caac.21262
2. Siegel RL, Miller KD. Cancer Statistics, 2017. *CA Cancer J Clin* (2017) 67:7–30. doi: 10.3322/caac.21387
3. Brawley OW. Avoidable cancer deaths globally. *CA Cancer J Clin* (2011) 61:67–8. doi: 10.3322/caac.20108
4. Oser MG, Niederst MJ, Sequist LV. Transformation from non-small cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin. *Lancet Oncol* (2015) 16:e165–172. doi: 10.1016/S1470-2045(14)71180-5
5. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc* (2008) 83:584–94. doi: 10.4065/83.5.584
6. Curran WJ Jr, Paulus R, Langer CJ, Komaki R, Lee JS, Hauser S, et al. Sequential vs. concurrent chemoradiation for stage III non-small cell lung cancer: randomized phase III trial RTOG 9410. *J Natl Cancer Inst* (2011) 103:1452–60. doi: 10.1093/jnci/djr325
7. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* (2015) 136:E359–86. doi: 10.1002/ijc.29210
8. Siegel RL, Miller KD. Cancer statistics, 2018. *CA Cancer J Clin* (2018) 68:7–30. doi: 10.3322/caac.21442
9. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* (2016) 11:39–51. doi: 10.1016/j.jtho.2015.09.009
10. Sculier JP, Chansky K, Crowley JJ, Van Meerbeeck J, Goldstraw P. The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th Edition of the TNM Classification of Malignant Tumors and the proposals for the 7th Edition. *J Thorac Oncol* (2008) 3:457–66. doi: 10.1097/JTO.0b013e31816de2b8
11. Chansky K, Sculier JP, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P. The International Association for the Study of Lung Cancer Staging Project: prognostic factors and pathologic TNM stage in surgically managed non-small cell lung cancer. *J Thorac Oncol* (2009) 4:792–801. doi: 10.1097/JTO.0b013e3181a7716e
12. Kawaguchi T, Takada M, Kubo A, Matsumura A, Fukai S, Tamura A, et al. Performance status and smoking status are independent favorable prognostic factors for survival in non-small cell lung cancer: a comprehensive analysis of 26,957 patients with NSCLC. *J Thorac Oncol* (2010) 5:620–30. doi: 10.1097/JTO.0b013e3181d2dcd9
13. Higashi K, Ueda Y, Arisaka Y, Sakuma T, Nambu Y, Oguchi M, et al. 18F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resected non-small cell lung cancer. *J Nucl Med* (2002) 43:39–45. doi: 10.1587/transcom.E96.B.802
14. Liang W, Zhang L, Jiang G, Wang Q, Liu L, Liu D, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol* (2015) 33:861–9. doi: 10.1200/JCO.2014.56.6661
15. Wang L, Dong T, Xin B, Xu C, Guo M, Zhang H, et al. Integrative nomogram of CT imaging, clinical, and hematological features for survival prediction of patients with locally advanced non-small cell lung cancer. *Eur Radiol* (2019) 29:2958–67. doi: 10.1007/s00330-018-5949-2
16. Valentini V, van Stiphout RG, Lammering G, Gambacorta MA, Barba MC, Bebenek M, et al. Nomograms for predicting local recurrence, distant metastases, and overall survival for patients with locally advanced rectal cancer on the basis of European randomized clinical trials. *J Clin Oncol* (2011) 29:3163–72. doi: 10.1200/JCO.2010.33.1595
17. Han DS, Suh YS, Kong SH, Lee HJ, Choi Y, Aikou S, et al. Nomogram predicting long-term survival after d2 gastrectomy for gastric cancer. *J Clin Oncol* (2012) 30:3834–40. doi: 10.1200/JCO.2012.41.8343
18. Chun FK, Briganti A, Graefen M, Montorsi F, Porter C, Scattoni V, et al. Development and external validation of an extended 10-core biopsy nomogram. *Eur Urol* (2007) 52:436–44. doi: 10.1016/j.eururo.2006.08.039
19. Goodgame B, Pillot GA, Yang Z, Shriki J, Meyers BF, Zoole J, et al. Prognostic value of preoperative positron emission tomography in resected stage I non-small cell lung cancer. *J Thorac Oncol* (2008) 3:130–4. doi: 10.1097/JTO.0b013e318160c122
20. Kim HR, Kim DJ, Lee WW, Jheon S, Sung SW. The significance of maximum standardized uptake values in patients with stage I pulmonary adenocarcinoma. *Eur J Cardiothorac Surg* (2009) 35:712–6. doi: 10.1016/j.ejcts.2008.12.030
21. Hoang JK, Hoagland LF, Coleman RE, Coan AD, Herndon JE 2nd, Patz EF Jr. Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography imaging in patients with advanced-stage non-small-cell lung carcinoma. *J Clin Oncol* (2008) 26:1459–64. doi: 10.1200/JCO.2007.14.3628
22. Inal A, Kucukoner M, Kaplan MA, Urakci Z, Karakus A, Komek H, et al. Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography in patients with advanced non-small cell lung cancer: single center experience. *J BUON* (2012) 17:724–8. doi: 10.1097/MD.0000000000009877
23. Hatt M, Majdoub M, Vallières M, Tixier F, Le Rest CC, Groheux D, et al. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med* (2015) 56:38–44. doi: 10.2967/jnumed.114.144055
24. Kirienko M, Gallivanone F, Sollini M, Veronesi G, Voulaz E, Antunovic L, et al. FDG PET/CT as theranostic imaging in diagnosis of non-small cell lung cancer. *Front Biosci (Landmark Ed)* (2017) 22:1713–23. doi: 10.2741/4567
25. Desseroit MC, Visvikis D, Tixier F, Majdoub M, Perdrisot R, Guillevin R, et al. Development of a nomogram combining clinical staging with (18)F-FDG PET/CT image features in non-small-cell lung cancer stage I-III. *Eur J Nucl Med Mol Imaging* (2016) 43:1477–85. doi: 10.1007/s00259-016-3325-5
26. Grootjans W, de Geus-Oei LF, Troost EG, Visser EP, Oyen WJ, Bussink J, et al. PET in the management of locally advanced and metastatic NSCLC. *Nat Rev Clin Oncol* (2015) 12:395–407. doi: 10.1038/nrclinonc.2015.75
27. Dou TH, Coroller TP, van Griethuysen JJM, Mak RH, Aerts HJWL. Peritumoral radiomics features predict distant metastasis in locally advanced NSCLC. *PLoS One* (2018) 13:e0206108. doi: 10.1371/journal.pone.0206108
28. Kirienko M, Cozzi L, Antunovic L, Lozza L, Fogliata A, Voulaz E, et al. Prediction of disease-free survival by the PET/CT radiomic signature in non-small cell lung cancer patients undergoing surgery. *Eur J Nucl Med Mol Imaging* (2018) 45:207–17. doi: 10.1007/s00259-017-3837-7
29. Wels MG, Lades F, Muehlberg A, et al. General purpose radiomics for multi-modal clinical research. In: *Proceedings of the SPIE 10950, medical imaging 2019: computer-aided diagnosis,10950-150* (2019). doi: 10.1117/12.2511856
30. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics or High-Throughput Image-based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
31. Sauerbrei W, Royston P. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* (2007) 26:5512–28. doi: 10.1002/sim.3148
32. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol* (2016) 34:2157–64. doi: 10.1200/JCO.2015.65.9128
33. Yang L, Yang J, Zhou X, Huang L, Zhao W, Wang T, et al. Development of a radiomics nomogram based on the 2D and 3D CT features to predict the survival of non-small cell lung cancer patients. *Eur Radiol* (2019) 29:2196–206. doi: 10.1007/s00330-018-5770-y
34. Liao S, Penney BC, Zhang H, Suzuki K, Pu Y. Prognostic value of the quantitative metabolic volumetric measurement on ^{18}F-FDG PET/CT in Stage IV nonsurgical small-cell lung cancer. *Acad Radiol* (2012) 19:69–77. doi: 10.1016/j.acra.2011.08.020
35. Hyun SH, Ahn HK, Kim H, Ahn MJ, Park K, Ahn YC, et al. Volume-based assessment by (18)F-FDG PET/CT predicts survival in patients with stage III

non-small-cell lung cancer. *Eur J Nucl Med Mol Imaging* (2014) 41:50–8. doi: 10.1007/s00259-013-2530-8
36. Duan XY, Wang W, Li M, Li Y, Guo YM. Predictive significance of standardized uptake value parameters of FDG-PET in patients with non-small cell lung carcinoma. *Braz J Med Biol Res* (2015) 48:267–72. doi: 10.1590/1414-431x20144137
37. Kim DH, Jung JH, Son SH, Kim CY, Hong CM, Oh JR, et al. Prognostic Significance of Intratumoral Metabolic Heterogeneity on 18F-FDG PET/CT in Pathological N0 Non-Small Cell Lung Cancer. *Clin Nucl Med* (2015) 40:708–14. doi: 10.1097/RLU.0000000000000867
38. Tixier F, Hatt M, Valla C, Fleury V, Lamour C, Ezzouhri S, et al. Visual versus quantitative assessment of intratumor 18F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med* (2014) 55:1235–41. doi: 10.2967/jnumed.113.133389
39. Lee J, Li B, Cui Y, Sun X, Wu J, Zhu H, et al. A Quantitative CT Imaging Signature Predicts Survival and Complements Established Prognosticators in Stage I Non-Small Cell Lung Cancer. *Int J Radiat Oncol Biol Phys* (2018) 102:1098–106. doi: 10.1016/j.ijrobp.2018.01.006
40. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* (2016) 281:947–57. doi: 10.1148/radiol.2016152234

Conflict of Interest: S-fD was employed by GE Healthcare China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

OPINION
published: 18 November 2020
doi: 10.3389/fonc.2020.556334

Optimization of Patient Management During the COVID-19 Pandemic: Chest CT Scan and PCR as Gatekeepers of the Radiation Therapy Workflow

Roger Sun[1], Samy Ammari[2], Sophie Bockel[1], Samir Achkar[1], Mansouria Merad[3], Laurent Dercle[4], Sofia Rivera[1], Cyrus Chargari[1] and Eric Deutsch[1]**

[1] Gustave Roussy, Département de Radiothérapie, INSERM 1030, Université Paris-Saclay, Villejuif, France, [2] Gustave Roussy, Département d'Imagerie Médicale, Université Paris-Saclay, Villejuif, France, [3] Gustave Roussy, Département d'Oncologie Médicale, Université Paris-Saclay, Villejuif, France, [4] Department of Radiology, Columbia University Irving Medical Center, New York, NY, United States

OPEN ACCESS

Edited by:
Randy Yeh,
Memorial Sloan Kettering Cancer Center, United States

Reviewed by:
Ilja Ciernik,
Städtische Klinikum Dessau, Germany

***Correspondence:**
Roger Sun
roger.sun@gustaveroussy.fr
Eric Deutsch
eric.deutsch@gustaveroussy.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 27 April 2020
Accepted: 23 October 2020
Published: 18 November 2020

Citation:
Sun R, Ammari S, Bockel S, Achkar S, Merad M, Dercle L, Rivera S, Chargari C and Deutsch E (2020) Optimization of Patient Management During the COVID-19 Pandemic: Chest CT Scan and PCR as Gatekeepers of the Radiation Therapy Workflow. Front. Oncol. 10:556334. doi: 10.3389/fonc.2020.556334

Keywords: coronavirus disease, radiotherapy, chest CT, workflow, PCR

INTRODUCTION

The coronavirus disease 2019 (COVID-19) has been rapidly spreading since the first patients were described in Wuhan, China in late December 2019 (1). As of mid-July 2020, nearly 15 million of confirmed cases and 600,000 deaths have been reported worldwide (2). This global pandemic and international health crisis has challenged healthcare providers to profoundly re-organize healthcare systems in order to ensure the continuity of essential treatments, while limiting the risk to patients and healthcare providers, and simultaneously handling shortages in personnel, beds, and equipment (3).

Increasing evidence suggests that cancer patients and especially those undergoing treatment might be at higher risk of developing severe forms of COVID-19 (4–7). Indeed, anticancer treatment such as chemotherapy and radiotherapy that induce neutropenia and/or lymphopenia, as well as targeted therapy or immune-check point inhibitors, may worsen the course of COVID-19, although this is still debated (3, 8) For this reason, cancer societies have recommended to adapt the management of cancer patients by de-escalating cytotoxic chemotherapy, delaying non-urgent treatments, and considering non-surgical options when feasible (9, 10). In radiation oncology, some indications remain non-deferrable, such as chemoradiotherapy for locally advanced tumors, brachytherapy, and urgent palliative treatment, particularly in patients with rapidly growing tumors (11). In addition, radiotherapy capacity may even need to be increased as an alternative treatment to compensate for surgery cancellations. While hypofractionated radiotherapy schedules may help to limit the number of hospital visits, daily sessions for at least 2 weeks remain the rule for most curative treatments, with some plans lasting up to 7–8 weeks (12, 13). Moreover, strategy relying only on patient selection and prioritizing treatment management alone may not be sufficient in a prolonged healthcare crises (10). Active and effective strategies of early detection of COVID-19

should therefore be discussed and tailored for radiation oncology departments, especially since a significant proportion of patients are asymptomatic (14).

SARS-COV-2 DETECTION AND COVID-19 DIAGNOSIS

COVID-19 is an infectious respiratory disease caused by a novel coronavirus called SARS-CoV-2 (15). The accepted standard routine laboratory method to diagnose COVID-19 consists of the use of reverse transcription polymerase chain reaction (rt-PCR) to detect viral RNA in respiratory samples. However, performance of nasopharyngeal and oropharyngeal swabs is questioned since several studies have shown a high number of false negative results (16, 17). In a cohort of 213 patients with mild to severe symptoms of COVID-19, Yang et al. reported positive rates ranging from 73.3% for nasal swabs performed less than 7 days after illness onset on patients with severe symptoms and 53.6% for nasal swabs performed more than 8 days after illness onset on patients with mild symptoms. Wang et al. reported low positive rates of 32% for pharyngeal swabs (126/398 patients) and 63% for nasal swabs (5/8 patients). Moreover, rt-PCR has other limitations including the availability of testing kits, the need for adapted infrastructure, training of personnel for high quality swab samples, and relative lengthy turnaround times for test results.

On the other hand, some authors have suggested computed tomography (CT) could be pivotal for the diagnosis and the screening of COVID-19 (18, 19). In a cohort of 1,014 patients, Ai et al. reported a positive rate of 59% for rt-PCR (601/1014) and 88% for chest CT (888/1014), with 97% of positive rt-PCR patients having also a positive chest CT (580/601) (18). The higher sensitivity of chest CT compared to rt-PCR was confirmed in a recent meta-analysis [94% (5% CI: 91%, 96%) vs. 89% (95% CI: 81%, 94%), respectively] (20). However, this study highlighted the low specificity of chest CT [37% (95% CI: 26%, 50%)] with poor positive predictive value (PPV) for populations with low expected prevalence of COVID-19, limiting its routine use for mass screening and diagnosis in the overall population, especially since CT scans also expose people to ionizing radiation. Therefore, the general consensus is to not recommend chest CT for diagnosing COVID-19 and to reserve CT for patients with worsening symptoms (21, 22). However, several authors have suggested that imaging could help patient triage in a resource-constrained environment, especially when the pre-test probability of COVID-19 is high (21, 22). Nevertheless, with a high negative predictive value (NPV) estimated between 90.6% and 99.8%, chest CT imaging could still be very valuable in specific situations, such as in screening patients undergoing radiotherapy.

Recently, serological tests have become more widespread (23). Detection of IgM antibodies associated with symptoms is highly suggestive of SARS-CoV-2 infection. IgM antibodies may be detectable around 5 to 10 days after the onset of symptoms, with a mean time for seroconversion of 10 to 12 days for IgM, and 12 to 14 days for IgG (23–26). However, whether an antibody response with neutralizing antibodies is associated with protective immunity is still unclear, and higher IgG concentration have been reported in patients with severe COVID-19 infection (27). While these tests were not recommended for systematic screening by the World Health Organization (28), they were recommended by the French Haute Autorité de Santé in patients with negative RT-PCR and clinical suspicion for COVID-19 (29).

OPTIMIZING MEDICAL IMAGING IN RADIOTHERAPY DURING THE COVID-19 PANDEMIC

Feasibility and Benefit of Chest CT in Radiation Therapy

Radiation therapy requires each patient to undergo a simulation CT scan for treatment planning. This CT allows the visualization and delineation of target tumor volumes to irradiate and normal tissues to spare, the calculation and the optimization of the radiation dose, and the reproducibility of patient positioning (30). An additional chest CT scan or extending the scan coverage of a simulation CT could therefore be easily performed for each patient without modifying patient's visit or the radiotherapy workflow. Moreover, the expected additional delivered radiation dose to the chest is low, around 12 mGy [interquartile range (IQR), 7–17 mGy] (31, 32), and may be considered as negligible for patients undergoing radiotherapy. Indeed, even for non-thoracic treatment, the estimated dose at 30 cm from the irradiated field is estimated to be 0.05%–0.7% of the delivered dose (33). For example, the estimated dose received to the chest for a pelvic treatment of 45 Gy would correspond to 2–26 times the dose of a chest CT. In addition, free-breathing chest CTs are routinely performed during the CT simulation for patients with lung and breast cancers, metastases to the thoracic and lumbar spine, and sometimes for head and neck cancers. In certain indications, such as in breast cancer, deep inspiration breath hold radiotherapy technique allows acquisition of breath-hold chest CT.

For patients with other cancers, the additional chest CT could also serve as a baseline for reference, in addition to screening for COVID-19. Indeed, cancer patients may have progressive cancer- or cancer-treatment related lung abnormalities, which may be complicated to interpret if the patient becomes symptomatic. Moreover, the rate of incidental findings in cancer patients may be relatively high: a study of 510 patients reported incidental findings in up to 28% of patients, including 3.4% of patients with a significant finding that either changed the cancer therapy or required immediate treatment (34).

Early Detection of COVID-19 in Patients Undergoing Radiotherapy

The lack of large epidemiological studies and the difficulty of estimating the true incidence of COVID-19 cases make it

challenging to determine whether cancer patients are more or less vulnerable to SARS-CoV-2 in comparison with the overall population. Indeed, among COVID-19 patients, an important proportion of patients are asymptomatic, but may still be contagious (35–37). A voluntary screening on 10,797 patients in Iceland, with 87 COVID-19 positive patients (0.8%) of whom 57% were asymptomatic (14). Moreover, with an incubation period between 3 to 6 days (15, 38), a study estimated that up to 44% [95% CI (25–69)] of secondary cases may have been infected by pre-symptomatic cases (36). Cancer patients undergoing treatment may have compromised immunity, potentially increasing their morbidity and mortality from COVID-19. Miyashita et al. reported 5,688 COVID-19 patients, including 334 patients with cancer. Intubation was more frequent in cancer patients [RR: 1.89 95% CI (1.37–2.61)], but there was no significant excess risk of death (5). Kuderer et al. reported in a study of 928 patients with history of cancer that there was an increased 30-day mortality in patients with active cancer [progressing vs. remission: Odds ratio = 5.20, 95% CI (2.77–9.77)] (39).

Radiotherapy treatment requires frequent hospital visits for the patients, a flow of multiple patients per day on the same machine, and involvement of an entire team of healthcare professionals including a medical dosimetrist, medical physicist, radiation oncologist, radiation therapists, oncology nurses, caregivers, and medical secretaries. Strategies for early detection of COVID-19 are therefore of utmost importance to ensure patient and medical staff safety during the COVID-19 pandemic, especially in regions with high prevalence and documented community spread.

Proposal of COVID-19 Testing for Asymptomatic Patients in Radiation Oncology Departments

A workflow including systematic nasopharyngeal swab and chest CT for asymptomatic patients is currently being evaluated in our Radiotherapy department, as the prevalence of COVID-19 was high in our region, with up to 11.9% of individuals infected (range: 7.6% to 19.4%) according to epidemiologic models (40). Inspiratory breath hold chest acquisition is made during the simulation CT, with the same CT acquisition protocols conventionally used in radiology. While free-breathing acquisitions are usually done in radiotherapy, they are more susceptible to motion artifacts and may obscure or even mimic subtle ground-glass opacities commonly seen in COVID-19, and regular cone-beam CT (CBCT) positioning review may require comparison with reference chest CT. Images are then reviewed by both the treating radiation oncologist and an experienced radiologist for imaging findings suggestive of COVID-19 such as ground-glass opacity, condensation, reticulation, interlobular or intralobular septal thickening, nodules, and distribution of the lesions (41). For suspected COVID-19 patients on either radiological findings or clinical symptoms but negative RT-PCR, directed nasopharyngeal swab and inflammatory blood test are repeated before the onset of radiotherapy. This procedure allows the classification of patients into three different categories with specific management according to the COVID-19 probability (**Figure 1**):

1) Confirmed or probable cases: We propose using a dedicated treatment room with adapted personal protective equipment {protective masks against inhalation of droplets, such as FFP2 (filtering facepiece type 2) masks [standard NF EN 149 (42)], gown, gloves, eye protection, and apron} and a dedicated accelerator for the irradiation of suspected and/or confirmed COVID-19 patients, which may help to avoid deferring treatment. On a case by case basis taking into account symptoms, deferring radiotherapy after the acute phase of infection is discussed. However, this approach using a dedicated room for confirmed or probable cases is only feasible in large centers with a high number of accelerators.
2) Uncertain cases: We propose in these patients close monitoring of clinical symptoms and repeating nasopharyngeal swabs. Further investigation might be proposed according to the multidisciplinary team, such as CT follow-up or other microbiology investigation and specimens. These patients are preferably scheduled for treatment at the end of the day.
3) Unlikely cases: We propose using standard personal protective equipment and routine radiotherapy treatment protocol.

In addition to CT scan and rt-PCR screening, we have undertaken a global reorganization of our radiation oncology department with implementation of barrier precautions and social distancing. Prior to entry, patients are asked to answer a short checklist to screen for clinical symptoms, such as fevers, chills, cough, etc., and reception staff measure the temperature of patients at the entrance. All medical staff and patients are required to wear surgical masks in the department. The flow of patients in the radiation oncology department has also been rethought. The simulation CT is dedicated to COVID-19 patients only on Friday, allowing a two-day interval before using the scanner again for patients without COVID-19. Whenever possible, we encourage COVID-19 patients to wait for their session in their car or in the taxi in order to limit the time spent in the waiting room, and therefore, limiting the risk of SARS-CoV-2 transmission.

These screening strategies have the most potential and maximum benefit during the peak of COVID-19 cases and adjustments need to be made according to the evolution of the pandemic and the pre-test probability of COVID-19 in a particular area, as low prevalence of COVID-19 leads to a low PPV, especially for chest CT (20). For instance, while systematic baseline RT-PCR testing is recommended in Canada by the Ontario provincial Ministry of Health guidelines for cancer patients undergoing immunosuppressive cancer treatment, including radiation therapy, high priority testing criteria for asymptomatic patients in the event of testing limitations were defined, included age ≥ 60 years, performance status ≥ 2, comorbid conditions or impaired immunity, significant smoking history and lung tissue in the radiation treatment volume (43).

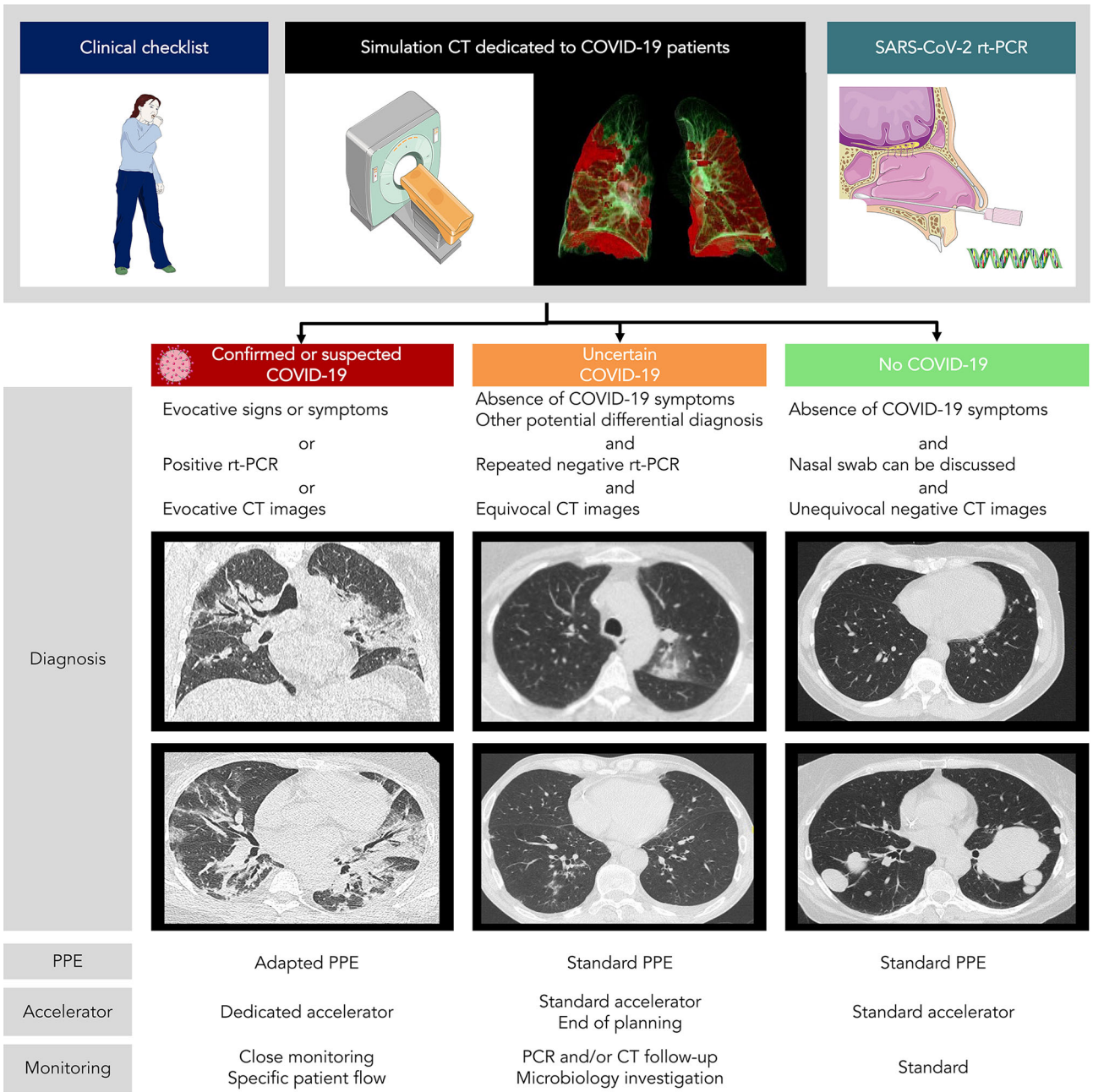


FIGURE 1 | COVID-19 patients management in radiotherapy according to chest CT + rt-PCR screening. PPE, personal protective equipment.

CONCLUSION

The COVID-19 pandemic has significantly impacted the delivery of care to cancer patients, leading to delays in providing adequate treatment. While postponing treatments and providing remote consultations were feasible in the early stages of the pandemic, these temporary measures are no longer sustainable as the pandemic continues. Hospitals and medical services need to develop and adopt long-term strategies to continue providing cancer care during the pandemic. In patients undergoing radiotherapy, we propose using chest CT and rt-PCR screening for early detection of COVID-19, especially since a number of patients are asymptomatic and cancer patients might be more vulnerable than the overall population. Rethinking the flow of patients is also critical to allow the continuation of care, with implementation of barrier precautions and social distancing.

AUTHOR CONTRIBUTIONS

RS and ED prepared the first draft of the manuscript. LD conceived the figure. SB, SR, SAm, MM, SAc, LD, and CC reviewed the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank Gianfranco Brusadin, Simon Corbin, Frederik Hubert, Virgine Chiesa, Najma Douir, Gaëlle Mevel, and all the radiation therapists of the department for their help and the care they provide to patients.

REFERENCES

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* (2020) 382 (8):727–33. doi: 10.1056/NEJMoa2001017
2. World Health Organization. *Novel Coronavirus (2019-nCoV) situation reports*. Situation Report (2020). p. 183. Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports Last Accessed 21 July 2020.
3. van de Haar J, Hoes LR, Coles CE, Seamon K, Fröhling S, Jäger D, et al. Caring for patients with cancer in the COVID-19 era. *Nat Med* (2020) 26:1–7. 16 avr. doi: 10.1038/s41591-020-0948-7
4. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol* (2020) 21(3):335–7. doi: 10.1016/S1470-2045(20)30096-6
5. Miyashita H, Mikami T, Chopra N, Yamada T, Chernyavsky S, Rizk D, et al. Do Patients with Cancer Have a Poorer Prognosis of COVID-19? An Experience in New York City. *Ann Oncol* (2020) 31(8):1088–9. doi: 10.1016/j.annonc.2020.04.006
6. Tian Y, Qiu X, Wang C, Zhao J, Jiang X, Niu W, et al. Cancer associates with risk and severe events of COVID-19: A systematic review and meta-analysis. *Int J Cancer* (2020). doi: 10.1002/ijc.33213
7. Li Q, Chen L, Li Q, He W, Yu J, Chen L, et al. Cancer increases risk of in-hospital death from COVID-19 in persons <65 years and those not in complete remission. *Leukemia* (2020) 34(9):2384–91. doi: 10.1038/s41375-020-0986-7
8. Lee LYW, Cazier JB, Starkey T, Turnbull CDUK Coronavirus Cancer Monitoring Project Team, , Kerr R, et al. COVID-19 mortality in patients with cancer on chemotherapy or other anticancer treatments: a prospective cohort study. *Lancet Lond Engl* (2020) 395(10241):1919–26. doi: 10.1016/S0140-6736(20)31173-9
9. Burki TK. Cancer guidelines during the COVID-19 pandemic. *Lancet Oncol* (2020) 21(5):629–630. doi: 10.1016/S1470-2045(20)30217-5
10. Hanna TP, Evans GA, Booth CM. Cancer, COVID-19 and the precautionary principle: prioritizing treatment during a global pandemic. *Nat Rev Clin Oncol* (2020) 17(5):268–70. doi: 10.1038/s41571-020-0362-6
11. Chargari C, Chopra S, Viswanathan AN, Deutsch E. Brachytherapy Issues And Priorities In The Context of COVID-19 Outbreak. *Adv Radiat Oncol* (2020) 5(4):640–3. doi: 10.1016/j.adro.2020.04.034
12. Zaorsky NG, Yu JB, McBride SM, Dess RT, Jackson WC, Mahal BA, et al. Prostate Cancer Radiotherapy Recommendations in Response to COVID-19. *Adv Radiat Oncol* (2020) 5(4):659–65. doi: 10.1016/j.adro.2020.03.010
13. Thomson DJ, Palma D, Guckenberger M, Balermpas P, Beitler JJ, Blanchard P, et al. Practice recommendations for risk-adapted head and neck cancer radiotherapy during the COVID-19 pandemic: an ASTRO-ESTRO consensus statement. *Int J Radiat Oncol Biol Phys* (2020) 107(4):618–27. doi: 10.1016/j.radonc.2020.04.019
14. Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. *N Engl J Med* (2020) 382(24):2302–15. doi: 10.1101/2020.03.26.20044446
15. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* (2020) 382(13):1199–207. doi: 10.1056/NEJMoa2001316
16. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* (2020) 323(18):1843–4. doi: 10.1001/jama.2020.3786
17. Yang Y, Yang M, Shen C, Wang F, Yuan J, Li J, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *medRxiv* (2020) 2020.02.11.20021493. doi: 10.1101/2020.02.11.20021493
18. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* (2020), 296(2):E32–40. doi: 10.1148/radiol.2020200642
19. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* (2020) 296(2):E115–E117. doi: 10.1148/radiol.2020200432
20. Kim H, Hong H, Yoon SH. Diagnostic Performance of CT and Reverse Transcriptase-Polymerase Chain Reaction for Coronavirus Disease 2019: A Meta-Analysis. *Radiology* (2020) 296(3):E145–E155. doi: 10.1148/radiol.2020201343
21. Hermans JJR, Groen Jw, Zwets E, Boxma-De Klerk BM, Van Werkhoven JM, Ong DSY, et al. Chest CT for triage during COVID-19 on the emergency department: myth or truth? *Emerg Radiol* (2020) 20:1–11. doi: 10.1007/s10140-020-01821-1
22. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational Consensus Statement from the Fleischner Society. *Radiology* (2020) 296(1):172–80. doi: 10.1148/radiol.2020201365
23. Peeling RW, Wedderburn CJ, Garcia PJ, Boeras D, Fongwen N, Nkengasong J, et al. Serology testing in the COVID-19 pandemic response. *Lancet Infect Dis* (2020) 20(9):e245–9. doi: 10.1016/S1473-3099(20)30517-X
24. To KK-W, Tsang OT-Y, Leung W-S, Tam AR, Wu T-C, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* (2020) 20(5):565–74. doi: 10.1016/S1473-3099(20)30196-1
25. Long Q-X, Liu B-Z, Deng H-J, Wu G-C, Deng K, Chen Y-K, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med* (2020) 26 (6):845–8. doi: 10.1038/s41591-020-0897-1
26. Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clin Infect Dis* (2020) ciaa344. doi: 10.1093/cid/ciaa344
27. Zhang B, Zhou X, Zhu C, Feng F, Qiu Y, Feng J, et al. Immune phenotyping based on neutrophil-to-lymphocyte ratio and IgG predicts disease severity and outcome for patients with COVID-19. *medRxiv* (2020) 7:157. doi: 10.1101/2020.03.12.20035048
28. World Health Organization Scientific Brief. *Advice on the use of point-of-care immunodiagnostic tests for COVID-19*. Available at: https://www.who.int/news-room/commentaries/detail/advice-on-the-use-of-point-of-care-immunodiagnostic-tests-for-covid-19 [cité 22 juill 2020].
29. Minaya Flores P, Lasserre A, Zeghari-Squalli N, Carbonneil C, Morin-Surroca M, Dalour S. *Place des tests sérologiques rapides (TDR, TROD, autotests) dans la stratégie de prise en charge de la maladie COVID-19*. Haute Autorité de santé (HAS) (2020). p. 34. Available at: https://www.has-sante.fr/upload/docs/application/pdf/2020-05/rapport_tests_serologiques_rapides_covid-19_vd.pdf.
30. Pereira GC, Traughber M, Muzic RF. The role of imaging in radiation therapy planning: past, present, and future. *BioMed Res Int* (2014) 2014:231090. doi: 10.1155/2014/231090
31. Mettler FA, Huda W, Yoshizumi TT, Mahesh M. Effective Doses in Radiology and Diagnostic Nuclear Medicine: A Catalog. *Radiology* (2008) 248(1):254–63. doi: 10.1148/radiol.2481071451
32. Smith-Bindman R, Moghadassi M, Wilson N, Nelson TR, Boone JM, Cagnon CH, et al. Radiation Doses in Consecutive CT Examinations from Five University of California Medical Centers. *Radiology* (2015) 277(1):134–41. doi: 10.1148/radiol.2015142728
33. Benadjaoud MA, Bezin J, Veres A, Lefkopoulos D, Chavaudra J, Bridier A, et al. A multi-plane source model for out-of-field head scatter dose

calculations in external beam photon therapy. *Phys Med Biol* (2012) 57 (22):7725–39. doi: 10.1088/0031-9155/57/22/7725
34. Caraiani C, Pop A, Calin A, Ciobanu L, Militaru C, Berghe A, et al. Incidental findings during follow-up scans in oncological patients. *Clujul Med* (2018) 91 (3):293–9. doi: 10.15386/cjmed-931
35. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* (2020) 581(7809):465–9. doi: 10.1038/s41586-020-2196-x
36. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* (2020) 26(5):672–5. doi: 10.1038/s41591-020-0869-5
37. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* (2020) 368(6491):eabb6936. doi: 10.1126/science.abb6936
38. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance* (2020) 25(5):2000062. doi: 10.2807/1560-7917.ES.2020.25.5.2000062
39. Kuderer NM, Choueiri TK, Shah DP, Shyr Y, Rubinstein SM, Rivera DR, et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. *Lancet* (2020) 395(10241):1907–18. doi: 10.1016/S0140-6736(20)31187-9
40. Salje H, Kiem CT, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. *Science* (2020) 369 (6500):208–11. doi: 10.1126/science.abc3517
41. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, et al. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. *Radiology* (2020) 295(3):200463. doi: 10.1148/radiol.2020200463
42. Lepelletier D, Grandbastien B, Romano-Bertrand S, Aho S, Chidiac C, Géhanno JF, et al. (2020). What face mask for what use in the context of COVID-19 pandemic? The French guidelines. *J Hosp Infect* 105(3):414–8. Advance online publication. doi: 10.1016/j.jhin.2020.04.036
43. Madariaga A, McMullen M, Sheikh S, Kumar R, Liu F-F, Zimmermann C, et al. COVID-19 testing in cancer patients: Does one size fit all? *Clin Cancer Res* (2020) 26(18):4737–42. doi: 10.1158/1078-0432.CCR-20-2224

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor is currently organizing a Research Topic with the authors SA and LD.

frontiers in Oncology

ORIGINAL RESEARCH
published: 08 December 2020
doi: 10.3389/fonc.2020.578895

Radiomics Feature Activation Maps as a New Tool for Signature Interpretability

Diem Vuong[1], Stephanie Tanadini-Lang[1], Ze Wu[1], Robert Marks[1], Jan Unkelbach[1], Sven Hillinger[2], Eric Innocents Eboulet[3], Sandra Thierstein[3], Solange Peters[4], Miklos Pless[5], Matthias Guckenberger[1] and Marta Bogowicz[1]*

[1] Department of Radiation Oncology, University Hospital Zurich and University of Zurich, Zurich, Switzerland, [2] Department of Thoracic Surgery, University Hospital Zurich and University of Zurich, Zurich, Switzerland, [3] Department of Clinical Trial Management, Swiss Group for Clinical Cancer Research (SAKK) Coordinating Center, Bern, Switzerland, [4] Department of Oncology, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland, [5] Department of Medical Oncology, Kantonsspital Winterthur, Winterthur, Switzerland

OPEN ACCESS

Edited by:
Laurent Dercle,
Columbia University Irving Medical Center, United States

Reviewed by:
Hyunjin Park,
Sungkyunkwan University, South Korea
Rivka R. Colen,
University of Pittsburgh Medical Center, United States

***Correspondence:**
Diem Vuong
diem.vuong@usz.ch

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 01 July 2020
Accepted: 22 October 2020
Published: 08 December 2020

Citation:
Vuong D, Tanadini-Lang S, Wu Z, Marks R, Unkelbach J, Hillinger S, Eboulet EI, Thierstein S, Peters S, Pless M, Guckenberger M and Bogowicz M (2020) Radiomics Feature Activation Maps as a New Tool for Signature Interpretability. Front. Oncol. 10:578895. doi: 10.3389/fonc.2020.578895

Introduction: In the field of personalized medicine, radiomics has shown its potential to support treatment decisions. However, the limited feature interpretability hampers its introduction into the clinics. Here, we propose a new methodology to create radiomics feature activation maps, which allows to identify the spatial-anatomical locations responsible for signature activation based on local radiomics. The feasibility of this technique will be studied for histological subtype differentiation (adenocarcinoma *versus* squamous cell carcinoma) in non-small cell lung cancer (NSCLC) using computed tomography (CT) radiomics.

Materials and Methods: Pre-treatment CT scans were collected from a multi-centric Swiss trial (training, n=73, IIIA/N2 NSCLC, SAKK 16/00) and an independent cohort (validation, n=32, IIIA/N2/IIIB NSCLC). Based on the gross tumor volume (GTV), four peritumoral region of interests (ROI) were defined: lung_exterior (expansion into the lung), iso_exterior (expansion into lung and soft tissue), gradient (GTV border region), GTV+Rim (GTV and iso_exterior). For each ROI, 154 radiomic features were extracted using an in-house developed software implementation (Z-Rad, Python v2.7.14). Features robust against delineation variability served as an input for a multivariate logistic regression analysis. Model performance was quantified using the area under the receiver operating characteristic curve (AUC) and verified using five-fold cross validation and internal validation. Local radiomic features were extracted from the GTV+Rim ROI using non-overlapping 3x3x3 voxel patches previously marked as GTV or rim. A binary activation map was created for each patient using the median global feature value from the training. The ratios of activated/non-activated patches of GTV and rim regions were compared between histological subtypes (Wilcoxon test).

Results: Iso_exterior, gradient, GTV+Rim showed good performances for histological subtype prediction ($AUC_{training}$=0.68–0.72 and $AUC_{validation}$=0.73–0.74) whereas GTV and lung_exterior models failed validation. GTV+Rim model feature activation maps showed

that local texture feature distribution differed significantly between histological subtypes in the rim (p=0.0481) but not in the GTV (p=0.461).

Conclusion: In this exploratory study, radiomics-based prediction of NSCLC histological subtypes was predominantly based on the peritumoral region indicating that radiomics activation maps can be useful for tracing back the spatial location of regions responsible for signature activation.

Keywords: lung cancer, computed tomography, peritumoral radiomics, radiomics activation maps, local radiomics, interpretability

INTRODUCTION

Personalization of therapy options for patients with oncological diseases has gained great importance in recent years. Differentiation of non-small cell lung cancer (NSCLC) patients into histological subtypes, i.e., lung adenocarcinoma (ADC, ~50%) and squamous cell carcinoma (SCC, ~40%) (1) is for example an important factor in the choice of systemic treatments (2). Current biomarker assessments are often based on invasive interventions to extract a single-pin-pointed measurement. Consequently, there are many clinical scenarios with a clinical need for alternatives to tissue-based assessment of tumor histology: e.g., challenging anatomical locations for biopsy, unfavorable risk-benefit ratio for biopsy, history of more than one malignancy, or characterization of two simultaneously identified lung nodules.

Quantitative, image-based biomarkers, so-called radiomic features, can potentially overcome these obstacles (3–6). Extracted from medical images such as computed tomography (CT), those features rely on mathematical definitions to depict image-related characteristics. Features can often be subdivided into four main types: shape, intensity, texture and filtered based features, providing a 3D profile of the region of interest (ROI) (2). Radiomics has shown increasingly its potential usefulness in diagnosis, prognosis and response assessment (4, 6–8). For example, Aerts *et al.* showed that CT based radiomics was able to predict overall survival (OS) in NSCLC and head and neck cancer patients (concordance index=0.65, 0.69, respectively) treated with radiochemotherapy (5). Further, radiomics was reported prognostic for NSCLC patients treated with targeted therapies such as nivolumab, docetaxel and gefitinib with promising results (9, 10). Next to OS, other endpoints have been reported such as disease-free survival (11) or distant metastasis (12). Moreover, radiomics has shown to be useful for response assessment, i.e., in prediction of pathological complete response (13, 14). Identified radiomics features prognostic for survival in NSCLC were associated with image related tumor heterogeneity in CT imaging (15), i.e., entropy (16) or busyness (7) based on filtered images.

However, the quantitative and highly complex methodical nature of radiomics is a two-edged sword. Compared to manually assessed measures in radiological reports, these radiomic features lack in their interpretability, challenging the methodology to emerge from a research topic to a useful tool in clinical settings.

Gradually, this hurdle has been recognized and few research groups have attempted to improve the feature interpretability. One strong motion is to correlate radiomic features with known biological markers such as human papillomavirus (17, 18) or epidermal growth factor receptor (19–21). However, the biological data is often only of limited availability. In contrast, local radiomic features can be used to provide more spatial information about given signatures. Local radiomics refers to the extraction of radiomic features from small sub-regions (patches), which cover the complete ROI. Compared to traditional global radiomic features, the spatial location of these patches is known and hence differences in radiomics signatures can be determined on a smaller spatial scale. Bogowicz et al. for example showed that local radiomics differed substantially between recurrent to non-recurrent regions in head and neck cancer treated with radiotherapy (22). Local radiomics may not only serve as a detection tool, but the additional spatial information obtained from the patches potentially allows to trace the regions which are most revealing for a particular radiomics signature.

It is the aim of this exploratory study to create and analyze CT radiomics signature activation maps using local radiomics. As a case-study, we built tumoral and peritumoral radiomics models using a multi-centric imaging dataset to predict NSCLC histological subtypes. Local radiomic features were extracted for the model features to create radiomics feature activation maps. These maps were assessed to evaluate whether the tumoral or peritumoral region is more informative for NSCLC histology differentiation in pre-treatment CT.

MATERIALS AND METHODS

Patient and Imaging Characteristics

Patient and imaging characteristics were integrated from a previous study (23). For the training cohort, pre-treatment CT scans were collected from 73 stage IIIA/N2 NSCLC patients from a prospective Swiss multi-centric randomized phase 3 trial (SAKK 16/00 (5), neoadjuvant chemotherapy or radiochemotherapy prior to surgery). For the validation cohort, CT scans of 32 stage IIIA/N2 or IIIB NSCLC patients were included (induction radiochemotherapy or chemotherapy only prior to surgery) which were treated at the University Hospital Zurich (USZ). Patients with histological subtypes ADC and SCC were selected for this study. Histology as well as patient staging [6th edition of the tumor-node-metastasis

(TNM) classification] were defined according to the SAKK 16/00 protocol (5). Patients were similarly distributed between ADC and SCC subtype in the training and validation cohort (61.6 and 56.3% of ADC patients in training and validation, respectively) (**Supplement A, Table 1**).

Patients received non-contrast enhanced, non-gated pre-treatment CT scans reconstructed with filtered-back projection (FBP) using standard convolution kernel. Due to the multi-centric imaging set, we defined the standard kernel as follows: GE—STANDARD, Siemens—B30f/B31f, Toshiba—FC18, and Philips—B, similarly to the phantom study of Mackin and Ger et al. (24, 25). CT spatial resolution varied between 0.98 and 1.37 mm in-plane and 0.6 to 5.0 mm slice thickness. Patients from the validation cohort received a non-contrast enhanced average CT and were imaged on CT scanner Discovery RX, STE, 690 and Biograph 128 Edge, 128, 40, 6 and SOMATOM Definition AS, from GE MEDICAL SYSTEMS and SIEMENS. Scans were reconstructed with FBP and a smooth kernel (STANDARD, I30f, B31f). CT spatial resolution was 0.98, 1.17, 1.37 mm in-plane and 2 and 3.27 mm slice thickness.

Delineation

Five ROIs were defined for this study (**Figure 1**):

A. **GTV**: visual extent of the gross tumor volume (GTV)
B. **lung_exterior**: 0.8 cm expansion from the GTV into lung tissue only
C. **iso_exterior**: 0.8 cm expansion from the GTV into lung and soft tissue
D. **gradient:** 0.4 cm contraction and 0.8 cm expansion from the GTV
E. **GTV+Rim**: union of GTV and 0.8 cm expansion from GTV (iso_exterior)

The GTV ROI was manually delineated on the CT scans by an experienced physician using MIM VISTA (Version 6.7.9., MIM Software Inc., Cleveland, USA) with the lung window level and the support of registered PET images. All ROIs except for the GTV will be referred to as peritumoral ROIs. These peritumoral ROIs were created using an in-house developed MIM workflow. Anatomical structures which would strongly disturb the analysis, e.g., consisting of large air cavities (bronchi) or dense structures (bones) were manually excluded from all ROIs. Further, patients were excluded from the gradient analysis if the gradient regions comprised the entire GTV (**Figure 1D**).

Robustness Study

The creation of the peritumoral ROIs was based on the manual delineation of the GTV, therefore a robustness study was performed to study the impact of inter-observer delineation variability on the radiomic features. A separate set of eleven patients were used as described in the study of Pavic et al. (26). Three independent observers from USZ manually delineated the GTV. The same MIM workflow was used to create the peritumoral ROIs with the GTV of the three observers as an input. The intra-class correlation coefficient (ICC) was used as stability measure as described in Pavic et al. (26). However, a stricter acceptance level of 0.9 was chosen, i.e., radiomic features with ICC > 0.9 were considered stable.

Radiomics

Pre-treatment CT scans were resampled to 3.75 mm cubic voxels, the 75th percentile of slice thicknesses in the training dataset using linear interpolation. Radiomic calculations were performed using an in-house developed software implementation (Z-Rad) based on Python programming language v 2.7.14 (for details on the software and features, please consult: https://medical-physics-usz.github.io). A Hounsfield unit (HU) range of −1,024 to 200 HU was chosen to exclude bone structures which could not be accounted for manually. Since the expansion and contraction parameters for the peritumoral ROIs were fixed, no shape features were considered for the analysis. Further, due to the small number of voxels in each direction, no wavelet features were included. Hence, a total of 154 radiomic features were calculated, i.e., intensity (n = 17) and texture (n = 137). Feature definitions were standardized according to the image biomarker standardization initiative (IBSI, version 11) (27). A fixed bin size of 20 HU was used to discretize the grey level values

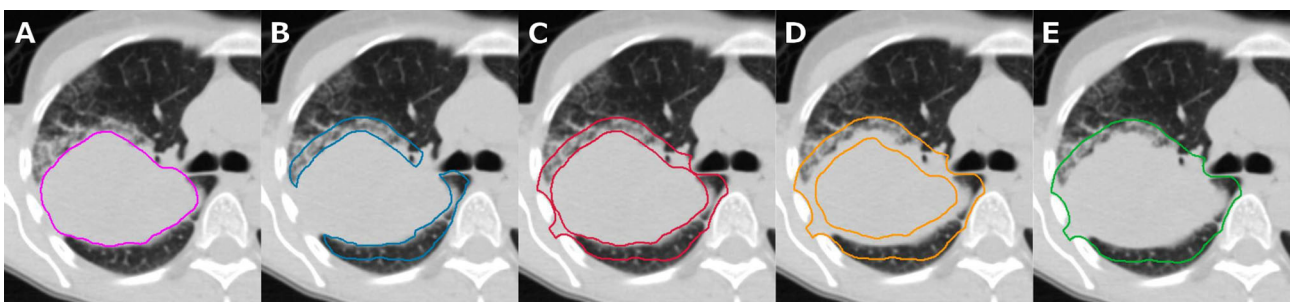


FIGURE 1 | Same axial slice of a patient in our cohort shown for tumoral and peritumoral region of interests (ROIs), i.e., visual extent of the primary tumor (**A**, GTV), 0.8 cm expansion into lung tissue region inside the lung (**B**, lung_exterior), 0.8 cm expansion into lung and soft tissue (**C**, iso_exterior), 0.4 cm contraction and 0.8 cm expansion from the GTV (**D**, gradient), and primary tumor including iso_exterior (**E**, GTV+Rim).

for texture analysis, resulting in approximately 60 bins, which has been shown to reduce intrinsic noise in the images while preserving essential texture (28).

Statistical Analysis for Global Radiomics

To reduce the number of features, principal component analysis (PCA) was performed as a feature reduction method (29). The retained principal components were defined based on the 95% data variance. The feature which correlated the most with the selected principal component was used as a surrogate (the largest Pearson correlation coefficient). Univariate logistic regression analysis was performed to determine individual prognostic power of each features, separately. The significance level was 0.05, with no correction for multiple testing. Based on features with highest prognostic power per principal component group, a multivariate logistic regression model was built with backward selection using Akaike information criterion (AIC) which balances the goodness of fit of the model and its simplicity (30). The discriminatory power of the models was quantified using the area under the receiver operating characteristic curve (AUC) along with its 95% confidence interval (CI). Model performance was verified using 5-fold cross validation. Folds were chosen randomly without repetition. The generalizability of the models was verified in the validation cohort. Statistical analysis, model building and validation were performed with R [Version 3.5.1, packages: base, survival (31), survcomp (32), boot (33), pROC (34), and glmnet (35)].

Creation of Activation Maps Based on Local Radiomics

Local radiomic features were extracted from the GTV+Rim ROI using non-overlapping patches of size 3x3x3 voxels. This size of the patches allowed a meaningful calculation of the texture features (minimum number of voxels in each direction) as well as a meaningful overlap with the rim region (0.8 cm margin in each direction and 3.75 mm voxel size). The placement of the patches was automatically optimized to cover the entire ROI with a minimum number of patches. Patches with a low number of informative voxels (< 9 voxels) of the ROI were discarded. The overlap of the patches with the GTV was assessed, i.e., 100% referred to patches comprising only the GTV and 0% to patches comprising only normal tissue. This tool is intended to determine whether the radiomics signature for the prediction of histological subtypes originates from a certain predefined region. The signatures of patches with mixed overlap (10% to 90%) contain ambiguous information and were therefore discarded to clearly distinguish patches spatially assigned rim or GTV. Finally, patches with overlap lower than 10% or larger than 90% were labeled as rim and GTV, respectively. A binary feature activation map was created for each individual patient using the respective median of the global (standard) feature value in the training cohort, i.e., patches with feature value larger than the median were considered activated. The ratios of activated/non-activated patches for the normal tissue and the GTV were compared in the validation cohort between the histology types (Wilcoxon test), considering only patients with at least 27 patches and a minimum 3 patches per region (**Figure 2**, a more detailed description can be found in **Supplement D**).

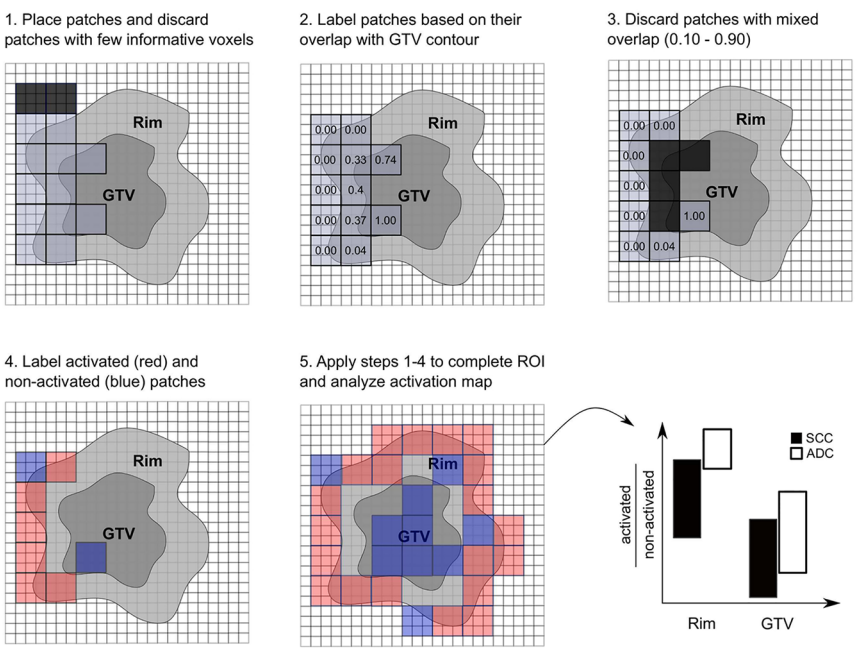


FIGURE 2 | Scheme of radiomics feature activation map creation. Patches were optimally placed and patches with few informative voxels (< 9) were discarded (1), patches were labeled according to their overlap with the gross tumor volume (GTV) contour (2), and patches with mixed overlap were discarded (3). Patches were labeled activated (red) if their feature value was larger than the global median and were labeled non-activated (blue) if their feature value was smaller than the global median. The activation ratio was analyzed per region and histological subtype (5).

RESULTS

Modeling and Validation

Robust features were identified from the inter-observer variability robustness studies. Overall, the number of stable features for each ROIs were found to be moderate, i.e., GTV (49.7%), lung_exterior (57.6%), iso_exterior (57.6%), gradient (55.8%), and GTV+Rim (74.5%). The analysis can be found in the **Supplement B**. Results of the univariate analysis of the robust features selected in the feature selection step are shown in **Table 1**. Features marked with an asterisk were the final features retained after backward selection. Overall good univariate performances on the training set were observed (AUC = 0.61 to 0.72).

Different methods of feature selection were tested. The PCA + univariate logistic regression feature selection method led to simpler models. For majority of the ROIs, the models using PCA + univariate logistic regression performed best compared to other feature selection methods (**Supplement C**).

For all regions, a logistic regression model could be built. The five-fold cross validation performance was [mean AUC (range)]: GTV [0.625 (0.23–1.00)], lung_exterior [0.72 (0.68–0.78)], iso_exterior [0.67 (0.46–0.84)], gradient [0.70 (0.48–0.82)], GTV+Rim [0.67 (0.48–0.84)]. The models based on the GTV and lung_exterior ROI were the only models which could not be validated on the validation cohort with 95% CI covering AUC = 0.5, i.e., a performance of a random predictor. Iso_exterior, gradient and the GTV+Rim showed acceptable performances in the range of 0.68–0.72 in the training and 0.73–0.74 in the validation cohort (**Figure 3**).

Model Features

In **Table 1**, the coefficients of the final model features are listed. The final models consisted of one or two features. Each model consisted of one texture features. These texture features can be associated with the texture heterogeneity in the ROI. For example, the gray level coocurrence matrix (GLCM) inverse variance in the GTV model is small if there is higher variance (**Figure 4**). The median GLCM inverse variance was lower for ADC compared to SCC, i.e., ADC were more likely to have heterogeneous and SCC more homogeneous patterns (**Figure 4**). For all regions, iso_exterior, gradient and the GTV+Rim, one texture feature (GLSZM_zone size non-uniformity normalized) was present in all three models. This feature counts the homogeneous zones of the same size over the different zone sizes and is low in patterns where zone counts are equally distributed along zone sizes, i.e., more heterogeneous patterns (**Figure 4**). In the models, the higher texture value (more homogeneous pattern) was associated more with SCC patients. Further, since this feature was present in the all three models, this feature will most likely be associated in the tumor adjacent region within the stable performing GTV+Rim model. Using the activation maps we further validate this assumption (see next section). Interestingly, the iso_exterior the 90% percentile intensity feature was significant more relevant in the model compared to the texture features whereas in the GTV+Rim model the opposite was observed (**Figure 4**).

Analysis of Radiomics Feature Activation Maps

The activation map analysis of the full radiomics signature indicated a greater importance of the rim region compared to the GTV (p=0.0541 and p=0.302 for rim and GTV, respectively). A closer analysis on the individual features showed that visually the texture feature was more activated on the adjacent region of the tumor, the intensity median more in the tumoral region. The median split values from the training cohort was 0.526 for GLSZM_zone size non-uniformity normalized and −158 HU for intensity_median. There was a significant difference in the activation ratio in the rim region when comparing ADC *vs.* SCC patients (p=0.048), however the ratio was non-significant in the tumor region (p=0.461). No significant difference in activated/non-activated ratio was observed in both regions for intensity median (**Figure 5**).

DISCUSSION

Ideally, clinically useful prognostic models should be performing reliably and be comprehensive. With the growing complexity of

TABLE 1 | Overview of univariate and multivariate analysis shown for all region of interests (ROIs) considered.

		Univariate		Multivariate	
ROI	**Features**	**AUC**	**p-value**	**Coefficient**	**p-value**
GTV	GLRLM_run entropy	0.64	0.034		
	GLCM_inverse variance*	0.65	0.035	−10.83	0.035
lung_exterior	GLSZM_zone percentage	0.63	0.046		
	GLCM_contrast	0.65	0.043		
	GLCM_homogeneity normalized*	0.72	0.004	52.51	0.004
	NGLDM_low dependence emphasis	0.66	0.025		
iso_exterior	Intensity_median	0.65	0.028		
	GLCM_correlation	0.61	0.043		
	GLSZM_zone size non-uniformity normalized*	0.68	0.015	−12.84	0.112
	Intensity_percentile_90*	0.68	0.025	−0.01	0.072
gradient	GLSZM_zone size non-uniformity normalized*	0.68	0.046	−20.26	0.046
GTV+Rim	Intensity_median*	0.63	0.026	0.002	0.144
	GLSZM_zone size non-uniformity normalized*	0.69	0.010	−16.75	0.026

Only features are listed which had a significant performance in the univariate analysis per principal component group. Features with an asterisk were retained in the final models after backward selection and their coefficients and p-values in the multivariate analysis are listed.

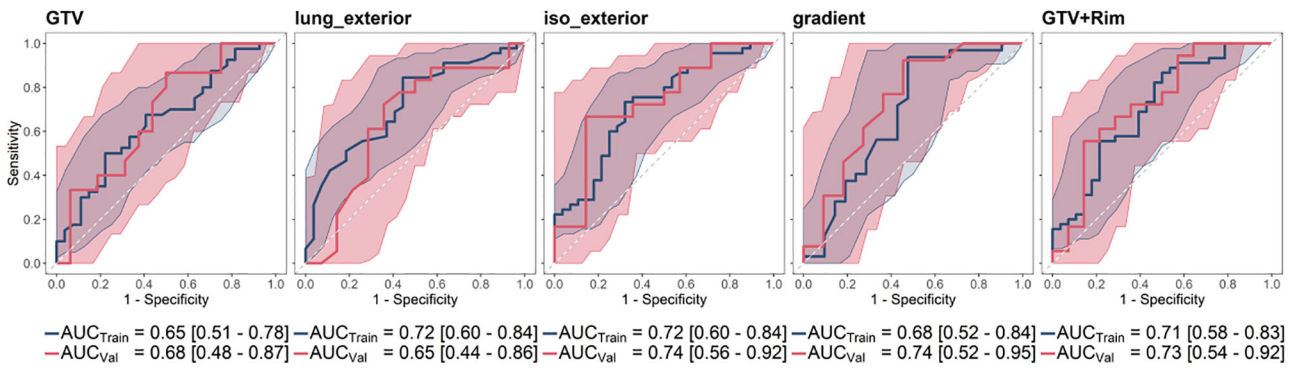


FIGURE 3 | Receiver operating characteristic curve (ROC) curves and corresponding mean AUC [95% confidence interval] of the analyzed ROIs shown for training (blue) and validation cohort (red). The radiomic models based on GTV and lung_exterior could not be successfully validated in the validation cohort. Iso_exterior, gradient and GTV+Rim models had good performances in both cohorts.

hand-crafted radiomic features, the feature interpretability becomes more relevant for its successful incorporation into clinical settings. Tools allowing feature interpretability may help in filtering false positive results in signature validation or clinical use.

In this exploratory study, we used a new local radiomics approach to create radiomics feature activation maps to locate the regions responsible for signature activation. On a local scale we were able to study whether the peritumoral or the tumoral radiomics was more informative for NSCLC histology differentiation in CT. To our knowledge, this is the first study to correlate peritumoral radiomic features with NSCLC histological subtypes. Multivariate logistic regression models were built for each ROI using features robust against inter-observer delineation variability. Iso_exterior, gradient as well as the combination of GTV and iso_exterior (GTV+Rim) showed acceptable performances in the range of AUC = 0.68–0.72 in the training and AUC=0.73–0.74 in the validation cohort whereas GTV and lung_exterior ROI models failed to validate. GTV+Rim radiomics feature activation maps for each patient showed that the rim region was more informative compared to tumoral radiomics to differentiate ADC and SCC.

CT based tumoral radiomic models have shown to be able to discriminate NSCLC histological subtypes, i.e., capturing that ADC cells are more loosely organized while SCC is more densely structured (36). Model performances however were not consistent across different studies ranging from moderate (37, 38) to good (36, 39–41). Possible explanation for worse performance of our GTV-based model compared to others may lay in the different imaging settings used. Reported models incorporated contrast-enhanced (17), respiratory-gated (15) CT scans or more complex modeling techniques such as Bayesian network (37). In our dataset, we selected a subset of patients with similar reconstruction settings resulting in small inter-scanner effects similar to previous studies (24, 25) potentially influencing the performance of the model. Further, scans were acquired in free-breathing which can introduce blurring to the final image (42, 43). However, in agreement with previous studies, ADC and SCC patients had a different tissue structure, i.e., the median of the mean intensity was significantly smaller for patients with ADC compared to SCC ($p < 0.05$). The final GTV model feature (GLCM_inverse variance) was lower for ADC patients compared to SCC patients, reflecting the more loosely structured tumor in ADC patients compared to more densely structured tissue of SCC patients. Lower GLCM_inverse variance feature can be associated with higher heterogeneity in the tissue in agreement with other studies (41), i.e., higher entropy values (associated with higher tumor heterogeneity) were observed to be associated with ADC tumors (40).

We hypothesized that peritumoral radiomics can depict better the known association between the anatomical tumor location and histological subtypes, i.e., ADC occur in more peripheral regions while SCC are often located centrally (1). This association is assumed to be most evident in iso_exterior ROI where the captured adjacent soft tissue structures can reflect the periphery or centrality of the primary tumor location. Indeed, the median 90% percentile in ADC was lower compared to SCC indicating less dense structures in the ROI. Further, it has been shown that the microscopic tumor extension in the peritumoral region differs between ADC and SCC. As a result, it has been suggested to use different margin sizes when treating ADC and SCC tumors to cover 95% of the microscopic tumor extension (8 and 6 mm margin for ADC and SCC, respectively) (44). With the chosen 8 mm margin for the iso_exterior region, this peritumoral model may depict this difference in the cell distribution.

The presence of the GLSZM_zone size non-uniformity normalized feature in the peritumoral ROI models indicated that this feature varies stronger between different histologies in the rim rather than in the GTV region in the GTV+Rim model. The activation maps of the GTV+Rim model confirmed this observation, i.e., the distribution of local texture feature of ADC differed significantly from SCC for the rim ($p=0.048$) but not for the GTV ($p=0.461$), irrespective that the feature threshold was based on the global feature values and no feature scaling was applied.

To account for the inter-observer delineation variability, a robustness study was performed for the primary tumor and peritumoral ROIs. The peritumoral feature stability was moderate, interestingly however, similar or even more stable

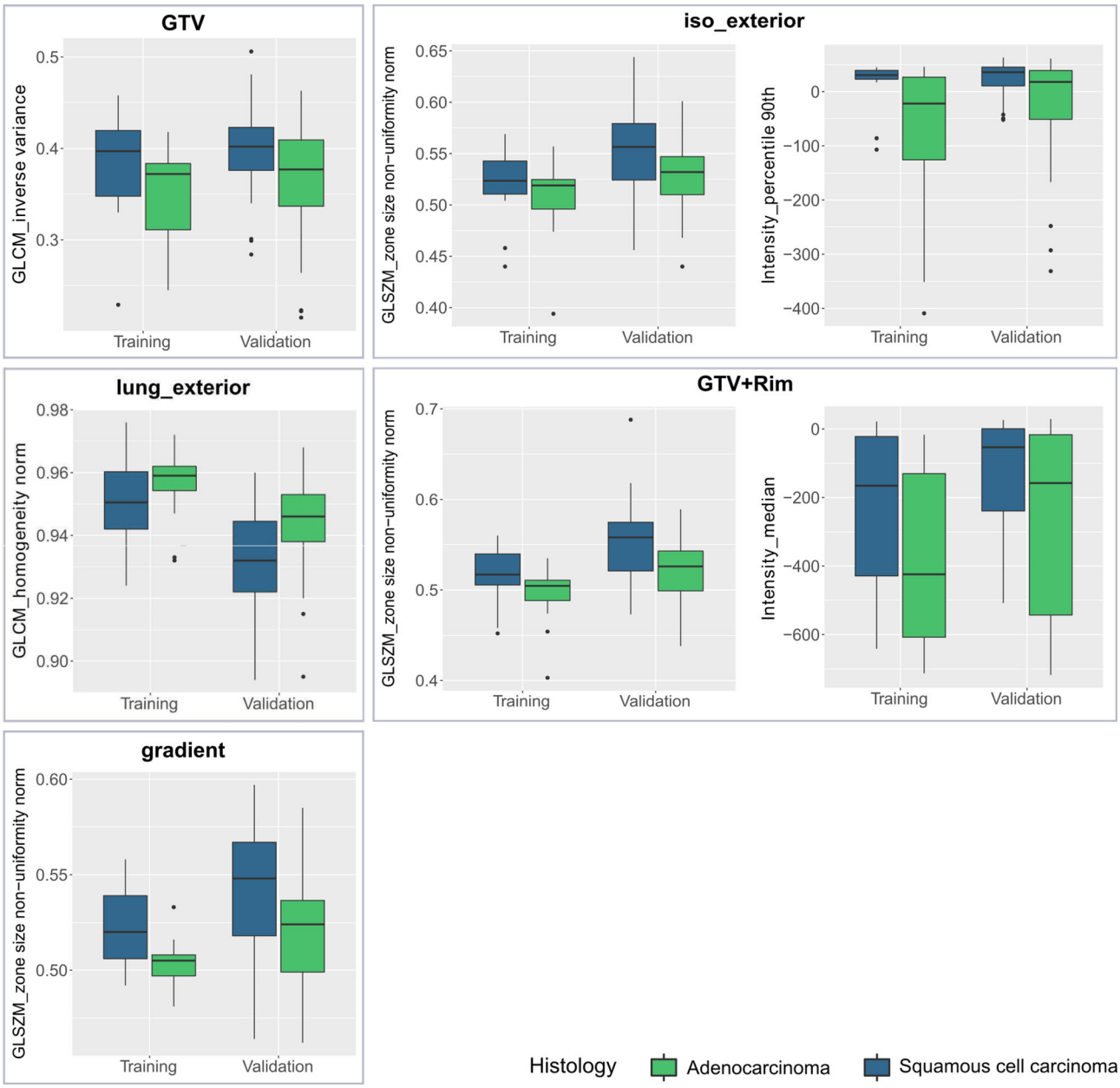


FIGURE 4 | Boxplots of each studied region of interest (ROI) stratified by histological subtype (adenocarcinoma, squamous cell carcinoma) and dataset (training, validation) for the final model features.

than the primary tumor radiomics. A possible reason can be a different amount of lung tissue in the primary tumor delineation, which can result in higher sensitivity to manual delineations of the primary tumor compared to peritumoral regions where substantial lung tissue was *a priori* present. Further, an increased stability for larger ROI sizes can be observed. This observation is in agreement with Tunali et al.'s inter-observer variability study, where, however, the initial primary tumor contours were delineated using three semi-automatic segmentation methods (45). Arguably, the strict acceptance level ICC > 0.9, could have discarded potential useful features. However, due to the small cohort of 11 patients for the robustness analysis, the strict acceptance level helps ensuring that results were not affected by the small sample size. Recently, it has been shown that discarding features based on their robustness will lead to different models compared to modeling using a standardized imaging allowing to include all features (23).

The localization of signature relevant regions in the context of activation maps has been established in deep learning methods.

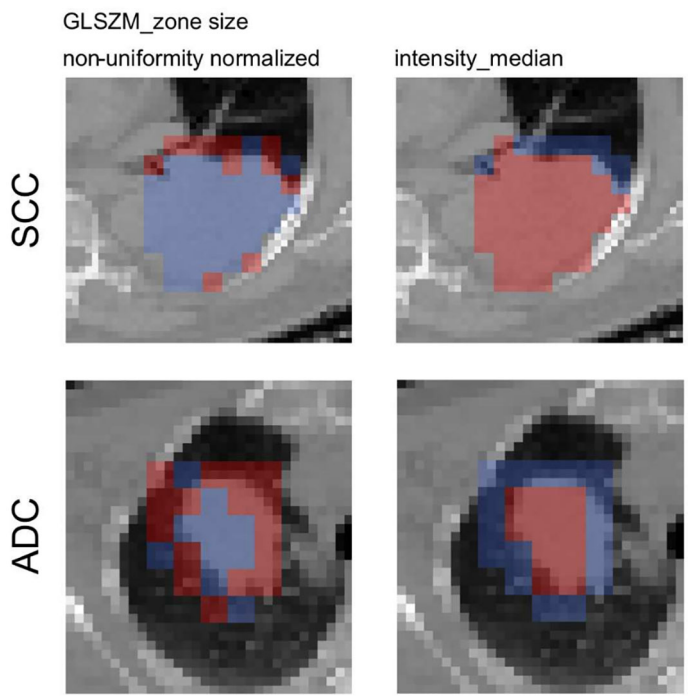


FIGURE 5 | Axial slices of the feature activation maps overlaid with the corresponding CT scan of a squamous cell carcinoma patient (SCC) and an adenocarcioma patient (ADC) from the validation cohort. Activated (red) patches had feature values larger than the median feature value from the training, whereas non-activated (blue) patches had feature values smaller than the median. The texture feature was activated mostly in the rim region whereas the intensity_median was activated mostly in the tumor region.

Activation maps are pre-dominantly used to identify areas of interest used from the neural network to perform its class prediction. Introduction of such activation maps into the field of radiomics may provide an addition for clinical interpretability of radiomic models. In the context of peritumoral radiomics for example, where various peritumoral region definitions were reported in different sites (46), no clear strategy was available to determine the most promising region other than to model and validate each region individually. Therefore the tool presented in this study may guide the user to select the most relevant region in a more efficient way. Further, the presented tool can be not only applied on individual features but could be useful to interpret a complete signature for example by combining the activation maps of the model features. However, in our case, the texture feature had a more important role for the modeling compared to the intensity feature, therefore we did not include an analysis combining both activation maps.

It is important to study the link between local and global features. For intensity features, the global features do not necessary have to reflect the spatial saliency on a local scale, as they are not scale-invariant. For texture features, the distribution of discretized intensities need to be preserved between the local and global approach. For example, in our study, the same discretization for local and global radiomics was used (fixed bin size of 20 HU with bin 0 corresponding to minimum intensity in the entire ROI). For majority of texture features the link between global and local can be argued on the basis that they are calculated on the relationship of a single voxel to its immediate neighbor (e.g., GLCM, NGTDM) consistent with the definition of our patches (3x3x3 voxels). The patch size should be adjusted in situations where larger distances are used for the texture metrics calculation. In more complex metrics (e.g., GLSZM, GLRLM), further analysis is required to study the main driving factor of the feature values. However, in our study both iso_exterior and the GTV+Rim model shared the same feature (GLSZM_zone_size_non-uniformity_normalized) indicating a close link of that feature to the rim region, further also observed on the local scale. Similarly, another study showed that CT based local radiomics was useful to identifying subregions of head and neck tumors associated with different degrees of radiation curability, i.e., local features differed significantly between recurrent region and controlled (non-recurrent) region (22). In that study, heterogeneity on both global and local scale was linked to worse prognosis. Irrespectively, a closer investigation is needed to identify the optimal activation threshold.

This study has its limitations. Higher complexity features such as wavelet features were not included, since a minimum number of voxels in each direction is needed to provide a meaningful analysis. The published tumoral radiomic models consisted of filter-based features such as law-features or wavelet features. However, these features were also more sensitive to

delineation variability (45). Further, strict cut-off values were chosen to differentiate patches originated from the GTV and rim (10% and 90%). These results will likely change when using different cut-off values. Out of the scope of this exploratory study was the use of different margin sizes for the definition of peritumoral ROI as well as the inclusion of clinical known prognostic factors which might have improved the presented model performances. A further limitation is our assumption that the tumor spreads isotropically radial from the primary tumor center of mass. However, we distinguished a tumor spread into the lung-only regions with an isotropic spread. Lastly, the small sample size could have impacted the results, further analysis incorporating more imaging data would be desired.

CONCLUSION

In this exploratory study we have shown that feature activation maps using local radiomics proved to be useful for tracing back the spatial location of regions responsible for signature activation. Radiomics feature activation map analysis indicated that the rim region, which is anatomically the tumor invasion front, was more relevant for histological subtype prediction than the GTV in CT imaging.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: radiomics data generated for this study are available on request from the corresponding author. Clinical data were provided by SAKK and are available upon request. Requests to access these datasets should be directed to DV, diem.vuong@usz.ch.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by all involved Swiss cantonal ethics committees with the following reference numbers: EKNZ PB_2016-01071, KEK ZH PB_2016-00412, KEK Bern PB_2016-01072, CER-VD PB_2016-01078, CCER PB_2016-01073, EKOS PB_2016-01075, Comitato Etico Cantonale Bellinzona PB_2016-01077 for the training cohort as well as KEK ZH 2018-02405 for the validation cohort. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Concept and design: DV, MB, ST-L. Acquisition and analysis: DV, MB, ZW, RM, JU. Data interpretation: DV, MB, ST-L, MG. Data collection: EE, ST, SH, MP, SP. Drafting of manuscript: DV, MB, ST-L. Critical revision of the manuscript: all authors. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Swiss National Science Foundation (310030_173303).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.578895/full#supplementary-material

REFERENCES

1. Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong K-K. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer* (2014) 14 (8):535–46. doi: 10.1038/nrc3775
2. Standfield L, Weston AR, Barraclough H, Kooten MV, Pavlakis N. Histology as a treatment effect modifier in advanced non-small cell lung cancer: A systematic review of the evidence. *Respirology* (2011) 16(8):1210–20. doi: 10.1111/j.1440-1843.2011.02025.x
3. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
4. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, ven Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141
5. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5006
6. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
7. Bogowicz M, Vuong D, Huellner MW, Pavic M, Andratschke N, Gabrys HS, et al. CT radiomics and PET radiomics: ready for clinical implementation? *Q J Nucl Med Mol Imag. Off Publ Ital Assoc Nucl Med AIMN Int Assoc Radiopharmacol IAR Sect Soc Of* (2019) 63(4):355–70. doi: 10.23736/s1824-4785.19.03192-3
8. Thawani R, McLane M, Beig N, Ghose S, Prasanna P, Velcheti V, et al. Radiomics and radiogenomics in lung cancer: A review for the clinician. *Lung Cancer* (2018) 115:34–41. doi: 10.1016/j.lungcan.2017.10.015
9. Aerts HJWL, Grossmann P, Tan Y, Oxnard GR, Rizvi N, Schwartz LH, et al. Defining a Radiomic Response Phenotype: A Pilot Study using targeted therapy in NSCLC. *Sci Rep* (2016) 6:33860. doi: 10.1038/srep33860
10. Dercle L, Fronheiser M, Lu L, Du S, Hayes W, Leung DK, et al. Identification of Non–Small Cell Lung Cancer Sensitive to Systemic Cancer Therapies Using Radiomics. *Clin Cancer Res* (2020) 26(9):2151–62. doi: 10.1158/1078-0432.CCR-19-2942
11. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* (2016) 281(3):947–57. doi: 10.1148/radiol.2016152234
12. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol* (2015) 114(3):345–50. doi: 10.1016/j.radonc.2015.02.015
13. Coroller TP, Agrawal V, Narayan V, Hou Y, Grossmann P, Lee SW, et al. Radiomic phenotype features predict pathological response in non-small cell

lung cancer. *Radiother Oncol* (2016) 119(3):480–6. doi: 10.1016/j.radonc.2016.04.004
14. Coroller TP, Agrawal V, Huynh E, Narayan V, Lee SW, Mak RH, et al. Radiomic-Based Pathological Response Prediction from Primary Tumors and Lymph Nodes in NSCLC. *J Thorac Oncol* (2017) 12(3):467–76. doi: 10.1016/j.jtho.2016.11.2226
15. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *Eur Radiol* (2012) 22(4):796–802. doi: 10.1007/s00330-011-2319-8
16. Dercle L, Ammari S, Bateson M, Durand PB, Haspinger E, Massard C, et al. Limits of radiomic-based entropy as a surrogate of tumor heterogeneity: ROI-area, acquisition protocol and tissue site exert substantial influence. *Sci Rep* (2017) 7(1):7952. doi: 10.1038/s41598-017-08310-5
17. Leijenaar RT, Bogowicz M, Jochems A, Hoebers FJ, Wesseling FW, Huang SH, et al. Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study. *Br J Radiol* (2018) 91(1086):20170498. doi: 10.1259/bjr.20170498
18. Bogowicz M, Riesterer O, Ikenberg K, Stieb S, Moch H, Studer G, et al. Computed Tomography Radiomics Predicts HPV Status and Local Tumor Control After Definitive Radiochemotherapy in Head and Neck Squamous Cell Carcinoma. *Int J Radiat Oncol Biol Phys* (2017) 99(4):921–8. doi: 10.1016/j.ijrobp.2017.06.002
19. Li S, Ding C, Zhang H, Song J, Wu L. Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. *Med Phys* (2019) 46 (10):4545–52. doi: 10.1002/mp.13747
20. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. *Cancer Res* (2017) 77(14):3922–30. doi: 10.1158/0008-5472.CAN-17-0122
21. Liu Y, Kim J, Balagurunathan Y, Li Q, Garcia AL, Stringfield O, et al. Radiomic Features Are Associated With EGFR Mutation Status in Lung Adenocarcinomas. *Clin Lung Cancer* (2016) 17(5):441–8.e6. doi: 10.1016/j.cllc.2016.02.001
22. Bogowicz M, Pavic M, Riesterer O, Finazzi T, Garcia Schüler H, Holz-Sapra E, et al. CT radiomics differentiates levels of radiocurability in tumor subvolumes in head and neck cancer, ESTRO 2020 Abstract book. Available at: https://cld.bz/8huStZo/262/ (Accessed November 15, 2020).
23. Vuong D, Bogowicz M, Denzler S, Oliveira C, Foerster R, Amstutz F, et al. Comparison of robust to standardized CT radiomics models to predict overall survival for non-small cell lung cancer patients. *Med Phys* (2020) 47(9):4045–53. doi: 10.1002/mp.14224
24. Ger RB, Zhou S, Chi P-CM, Lee HJ, Layman RR, Jones AK, et al. Comprehensive Investigation on Controlling for CT Imaging Variabilities in Radiomics Studies. *Sci Rep* (2018) 8(1):13047. doi: 10.1038/s41598-018-31509-z
25. Mackin D, Fave X, Zhang L, Yang J, Jones AK, Ng CS, et al. Harmonizing the pixel size in retrospective computed tomography radiomics studies. *PLoS One* (2017) 12(9):e0178524. doi: 10.1371/journal.pone.0178524
26. Pavic M, Bogowicz M, Würms X, Glatz S, Finazzi T, Riesterer O, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. *Acta Oncol* (2018) 57(8):1070–4. doi: 10.1080/0284186X.2018.1445283
27. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiol Publ Online March 10* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
28. Larue R, van Timmeren JE, de Jong EEC, Feliciani G, Leijenaar RTH, Schreurs WMJ, et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. *Acta Oncol* (2017) 56(11):1544–53. doi: 10.1080/0284186x.2017.1351624
29. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans R Soc Math Phys Eng Sci* (2016) 374 (2065):20150202. doi: 10.1098/rsta.2015.0202
30. Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* (1974) 19(6):716–23. doi: 10.1109/TAC.1974.1100705
31. Therneau TM, Lumley T. Package 'survival'. *R Top Doc* (2015) 128:112.
32. Canty A, Ripley B. boot: Bootstrap R (S-Plus) functions. *R Package Version* (2017) 1:3–20.
33. Schröder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* 2011 27(22):3206–8.
34. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf* (2011) 12(1):77. doi: 10.1186/1471-2105-12-77
35. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J J Stat Softw* (2010) 33(1):22. doi: 10.18637/jss.v033.i01 2010.
36. E L, Lu L, Li L, Yang H, Schwartz LH, Zhao B. Radiomics for Classifying Histological Subtypes of Lung Cancer Based on Multiphasic Contrast-Enhanced Computed Tomography. *J Comput Assist Tomogr* (2019) 43 (2):300–6. doi: 10.1097/RCT.0000000000000836
37. Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, et al. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front Oncol* (2016) 6:71. doi: 10.3389/fonc.2016.00071
38. Parmar C, Leijenaar RTH, Grossmann P, et al. Radiomic feature clusters and Prognostic Signatures specific for Lung and Head & Neck cancer. *Sci Rep* (2015) 5(1):1–10. doi: 10.1038/srep11044
39. Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, Fabro AT. Azevedo-Marques PM de. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. *Comput Methods Progr Biomed* (2018) 159:23–30. doi: 10.1016/j.cmpb.2018.02.015
40. Digumarthy SR, Padole AM, Gullo RL, Sequist LV, Kalra MK. Can CT radiomic analysis in NSCLC predict histology and EGFR mutation status? *Med (Baltimore)* (2019) 98(1):e13963. doi: 10.1097/MD.0000000000013963
41. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. *Eur Radiol* (2018) 28(7):2772–8. doi: 10.1007/s00330-017-5221-1
42. Oliver JA, Budzevich M, Zhang GG, Dilling TJ, Latifi K, Moros EG. Variability of Image Features Computed from Conventional and Respiratory-Gated PET/CT Images of Lung Cancer. *Transl Oncol* (2015) 8(6):524–34. doi: 10.1016/j.tranon.2015.11.013
43. Lafata K, Cai J, Wang C, Hong J, Kelsey CR, Yin F-F. Spatial-temporal variability of radiomic features and its effect on the classification of lung cancer histology. *Phys Med Biol* (2018) 63(22):225003. doi: 10.1088/1361-6560/aae56a
44. Giraud P, Antoine M, Larrouy A, Milleron B, Callard P, De Rycke Y, et al. Evaluation of microscopic tumor extension in non–small-cell lung cancer for three-dimensional conformal radiotherapy planning. *Int J Radiat Oncol* (2000) 48(4):1015–24. doi: 10.1016/S0360-3016(00)00750-1
45. Tunali I, Hall LO, Napel S, Cherezov D, Guvenis A, Gillies RJ, et al. Stability and reproducibility of computed tomography radiomic features extracted from peritumoral regions of lung cancer lesions. *Med Phys* (2019) 46 (11):5075–85. doi: 10.1002/mp.13808
46. Keek S, Sanduleanu S, Wesseling F, de Roest R, van den Brekel M, van der Heijden M, et al. Computed tomography-derived radiomic signature of head and neck squamous cell carcinoma (peri)tumoral tissue for the prediction of locoregional recurrence and distant metastasis after concurrent chemo-radiotherapy. *PLoS One* (2020) 15(5):e0232639. doi: 10.1371/journal.pone.0232639

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 07 January 2021
doi: 10.3389/fonc.2020.598253

Development and Validation of a Radiomic-Based Model for Prediction of Intrahepatic Cholangiocarcinoma in Patients With Intrahepatic Lithiasis Complicated by Imagologically Diagnosed Mass

Beihui Xue[1], Sunjie Wu[1], Minghua Zheng[2], Huanchang Jiang[3], Jun Chen[3], Zhenghao Jiang[3], Tian Tian[3], Yifan Tu[3], Huanhu Zhao[4], Xian Shen[5], Kuvaneshan Ramen[6], Xiuling Wu[7], Qiyu Zhang[8], Qiqiang Zeng[5] and Xiangwu Zheng[1]**

[1] Radiological Department, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [2] Department of Hepatology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [3] The First Clinical Medical College of Wenzhou Medical University, Wenzhou, China, [4] School of Pharmacy, Minzu University of China, Beijing, China, [5] The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, [6] Dr A. G Jeetoo Hospital, Port Louis, Mauritius, [7] Department of Pathology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [8] Department of Hepatobiliary Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Pierre Decazes, Centre Henri Becquerel Rouen, France
Zhongxiang Ding, Zhejiang University, China

***Correspondence:**
Qiqiang Zeng
zengqiqiangwz@163.com
Xiangwu Zheng
XiangwuZheng3@163.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 24 August 2020
Accepted: 11 November 2020
Published: 07 January 2021

Citation:
Xue B, Wu S, Zheng M, Jiang H, Chen J, Jiang Z, Tian T, Tu Y, Zhao H, Shen X, Ramen K, Wu X, Zhang Q, Zeng Q and Zheng X (2021) Development and Validation of a Radiomic-Based Model for Prediction of Intrahepatic Cholangiocarcinoma in Patients With Intrahepatic Lithiasis Complicated by Imagologically Diagnosed Mass. Front. Oncol. 10:598253. doi: 10.3389/fonc.2020.598253

Background: This study was conducted with the intent to develop and validate a radiomic model capable of predicting intrahepatic cholangiocarcinoma (ICC) in patients with intrahepatic lithiasis (IHL) complicated by imagologically diagnosed mass (IM).

Methods: A radiomic model was developed in a training cohort of 96 patients with IHL-IM from January 2005 to July 2019. Radiomic characteristics were obtained from arterial-phase computed tomography (CT) scans. The radiomic score (rad-score), based on radiomic features, was built by logistic regression after using the least absolute shrinkage and selection operator (LASSO) method. The rad-score and other independent predictors were incorporated into a novel comprehensive model. The performance of the Model was determined by its discrimination, calibration, and clinical usefulness. This model was externally validated in 35 consecutive patients.

Results: The rad-score was able to discriminate ICC from IHL in both the training group (AUC 0.829, sensitivity 0.868, specificity 0.635, and accuracy 0.723) and the validation group (AUC 0.879, sensitivity 0.824, specificity 0.778, and accuracy 0.800). Furthermore, the comprehensive model that combined rad-score and clinical features was great in predicting IHL-ICC (AUC 0.902, sensitivity 0.771, specificity 0.923, and accuracy 0.862).

Conclusions: The radiomic-based model holds promise as a novel and accurate tool for predicting IHL-ICC, which can identify lesions in IHL timely for hepatectomy or avoid unnecessary surgical resection.

Keywords: intrahepatic cholangiocarcinoma, intrahepatic lithiasis, radiomics, risk factors, nomogram

INTRODUCTION

Intrahepatic cholangiocarcinoma (ICC) is the second most prevalent liver malignancy following hepatocellular carcinoma, and its global disease incidence is increasing (1, 2). The risk factors for ICC are complex, but recently intrahepatic lithiasis (IHL) has been confirmed as a strong risk factor. High Odds ratios (ORs) have been found for developing ICC due to hepatolithiasis, up to 50 in Korea (3), six in China (4), and seven in Italy (5). Studies have reported that about 2.3 to 13.0% of patients with hepatolithiasis end up developing cholangiocarcinoma (6–11), and 65–70% of patients in Taiwan who underwent resection for cholangiocarcinoma suffer from concomitant hepatolithiasis (12, 13).

It is very difficult for a clinical surgeon to identify ICC early in patients with IHL because there are no specific symptoms and radiological features. Although tissue biopsy can be used to confirm a histological diagnosis, it is not routinely recommended in ICC (14), especially in IHL-ICC where 'negative' biopsy results do not exclude ICC given the significant potential for sampling error. The preoperative diagnosis for IHL-ICC is mainly obtained from a combination of imaging, serum carcinoembryonic antigen (CEA), and cancer antigen 19-9 (CA 19-9). However, the current diagnostic accuracy of IHL-ICC is low, generally ranging from 30 to 65% (7, 10, 11, 15, 16). Recently, we have increased the diagnostic accuracy to 78.5% through developing a nomogram for patients with IHL complicated by imagologically diagnosed mass (17). Despite this improvement, the accuracy of preoperative imaging diagnosis in the nomogram was still low because it was performed by two radiologists based on their experience. In recent years, radiomics has been introduced in clinic to identify liver tumors (18); however, no studies have used the radiomic approach for diagnosing IHL-ICC. Therefore, there is an urgent need to develop a radiomic model capable of improving the diagnostic accuracy of IHL-ICC.

In this study, we aimed to identify the radiomic features of IHL-ICC, develop a predictive model that combined radiomic score (rad-score) and clinical features for preoperative identification of ICC among patients with IHL, and also to validate its predictive capacity in an independent data sets.

PATIENTS AND METHODS

Patients Selection

All patients involved in this retrospective study that constituted the training cohort were diagnosed with intrahepatic lithiasis (IHL) complicated by imagologically diagnosed mass (IM) (IHL-IM) and underwent hepatectomy at The First Affiliated Hospital of Wenzhou Medical University (WMU) from January 2005 to July 2019. The database from our hospital was screened meticulously to select the potentially eligible patients who were; (1) with pathological diagnosis of ICC or IHL and (2) with available high-quality contrast-enhanced computed tomography (CT) before surgical resection. The clinical characteristics of these qualified patients were recorded. This retrospective study was reviewed and approved by the Institutional Review Board (IRB) of the First Affiliated Hospital of WMU, and a waiver of written informed consent was granted by the IRB due to the retrospective nature of this study in which de-identified data were used and analyzed.

The patients for the external validation cohort were selected from the Second Affiliated Hospital of WMU, whose IRB approved the validation study.

Details for the recruitment and selection criteria of the patients included in this study were shown in **Figure 1**.

CT Image Acquirement, Tumor Segmentation, and Radiomic Feature Extraction

All patients were assessed with contrast-enhanced CT using the LifeX software tools (19). Two radiologists (BX and SW) who were blinded to the pathologic details, reviewed transverse CT

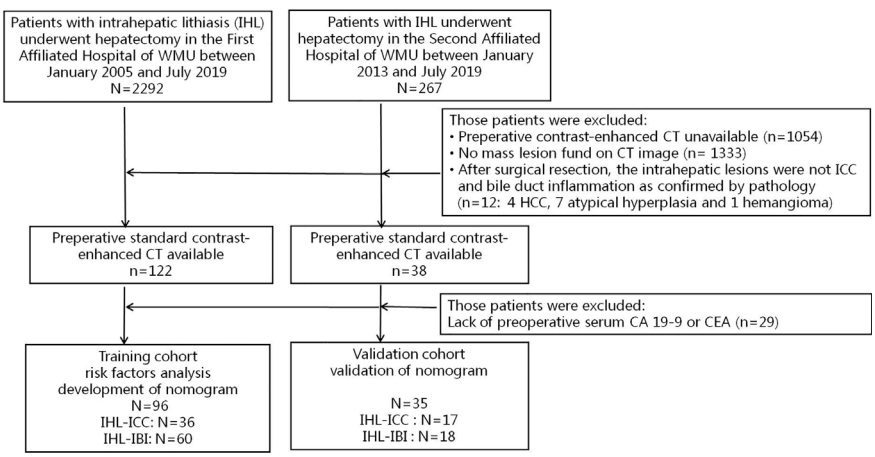


FIGURE 1 | Proceeding flow of enrollment.

images to determine respectively the features of the mass location and boundary.

The radiomic workflow is depicted in **Figure 2**. Image feature extraction was performed on retrieved arterial phase CT images (5 mm thickness). The pre-processing procedure [*i.e.*, the uniform of window width (200 Hu), window level (45 Hu), and pixel size (512 × 512)] was undertaken before feature extraction. Manual segmentation of tumor regions of interest (ROI) was carried out by two different radiologists (BX and SW). Each transverse slice consisted of cuts made along the primary tumor contour. A total of fifty-two quantified texture features were extracted, including features from histogram-based matrix and shape-based matrix from the first order and features from gray-level co-occurrence matrix (GLCM), gray-level zone length matrix (GLZLM), neighborhood gray-level dependence matrix (NGLDM), and gray-level run length matrix (GLRLM) from second or higher order (20). A detailed description of all these characteristics can be found in https://www.lifexsoft.org/index.php/resources/19-texture/radiomic-features. All original data about extracted features are displayed in the **Supplementary Material 1** and **Supplementary Material 2**.

Radiomic Feature Selection and Signature Construction

We devised a two-step procedure for dimensionality reduction and selection of robust features. Firstly, we calculated the intraobserver and interobserver reliability for each ROI based radiomic feature, extracted from 50 randomly chosen images. To assess interobserver reliability, the ROI segmentation was performed by two experts [one radiologist (reader 1, BX) and one hepatobiliary surgeon (reader 2, QZ)] who were blinded to both the clinical and pathologic details. To evaluate intraobserver reliability, reader 1 repeated the ROI segmentation and feature extraction procedure twice over a period of one month. The reliability was calculated by using intraclass correlation coefficient. Radiomic features with both intraobserver and interobserver intraclass correlation coefficient values greater than 0.55 (demonstrating at least moderate stability) were selected for subsequent investigation. Secondly, the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm was applied to the training cohort in order to determine which ICC-related features had non-zero coefficients while being cross-validated 10 times by the penalty parameter. A radiomic signature was generated *via* a linear combination of selected features weighted by their respective coefficients (21).

Development, Performance, and Validation of a Radiomic Nomogram

A radiomic model incorporating the radiomic signature, as well as independent risk factors that were obtained in our previous research for IHL-ICC (17), was constructed based on the results of the multivariate logistic regression analysis performed on the training cohort. A radiomic nomogram was then constructed in order to provide clinicians with a visual tool through the use of the selected covariates. Furthermore, a clinical model based on multivariate logistic regression analysis of candidate predictors, with the exception of radiomic signature, was developed. We calculated the area under the curve (AUC) of the receiver operating characteristic curve (ROC) to measure the discrimination performance of established models, and through the use of the DeLong algorithm (22), we compared the differences in AUC estimates between the various models. Calibration curves were graphed, through bootstrapping (resampled 1,000 times), to evaluate the predictive accuracy of the radiomic nomogram,

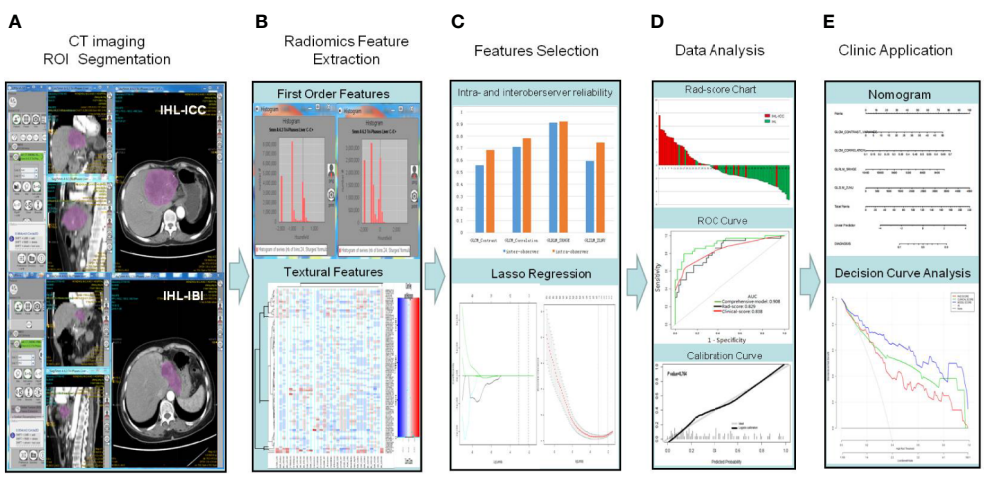


FIGURE 2 | Workflow of required steps in this current study. **(A)** Manual segmentation on arterial phase CT scans; **(B)** Quantification of tumor intensity, shape, and texture through radiomic features collected by LIFEx software from inside the defined tumor contours on CT images. **(C)** For feature selection, two successive steps are the reliability assessment regarding the extracted features, followed by the LASSO method. A radiomic signature was obtained by combining the selected features by their respective coefficients, linearly. **(D)** By measuring the area under a receiver operating characteristic (ROC) curve and the calibration curve, the performance of the prediction model can be analyzed. **(E)** A radiomic nomogram was built in order to provide clinicians with a visual tool through the use of the selected covariates, followed by decision curve.

followed by a Hosmer–Lemeshow test (P > 0.05 indicating good fit) (23). The performance of the radiomic model was then externally tested through an independent validation cohort.

Clinical Utility of the Radiomic Nomogram

The net benefits at different threshold probabilities were quantified by a decision curve analysis (DCA) (24), thereby estimating the clinical utility of the established models in the validation cohort.

Statistical Analysis

Numerical variables were compared by means of the *t*-test or Mann–Whitney U test, and categorical variables were compared using the $\chi 2$ test or Fisher's exact test, where appropriate. Univariate and multivariate Cox regression analyses were performed to determine predictors of IHL-ICC. All variables with a *p*-value <0.05 in univariate analysis were selected for multivariate analysis. Statistical analyses were performed with the R software (version 3.4.4, http://www.R-project.org), the EmpowerStats software (www.empowerstats.com, X&Y solutions, Inc. Boston MA). The R package "glmnet" was used to perform LASSO binary logistic regression analysis; the "rms" package, to create the nomogram; and the "pROC" package, to analyze ROC curves. A two-sided *p*-value <0.05 was considered statistically significant.

RESULTS

Demographic and Clinicopathological Characteristics

A total of 96 eligible patients were selected from the training cohort. Thirty-six of them were diagnosed with IHL-ICC, and 60 patients were diagnosed as IHL with intrahepatic biliary inflammation (IHL-IBI). Furthermore, 35 patients (17 IHL-ICC and 18 IHL-IBI) were included for validation. The detailed characteristics of the patients were summarized in **Table 1**. There were no significant differences regarding clinical and radiologic characteristics, in both the training and validation cohorts.

Feature Selection and Radiomic Signature Construction

Of 52 extracted radiomic features, four ICC-related features with non-zero coefficients in the LASSO logistic regression model were obtained from the training cohort. The radiomic score used to calculate the novel radiomic signature was obtained by means of the following formula: rad-score = 9.79113 + 0.06519 * GLCM_CONTRAST_VARIANCE + 5.97425*GLCM_CORRELATION-0.00151*GLRLM_SRHGE + 0.00098*GLZLM_ZLNU (**Figure 3A**).

Diagnostic Validation of Radiomic Signature and Clinical Prediction Models

The radiomic signature model exhibited promising discriminative ability for IHL-ICC and IHL-IBI in the training cohort. The AUC of the radiomic signature model was 0.829 [95% confidence interval (CI): 0.744, 0.910] with sensitivity 0.868, specificity 0.635, and accuracy 0.723 in training cohort (**Figure 3C**). Furthermore, by combining three independent factors (fever, CEA, and CA 19-9) in the training cohort, a clinical prediction model was constructed. The AUC of the nomogram for the clinical prediction model was 0.838 (95% CI, 0.747–0.928), with a sensitivity, specificity, and accuracy of 0.902, 0.647, 0.800 respectively (**Figure 3C**).

In the validation cohort, AUC of the radiomic signature model was 0.824 (95% CI: 0.768, 0.989) with sensitivity 0.824, specificity 0.778, and accuracy 0.800. The AUC of the nomogram for the clinical prediction model was 0.824 (95% CI, 0.681–0.966), with a sensitivity, specificity, and accuracy of 0.824, 0.722, 0.771 respectively (**Figure 3D**).

Development, Performance, and Validation of Prediction Models

A comprehensive model incorporating two kinds of independent predictors (radiomic signature and clinical features) was developed, by using the following formula: comprehensive model = −0.87516 + 0.84946*rad-score −1.02770*1 (if with fever) + 1.11976*2 (if 5μg/L≥CEA≥3.75μg/L) + 2.41799*(if CEA > 5μg/L) + 0.64579*(if 143.15 U/ml≥CA 19-9≥37 U/ml) + 1.56721*(if CA 19-9>143.15 U/ml), and presented as a nomogram (**Figure 3B**). The model is capable of indicating a good fit, as proved the Hosmer–Lemeshow test (p = 0.764), and the calibration of the nomogram was likewise well-calibrated, as illustrated in **Figure 2D**. In the training cohort, the comprehensive model displayed the highest discrimination between IHL-ICC and IHL-IBI with an AUC of 0.908 (95% CI: 0.833, 0.970) (sensitivity 0.771, specificity 0.923, and accuracy 0.862); the detected AUC value was higher than that of the radiomic signature model (AUC, 0.829; $p < 0.05$) and clinical prediction model and (AUC, 0.838; $p < 0.05$) (**Figure 3C**). In the validation cohort, the comprehensive model presented the greatest AUC (0.879; 95% CI: 0.768, 0.990) as well, which confirms that the comprehensive model is capable of better predictive efficacy than either the radiomic signature model (AUC, 0.824; $p < 0.05$) or clinical prediction model alone (AUC, 0.755; $p < 0.05$) (**Figure 3D**).

Clinical Use

The DCA for the radiomic nomogram, the clinical prediction model, and the comprehensive model are presented in **Figure 4**. The comprehensive model is capable of providing a better net benefit when predicting ICC in IHL patients, when compared with the other two models (demonstrated by the threshold probabilities of more than 10%), and particularly, in situations where there is no alternative prediction model available.

DISCUSSION

The accurate diagnosis for IHL patients with ICC is extremely important because it can facilitate the decision making with regard to surgical treatment at an early stage. The present work is the first attempt to propose a comprehensive model combined with radiomic and clinical signatures that can improve the

TABLE 1 | Demographic and clinical characteristics of the study population.

	Training cohort			Validation cohort		
	IHL-IBI	**IHL-ICC**	***P* value**	**IHL-IBI**	**IHL-ICC**	***P* value**
Demographic or Characteristic	(n = 60)	(n = 36)		(n = 18)	(n = 17)	
Age, mean (SD)			0.01			0.73
<60 y	29 (48.33%)	8 (22.22%)		8 (44.44%)	6 (35.29%)	
≥60 y	31 (51.67%)	28 (77.78%)		10 (55.56%)	11 (64.71%)	
Sex (F/M)	37/23	22/14	0.96	11/7	10/6	1
Smoking	7 (11.67%)	6 (11.67%)	1	2 (11.11%)	2 (11.76%)	1
Alcohol	8 (13.33%)	9 (25%)	0.15	3 (16.67%)	4 (23.53%)	1
Personal cancer history	2 (3.33%)	2 (5.56%)	1	1 (5.56%)	0	1
Family cancer history	1 (1.67%)	0 (0%)	1	1 (5.56%)	1 (5.88%)	1
Inflammatory attacks within half a year (≥2 times)	12 (20.00%)	9 (25.00%)	0.06	4 (22.22%)	4 (23.53%)	1
Lesion size (cm), mean (SD)	5.42 (1.88)	5.79 (1.63)	0.48	5.27 (2.45)	5.98 (1.66)	0.35
Location of hepatolithiasis						
Left lobe	40 (66.67%)	26 (72.22%)	0.57	11 (61.11%)	13 (76.47%)	0.47
Right lobe	15 (25.00%)	8 (22.22%)	0.76	4 (22.22%)	3 (17.65%)	1
Left and right lobes	3 (5.00%)	2 (5.56%)	1	2 (11.11%)	1 (5.88%)	1
Lobus caudatus	2(3.33%)	0	1	1(5.56%)	0	1
Symptoms						
Abdominal pain	53 (88.30%)	30 (83.33%)	0.7	14 (77.78%)	12 (70.59%)	0.71
Fever	31 (51.67%)	12 (33.3%)	0.08	4 (22.22%)	3 (17.65%)	1
Vomiting	20 (33.30%)	6 (16.67%)	0.08	6 (33.33%)	2 (11.76%)	0.22
Jaundice	8 (13.30%)	3 (8.30%)	0.68	2 (11.11%)	2 (11.76%)	1
Weight loss	1 (1.67%)	1 (2.78%)	0.61	0	1 (5.88%)	1
Laboratory						
ALK (U/L), mean (SD)	218.70 (224.42)	226.94 (183.96)	0.17	166.52 (263.94)	192.84 (218.13)	0.38
γ-GT (U/L), mean (SD)	218.80 (261.40)	221.56 (188.88)	0.31	269.71 (332.59)	177.41 (291.32)	0.30
ALT (U/L), mean (SD)	75.75 (80.52)	55.89 (76.81)	0.11	83.22 (141.27)	47.77 (52.48)	0.28
Albumin (g/dl), mean (SD)	35.37(5.65)	35.54 (4.63)	0.9	38.19 (4.26)	36.66 (5.33)	0.55
PT (second), mean (SD)	14.18 (1.65)	15.50 (10.53)	0.42	14.84 (1.37)	13.49 (0.77)	0.62
CA 19-9 (U/ml), median (IQR)	42.35 (11.18, 407.42)	902.8 (28.6, 2020.80)	<0.01	16.26 (7.08, 130.80)	92.96 (4.2, 1200)	<0.01
CEA (μg/L), median (IQR)	1.80 (1.20, 2.30)	5.50 (2.10, 35.10)	<0.01	2.21 (1.2, 3.11)	5.15 (2.13, 31.88)	<0.01
AFP (μg/L), median (IQR)	2.48 (1.70, 3.66)	3.25 (2.10, 4.21)	0.07	2.64 (1.68, 3.66)	2.55 (1.78, 3.42)	0.87
CA 125 (U/ml), mean (SD)	12.37 (0.99)	1399.36 (3440.36)	0.09	28.58 (139.99)	122.71 (165.13)	0.01
Complication						
HBsAg+	4 (6.67%)	4 (11.11%)	0.65	2 (11.11%)	1 (5.88%)	1
HBcAb+	16 (26.67%)	8 (22.22%)	0.84	10 (55.56%)	10 (58.82%)	1
Diabetes	7 (11.67%)	2 (5.56%)	0.53	1 (5.56%)	0	1
Clinical Score, mean (SD)	−1.31 (1.05)	0.94 (1.97)	<0.01	−1.34 (1.02)	0.87 (1.97)	<0.01
Radiomic Score, mean (SD)	−1.32 (1.23)	0.54 (1.62)	<0.01	−0.86 (1.18)	0.55 (1.82)	<0.01

current diagnostic accuracy standard of ICC in patients with IHL. The prediction model was validated internally and externally.

In a recent study, we had developed a nomogram to predict ICC for patients with IHL complicated by the presence of a imagologically diagnosed mass (17). However, the imagological diagnosis in the nomogram was made by radiologists. Even for experienced radiologists, the accuracy of diagnosis is still lower than 70% (10, 11, 15, 17). Detection of ICC in IHL is mainly dependent on imaging modalities because there are no specific symptoms in cases of IHL-ICC other than the clinical manifestation of hepatolithiasis.

ICC can be according to three types of morphological characteristics: mass-forming, periductal infiltrating, and intraductal growth. Of these three, mass-forming is the most common type and on CT scan, usually resembles a homogeneous low-attenuation mass with irregular peripheral enhancement, often accompanied by capsular retraction, peripheral intrahepatic duct dilation, and satellite nodules. If the dysplastic bile duct presents growth without mass formation, then it possesses the characteristics of a periductal infiltrating cholangiocarcinoma. Diffuse periductal thickening and increased enhancement can be observed in a dilated or irregularly narrowed intrahepatic duct. For patients with IHL, inflammatory pseudo-tumors or liver abscesses often occur at the site of intrahepatic stones, thus making it difficult to distinguish from mass-forming ICC, whereas proliferative cholangitis or inflammatory stenosis are difficult to distinguish from periductal infiltrating ICC. Furthermore, after long-term chronic inflammation, liver segments often become scarred and undergo fibrotic change (11), making IHL even more difficult to distinguish from ICC on imaging.

The radiomic technique can process high-throughput extraction of quantitative features that result in the conversion of images into mineable data and the subsequent analysis of these data for decision support, which draws a contrast with the traditional treatment of medical images as simple tools of visual interpretation. Radiomic data contain first-, second-, and higher-order statistics. The radiomic technique is very useful for IHL-ICC, which is highly heterogeneous and short of traditional imaging features. We used the LIFEx (A Freeware for Radiomic Feature Calculation) to implement these functions of ROI segmentation and radiomic feature extraction in a

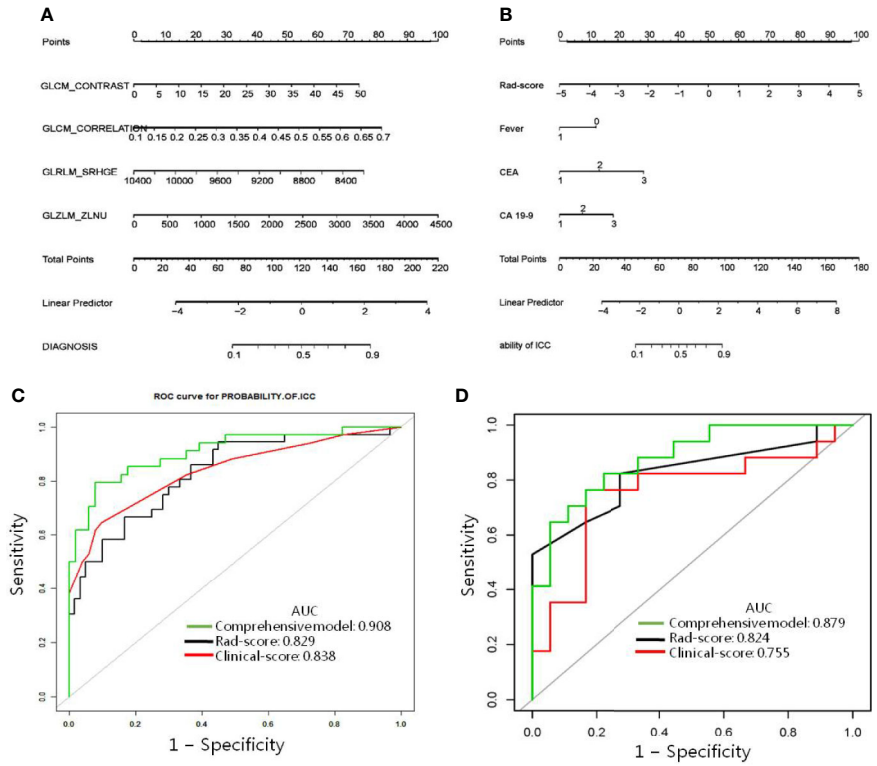


FIGURE 3 | Radiomic nomogram designed with receiver operating characteristic curves. **(A)** The radiomic nomogram and **(B)** the comprehensive model was developed in the training cohort for predicting IHL-ICC. Comparison of ROC among the radiomic nomogram, clinical model, and comprehensive model for the prediction of IHL-ICC in the **(C)** training and **(D)** validation cohorts.

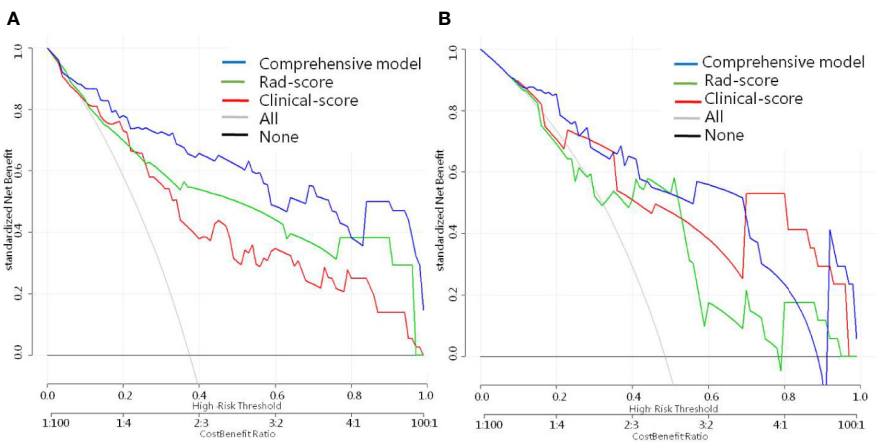


FIGURE 4 | Decision curve analysis for every model in **(A)** the training and **(B)** validation dataset. The net benefit is measured by the y-axis, which is calculated by summing the benefits (true-positive findings) and deducting the harms (false-positive findings), while weighting the harms associated to the relative damage of undetected IHL-ICC when compared with the damage of being mistakenly diagnosed with HL-ICC.

one-stop manner. The LASSO logistic regression algorithm can effectively solve the problem of multicollinearity among numerous extracted features and find meaningful feature parameters for a well constructed prediction model. In the present research, we got the higher-order radiomic features of IHL-ICC including GLCM_CONTRAST, GLCM_CORRELATION, GLRLM_SRHGE, and GLZLM_ZLNU that were obviously different from IHL-IBI. Finally, the radiomic model has improved the diagnostic accuracy for IHL-ICC to 0.72 which is higher than in our previous research (17) and others (10, 11, 15).

Furthermore, a comprehensive model incorporating two kinds of independent predictors (radiomic signature and clinical features) was developed for further improving the diagnostic accuracy for IHL-ICC. Based on our previous research, the clinical risk factor for IHL-ICC included biliary tract surgical history, fever, ascites, CA 19-9, and CEA. Here, we removed indicators that need to be judged subjectively, such as vomiting, and retained objective indicators including fever, CA 19-9, and CEA. The comprehensive model further improves the diagnostic accuracy to 86%, which is simpler and more convenient. As a non-invasive method, the comprehensive model for IHL-ICC would be a convenient application for clinicians.

There are several limitations to the present study. First, due to retrospective design and small sample, the potential selection bias cannot be excluded, which limits the accuracy and reliability of results. Second, when highlighting the outline of ROI areas, the variation between observed images should be deliberated. The inclusion of a computer-aided software, as used in this study, may help to reduce variation to some degree. Third, the texture features mined in this study were based solely on arterial phase CT images. Further investigation is needed to evaluate the performance of using either portal venous- or delayed-phase imaging or in combination, for predicting the malignant potential of IHL-IM. Furthermore, there are many different types of texture features and imaging processing software, so unifying the texture analysis would undoubtedly add rigor to the results obtained while spreading the application of this technology. Therefore, more investigation attempts are necessary for better estimation, especially large-scale prospective, and multicenter studies.

CONCLUSIONS

A prediction nomogram based on CT radiomics was created and validated in this study. It was suitably utilized in order to simplify the individualized prediction of malignancy in IHL-IM patients. The radiomic-based model holds promise as a novel and accurate tool for predicting IHL-ICC, which can identify lesions in IHL, in a timely fashion, determining if there is a need for hepatectomy, avoiding unnecessary surgical resection.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) of the First Affiliated Hospital of WMU. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.598253/full#supplementary-material

REFERENCES

1. Rizvi S, Khan SA, Hallemeier CL, Kelley RK, Gores GJ. Cholangiocarcinoma - evolving concepts and therapeutic strategies. *Nat Rev Clin Oncol* (2018) 15 (2):95–111. doi: 10.1038/nrclinonc.2017.157
2. Wood R, Brewster DH, Fraser LA, Brown H, Hayes PC, Garden OJ. Do increases in mortality from intrahepatic cholangiocarcinoma reflect a genuine increase in risk? Insights from cancer registry data in Scotland. *Eur J Cancer* (2003) 39(14):2087–92. doi: 10.1016/S0959-8049(03)00544-6
3. Lee TY, Lee SS, Jung SW, Jeon SH, Yun SC, Oh HC, et al. Hepatitis B virus infection and intrahepatic cholangiocarcinoma in Korea: a case-control study. *Am J Gastroenterol* (2008) 103(7):1716–20. doi: 10.1111/j.1572-0241.2008.01796.x
4. Zhou YM, Yin ZF, Yang JM, Li B, Shao WY, Xu F, et al. Risk factors for intrahepatic cholangiocarcinoma: a case-control study in China. *World J Gastroenterol* (2008) 14(4):632–5. doi: 10.3748/wjg.14.632
5. Donato F, Gelatti U, Tagger A, Favret M, Ribero ML, Callea F, et al. Intrahepatic cholangiocarcinoma and hepatitis C and B virus infection, alcohol intake, and hepatolithiasis: a case-control study in Italy. *Cancer Causes Control* (2001) 12(10):959–64. doi: 10.1023/A:1013747228572
6. Kim HJ, Kim JS, Suh SJ, Lee BJ, Park JJ, Lee HS, et al. Cholangiocarcinoma Risk as Long-term Outcome After Hepatic Resection in the Hepatolithiasis Patients. *World J Surg* (2015) 39(6):1537–42. doi: 10.1007/s00268-015-2965-0
7. Uenishi T, Hamba H, Takemura S, Oba K, Ogawa M, Yamamoto T, et al. Outcomes of hepatic resection for hepatolithiasis. *Am J Surg* (2009) 198 (2):199–202. doi: 10.1016/j.amjsurg.2008.08.020
8. Chen MF, Jan YY, Wang CS, Jeng LB, Hwang TL, Chen SC. Intrahepatic stones associated with cholangiocarcinoma. *Am J Gastroenterol* (1989) 84(4):391–5.
9. Suzuki Y, Mori T, Yokoyama M, Nakazato T, Abe N, Nakanuma Y, et al. Hepatolithiasis: analysis of Japanese nationwide surveys over a period of 40 years. *J Hepatobiliary Pancreat Sci* (2014) 21(9):617–22. doi: 10.1002/jhbp.116
10. Su CH, Shyr YM, Lui WY, P'Eng FK. Hepatolithiasis associated with cholangiocarcinoma. *Br J Surg* (1997) 84(7):969–73. doi: 10.1002/bjs.1800840717
11. Kubo S, Kinoshita H, Hirohashi K, Hamba H. Hepatolithiasis associated with cholangiocarcinoma. *World J Surg* (1995) 19(4):637–41. doi: 10.1007/BF00294744
12. Chen MF. Peripheral cholangiocarcinoma (cholangiocellular carcinoma): clinical features, diagnosis and treatment. *J Gastroenterol Hepatol* (1999) 14 (12):1144–9. doi: 10.1046/j.1440-1746.1999.01983.x
13. Chen MF, Jan YY, Jeng LB, Hwang TL, Wang CS, Chen SC, et al. Intrahepatic cholangiocarcinoma in Taiwan. *J Hepatobiliary Pancreat Surg* (1999) 6 (2):136–41. doi: 10.1007/s005340050096
14. Weber SM, Ribero D, O'Reilly EM, Kokudo N, Miyazaki M, Pawlik TM. Intrahepatic cholangiocarcinoma: expert consensus statement. *HPB (Oxford)* (2015) 17(8):669–80. doi: 10.1111/hpb.12441

15. Guglielmi A, Ruzzenente A, Valdegamberi A, Bagante F, Conci S, Pinna AD, et al. Hepatolithiasis-associated cholangiocarcinoma: results from a multi-institutional national database on a case series of 23 patients. *Eur J Surg Oncol* (2014) 40(5):567–75. doi: 10.1016/j.ejso.2013.12.006
16. Kim YT, Byun JS, Kim J, Jang YH, Lee WJ, Ryu JK, et al. Factors predicting concurrent cholangiocarcinomas associated with hepatolithiasis. *Hepatogastroenterology* (2003) 50(49):8–12.
17. Chen G, Yu H, Wang Y, Li C, Zhou M, Yu Z, et al. A novel nomogram for the prediction of intrahepatic cholangiocarcinoma in patients with intrahepatic lithiasis complicated by imagiologically diagnosed mass. *Cancer Manag Res* (2018) 10:847–56. doi: 10.2147/CMAR.S157506
18. Lewis S, Peti S, Hectors SJ, King M, Rosen A, Kamath A, et al. Volumetric quantitative histogram analysis using diffusion-weighted magnetic resonance imaging to differentiate HCC from other primary liver cancers. *Abdom Radiol (NY)* (2019) 44(3):912–22. doi: 10.1007/s00261-019-01906-7
19. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res* (2018) 78(16):4786–9. doi: 10.1158/0008-5472.CAN-18-0125
20. Chen L, Wang H, Zeng H, Zhang Y, Ma X. Evaluation of CT-based radiomics signature and nomogram as prognostic markers in patients with laryngeal squamous cell carcinoma. *Cancer Imag* (2020) 20(1):28. doi: 10.1186/s40644-020-00310-5
21. Ji GW, Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, et al. Biliary Tract Cancer at CT: A Radiomics-based Model to Predict Lymph Node Metastasis and Survival Outcomes. *Radiology* (2019) 290(1):90–8. doi: 10.1148/radiol.2018181408
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* (1988) 44(3):837–45. doi: 10.2307/2531595
23. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: *Applied Logistic Regression*, *2nd ed*. New York: Wiley (2005).
24. Ewout W, Steyerberg AJV. Decision curve analysis: a discussion. *Med Decis Making* (2008) 28(1):146–9. doi: 10.1177/0272989X07312725

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 12 January 2021
doi: 10.3389/fonc.2020.599050

Hormonal Receptor Immunochemistry Heterogeneity and ^{18}F-FDG Metabolic Heterogeneity: Preliminary Results of Their Relationship and Prognostic Value in Luminal Non-Metastatic Breast Cancers

Nicolas Aide[1,2†], Nicolas Elie[3†], Cécile Blanc-Fournier[4], Christelle Levy[5], Thibault Salomon[6] and Charline Lasnon[2,7]*

[1] Nuclear Medicine Department, University Hospital, Caen, France, [2] INSERM 1086 ANTICIPE, Normandy University, Caen, France, [3] Université de Caen Normandie, UNICAEN, SF 4206 ICORE, CMABIO3, Caen, France, [4] Pathology Department, François Baclesse Cancer Centre, Caen, France, [5] Breast Cancer Unit, François Baclesse Cancer Centre, Caen, France, [6] Nuclear Medicine Department, Hospital Centre, Versailles, France, [7] Nuclear Medicine Department, François Baclesse Cancer Centre, Caen, France

OPEN ACCESS

Edited by:
Laurent Dercle, Columbia University Irving Medical Center, United States

Reviewed by:
Laurence Champion, Hôpital René Huguenin, France
Antoine Girard, Centre Eugène Marquis, France

***Correspondence:**
Charline Lasnon
c.lasnon@baclesse.unicancer.fr

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 26 August 2020
Accepted: 12 November 2020
Published: 12 January 2021

Citation:
Aide N, Elie N, Blanc-Fournier C, Levy C, Salomon T and Lasnon C (2021) Hormonal Receptor Immunochemistry Heterogeneity and ^{18}F-FDG Metabolic Heterogeneity: Preliminary Results of Their Relationship and Prognostic Value in Luminal Non-Metastatic Breast Cancers. Front. Oncol. 10:599050. doi: 10.3389/fonc.2020.599050

Introduction: We aimed to investigate whether ^{18}F-FDG PET metabolic heterogeneity reflects the heterogeneity of estrogen receptor (ER) and progesterone receptor (PR) expressions within luminal non-metastatic breast tumors and if it could help in identifying patients with worst event-free survival (EFS).

Materials and methods: On 38 PET high-resolution breast bed positions, a single physician drew volumes of interest encompassing the breast tumors to extract SUV_{max}, histogram parameters and textural features. High-resolution immunochemistry (IHC) scans were analyzed to extract Haralick parameters and descriptors of the distribution shape. Correlation between IHC and PET parameters were explored using Spearman tests. Variables of interest to predict the EFS status at 8 years (EFS-8y) were sought by means of a random forest classification. EFS-8y analyses were then performed using univariable Kaplan-Meier analyses and Cox regression analysis. When appropriate, Mann-Whitney tests and Spearman correlations were used to explore the relationship between clinical data and tumoral PET heterogeneity variables.

Results: For ER expression, correlations were mainly observed with ^{18}F-FDG histogram parameters, whereas for PR expression correlations were mainly observed with gray-level co-occurrence matrix (GLCM) parameters. The strongest correlations were observed between $skewness_{\_ER}$ and $uniformity_{\_HISTO}$ ($\rho = -0.386$, $p = 0.017$) and $correlation_{\_PR}$ and $entropy_{\_GLCM}$ ($\rho = 0.540$, $p = 0.001$), respectively. The median follow-up was 6.5 years and the 8y-EFS was 71.0%. Random forest classification found age, clinical stage,

SUV_{max}, $skewness_{\_ER}$, $kurtosis_{\_ER}$, $entropy_{\_HISTO}$, and $uniformity_{\_HISTO}$ to be variables of importance to predict the 8y-EFS. Univariable Kaplan-Meier survival analyses showed that $skewness_{\_ER}$ was a predictor of 8y-EFS (66.7 ± 27.2 versus 19.1 ± 15.2, p = 0.018 with a cut-off value set to 0.163) whereas other IHC and PET parameters were not. On multivariable analysis including age, clinical stage and $skewness_{\_ER}$, none of the parameters were independent predictors. Indeed, $skewness_{\_ER}$ was significantly higher in youngest patients (ρ = –0.351, p = 0.031) and in clinical stage III tumors (p = 0.023).

Conclusion: A heterogeneous distribution of ER within the tumor in IHC appeared as an EFS-8y prognosticator in luminal non-metastatic breast cancers. Interestingly, it appeared to be correlated with PET histogram parameters which could therefore become potential non-invasive prognosticator tools, provided these results are confirmed by further larger and prospective studies.

Keywords: breast cancer, steroid receptors, image processing, computer-aided system, radiomics analysis; ^{18}F-FDG PET imaging

INTRODUCTION

Breast cancer is the most frequently diagnosed cancer in women (16% of all women's cancers) in all world regions.[1] Its incidence is rising as a result of longer life expectancy and changes in risk factors. Breast cancer treatment recommendations are based on histological subtype (ER-positive, HER-2 positive, or triple negative tumors), tumor grade, and stage of the disease. More recently, with the development of DNA microarray gene expression analysis, a molecular classification has been proposed and validated (1–3). However, its clinical use is limited, since these techniques are currently expensive as compared to conventional immunohistochemistry (IHC). An attempt to replicate molecular classification using conventional IHC characteristics of the tumor, including ER, PR, HER-2, and Ki67 showed low concordance with gene expressions profile (4, 5). When it comes to breast cancer staging, 2-deoxy-2[18F]-fluoro-D-glucose (^{18}F-FDG) PET/CT is a well-established examination for the initial staging of locally advanced breast cancer (6–9), as it displays excellent capabilities for extra-axillary nodal and distance metastases detection. On the contrary, for the local evaluation of primary breast lesion, ^{18}F-FDG PET/CT has so far been outperformed by echography and MRI mainly because of its lack of sensitivity (10, 11). However, with the newly growing development of metabolic heterogeneity features in nuclear medicine, the PET community is regaining interest in the value of ^{18}F-FDG PET/CT for the non-invasive biological characterization of primary breast tumors. Until now, PET radiomics have always been confronted with the expression of ER, PR, HER2, and Ki67 (12–15) and PET radiomics certainly seem to represent more than just a binary expression of receptors. Meanwhile, improvement in high-resolution scanning of pathological sections and digital imaging analysis is leading to the rise of digital-IHC. Even though it demands further validation and standardization, this technique can provide computation of texture and distribution parameters for hormonal receptors intra-tumoral heterogeneity (16, 17).

The objective of the present study was therefore to investigate (**i**) if PET metabolic heterogeneity features reflect the heterogeneity of ER and PR expression within luminal breast tumors and (**ii**) if PET metabolic heterogeneity features could help in non-invasively identifying patients with the worst event-free survival (EFS).

[1] https://www.who.int/topics/cancer/breastcancer/fr/index1.html.

Abbreviations: PET, Positron Emission Tomography; ER, Estrogen Receptor; PR, Progesterone Receptor; SUV, Standardized Uptake Value; IHC, Immunohistochemistry; EFS, Event Free Survival; RF, Random Forest Classification; GLCM, Gray-Level Co-occurrence Matrix; ^{18}F-FDG, ^{18}Fluorodeoxyglucose; HER-2, Human Epidermal Growth Factor Receptor-2; HR, High Resolution; VOI, Volume of Interest; NGLDM, Neighborhood Gray-Level Different Matrix; TF, Textural Feature; ROI, Region Of Interest; SD, Standard Deviation; CART, Classification And Regression Trees; OOB, Out-Of-Bag; ROC, Receiver Operating characteristic Curve; CT, Computed Tomography.

MATERIAL AND METHODS

Study Population

This study is an ancillary study to a previous monocentric and prospective one conducted in our PET unit (18). From April 2009 to June 2012, that study included newly diagnosed and histologically proven breast cancer for which surgery was indicated in first place without neo-adjuvant chemotherapy. It was approved by the Ethics Committee (CPP Nord Ouest III, reference 2009-10) and all patients gave informed and signed consent.

PET/CT Acquisitions

All ^{18}F-FDG PET/CT acquisitions were performed on a Biograph TrueV (Siemens Healthineers) before any treatment. Patients were fasted during at least 6 h. A high-resolution (HR) breast-dedicated bed position (6 min per bed position) was acquired 75 min after the radiopharmaceutical injection. Data were reconstructed using an algorithm with point spread function (PSF) modeling (HD; TrueX, Siemens Healthineers, 3 iterations,

and 21 subsets) with no post-filtering and a 512^2 matrix size leading to voxels of 1.3 × 1.3 × 1.9 mm (19).

PET-CT Analysis

Injected dose, time between injection and acquisition and capillary glycaemia were recorded to seek EANM recommendations fulfilment (20). A single observer delineated volumes of interest (VOIs) that encompassed the entire breast tumor by using a gradient-based method implemented in MIM software (MIM software, version 5.6.5). When multiple lesions were depicted, only the biggest lesion was considered. VOIs were then saved as DICOM RT structures and loaded in LifeX v5.10 software (21) (www.lifexsoft.org) to extract SUV_{max}, histogram parameters and the following TFs:

- Inverse difference, angular second moment, variance, correlation, entropy, dissimilarity from gray-level co-occurrence matrix (GLCM) that considers the arrangements of pairs of voxels
- coarseness, contrast and busyness from neighborhood gray-level different matrix (NGLDM) that corresponds to the difference of gray-level between one voxel and its 26 neighbors in 3 dimensions.

All textural features fulfilled the benchmark of the image biomarkers standardization initiative (22). Absolute resampling using 64 bins between 0 and 32 (corresponding to the maximum SUV units recorded within PET data) was used for all TFs leading to a size of bin 0.5 (23, 24).

Immunochemistry

Automated immunohistochemistry using a Ventana Bench Mark Ultra was performed on 4-µm-thick paraffin sections of tumor resection with clone SP1 Ventana for ER (pre-diluted) and clone 1E2 Ventana for PR (pre-diluted). The slides were controlled by an experienced pathologist.

Digital-Immunochemistry Computation

The ScanScope CS microscope slide scanner (Leica Biosystems) was used to digitize whole slide images of histological sections at 20 × (0.5 µm/pixel) and record them as tiled tiff images.

For each image, regions of interest (ROIs) were drawn using the ImageScope software (Leica Biosystems) in order to select only tumor tissues and remove the artifacts. The images were processed as reported in the previous study (25). Briefly, squares of 2000 pixels size corresponding to 1 mm^2 area were used in this study. The squares were generated to fit the area of the ROI. A ratio between the stained area (brown color) and the surface of tissue was computed and assigned to each square based on their coordinates. Local ratio computed for each square was ranked according to the following ten intervals: level 0 (0–10%), level 1 (>10–20%), level 2 (>20–30%), level 3 (>30–40%), level 4 (>40–50%), level 5 (>50–60%), level 6 (>60–70%), level 7 (>70–80%), level 8 (>80–90%), and level 9 (>90–100%). The ranks then formed the basis for the co-occurrence matrix used to compute Haralick texture parameters. The classical Haralick parameters (26) were computed from the normalized co-occurrence matrix: contrast, homogeneity, dissimilarity, entropy, energy, and correlation. The descriptors of the distribution shape were also computed: skewness and kurtosis.

Statistical Analysis

Quantitative data are presented as mean (standard deviation). Correlation between immunochemistry parameters and PET parameters were explored using Spearman correlation tests and matrixes. Variables of interest to predict the occurrence of an event at 8 years (EFS-8y) were sought by means of a random forest classification incorporating the following variables: age, histology, clinical stage, Elston and Ellis grade, molecular subtype classification (27), all immunochemistry parameters and all PET parameters. This analysis implemented classification and regression trees (CART, n = 100) as well as the bootstrapping aggregating (bagging) method previously proposed by Breiman (28–30). For the validation, i.e. the training accuracy, the internal check in RF itself was used, based on the prediction error using the Out-Of-Bag (OOB) estimates of classification error: the smaller the OOB error rate, the better the model is able to classify patients according to their EFS at 8 years (8y-EFS 0 and 8y-EFS 1). The importance of variables in classification was assessed by measuring the mean decrease accuracy (31) of class prediction. Variables of importance were compared between 8y-EFS 0 and 8y-EFS 1 groups using non-parametric Mann-Whitney tests. Receiving operating characteristics (ROC) analyses for 8y-EFS were then undertaken on variables identified as significantly different between groups to define optimal cut-off values based on the Youden index. Eight-year EFS analyses were finally performed using univariable Kaplan-Meier analyses, log-rank tests for comparison of survival curves and finally multivariable Cox regression analysis. The end-point used for survival analysis was the time from diagnosis until relapse or progression, unplanned retreatment, or death as a result of breast cancer. When appropriate, non-parametric Mann-Whitney tests and Spearman correlation tests were used to explore the relationship between clinical data and tumoral heterogeneity variables. Graph and statistical analysis were performed on XLSTAT Software (XLSTAT: Data Analysis and Statistical Solutions for Microsoft Excel. Addinsoft (2017)). For all statistical tests, we retained a two-tailed p value of less than 0.05 as statistically significant. Statistical process is summarized in **Figure 1**.

RESULTS

Patients and PET Characteristics

Sixty-three patients were referred for the staging of breast carcinoma from April 2009 to June 2012. Twenty-five patients were excluded from the analysis, leading to a final database of 38 patients. The causes of exclusion were as follows: PET-CT not performed prior to surgery (n = 8), metastatic tumors on initial staging (n = 4), missing data (n = 1), breast lesions not ^{18}F-FDG avid (n = 3), hormonal receptors (ER and PR) negative tumors (n = 7), IHC slide unusable (n = 1), and volume of interest too

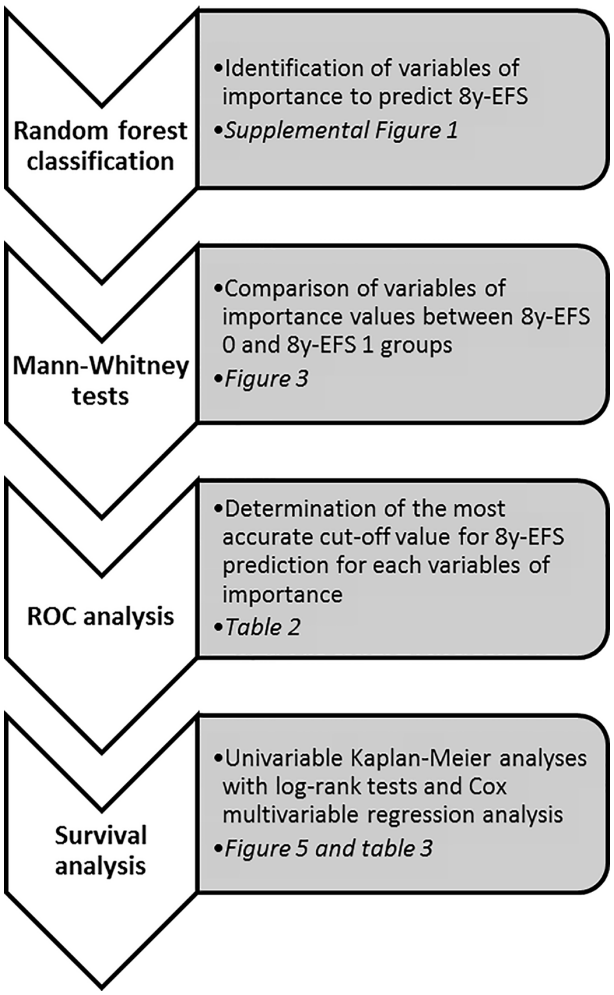


FIGURE 1 | Statistical process summary.

small to be analyzed with LifeX software (n = 1). Patient characteristics are displayed in **Table 1**. Thirty-four tumors were ER+/PR+ and 4 tumors were ER+/PR−. All patients underwent an adjuvant treatment: radiotherapy and hormonotherapy in 10 patients (26.3%) or chemotherapy, radiotherapy and hormonotherapy +/− trastuzumab in case of HER2+ tumors in 28 patients (73.7%). Mean injected dose and uptake time was 4.10 (0.56) MBq/kg and 81.6 (8.4) min, respectively.

Correlations Among Descriptors of the Distribution Shape and Haralick Texture Parameters of Estrogen and Progesterone Receptors Expression

Apart from skewness$_{_ER}$ that fairly correlated with both skewness$_{_PR}$ and kurtosis$_{_PR}$ with Spearman coefficients equal to 0.396 and 0.361 ($p = 0.015$ and $p = 0.026$), respectively, none of the ER and PR distribution descriptors or Haralick texture parameters were correlated to each other (**Figure 2A**).

Relation Between ^{18}F-FDG Textural Parameters and Intra-Tumoral Estrogen Receptors Expression

Relationship between variables can be seen in **Figure 2B**. Correlations were mainly observed with ^{18}F-FDG histogram parameters. Indeed, all PET histogram parameters were fairly correlated to kurtosis$_{_ER}$ with Spearman coefficients ranging from −0.338 to 0.410. Moreover, uniformity$_{_HISTO}$ was significantly but fairly correlated to skewness$_{_ER}$, contrast$_{_ER}$, quadratic entropy$_{_ER}$ and shannon entropy$_{_ER}$ ($\rho = -0.386$, $p = 0.017$; $\rho = 0.329$, $p = 0.044$; $\rho = 0.361$, $p = 0.027$, and $\rho = 0.333$, $p = 0.042$, respectively). Finally, entropy$_{_HISTO}$ was also fairly correlated to skewness$_{_ER}$ and quadratic entropy$_{_ER}$ ($\rho = 0.369$, $p = 0.023$; $\rho = -0.344$, $p = 0.035$, respectively).

When considering GLCM PET parameters, we observed correlations only between correlation$_{_ER}$ and both angular second moment$_{_GCLM}$ and entropy$_{_GLCM}$. Overall the PET parameter displaying the more numerous statistically significant correlations (n = 5) with intra-tumoral estrogen receptors expression was uniformity$_{_HISTO}$ with the strongest correlation being observed with skewness$_{_ER}$: $\rho = -0.386$, $p = 0.017$.

Relation Between ^{18}F-FDG Textural Parameters and Intra-Tumoral Progesterone Receptors Expression

Relationship between variables can be seen in **Figure 2C**. None of histogram PET parameters were correlated to intra-tumoral progesterone receptors expression parameters. Correlation$_{_PR}$ was the parameter displaying the maximal rate of statistically significant correlations with PET parameters (n = 7). It was fairly correlated to inverse difference$_{_GLCM}$, angular second moment$_{_GLCM}$, variance$_{_GLCM}$, entropy$_{_GLCM}$, dissimilarity$_{_GLCM}$, contrast$_{_NGLDM}$ and busyness$_{_NGLDM}$ ($\rho = -0.449$, $p = 0.005$; $\rho = -0.525$, $p = 0.001$; $\rho = 0.469$, $p = 0.003$; $\rho = 0.540$, $p = 0.001$; $\rho = 0.456$, $p = 0.004$; $\rho = 0.398$, $p = 0.014$; $\rho = -0.322$, $p = 0.049$).

Angular second moment$_{_GLCM}$ and entropy$_{_GLCM}$ were the PET parameters displaying the more numerous statistically significant correlations with intra-tumoral progesterone receptors expression. They both correlated to all IHC parameters, with the exception of contrast$_{_PR}$, homogeneity$_{_PR}$, and dissimilarity$_{_PR}$. The strongest correlation was observed between entropy$_{_GLCM}$ and correlation$_{_PR}$: $\rho = 0.540$, $p = 0.001$.

Survival Data Analysis

The statistical process for this specific part is summarized in **Figure 1**. The median follow-up was 6.5 years (range: 2.5–9.1 years) and with 11 recorded events, the 8y-EFS was 71.0% in the entire population. Among the 11 recorded events, 8 were metastatic recurrences, 2 were contralateral recurrences, and 1 was a local recurrence. The median time to recurrence from the date of diagnosis was 78 months ranging from 21 to 96 months. Of note, 4 deaths were recorded over the 8-year follow-up. Random forest classification found age, clinical stage, SUV_{max}, skewness$_{_ER}$, kurtosis$_{_ER}$, entropy$_{_HISTO}$, and uniformity$_{_HISTO}$ to be variables of importance to predict the 8y-EFS (**Supplemental Figure 1**).

TABLE 1 | Patients characteristics.

Characteristics	All patients (n = 38)	
Age (years, mean [min–max])	55 [32–80]	
Histology (n, %)		
Invasive ductal carcinoma	30	78.9
Invasive lobular carcinoma	2	5.3
Tubular carcinoma	1	2.6
Mixed carcinoma	5	13.2
Tumor stage (n, %)		
1	10	26.3
2	20	52.6
3	8	21.1
Nodal stage (n, %)		
0	11	28.9
1	17	44.7
2	5	13.2
3	5	13.2
Elston and Ellis grade (n, %)		
I	4	10.5
II	22	57.9
III	12	31.6
Molecular subtype classification (n, %)		
Luminal A	24	63.2
Luminal B/HER-2 negative	10	26.3
Luminal B/HER-2 positive	4	10.5

The OOB estimate was equal to 28.9%. Mean skewness$_{_ER}$ and mean entropy$_{_HISTO}$ were significantly higher ($p = 0.001$ and $p = 0.022$, respectively), whereas mean uniformity$_{_HISTO}$ was significantly lower ($p = 0.022$) in 8y-EFS_1 patients (**Figure 3**). There were no significant difference in SUV_{max} and kurtosis$_{_ER}$ values between 8y-EFS_0 and 8y-EFS_1 patients ($p = 0.760$ and $p = 0.052$, respectively). Representative images of PET and digital-immunochemistry images are displayed in **Figure 4**. On ROC analyses, optimal cut-off values for skewness$_{_ER,}$ entropy$_{_HISTO}$ and uniformity$_{_HISTO}$ to predict 8y-EFS were equal to 0.163, 1.23, and 0.066, respectively (**Table 2**). Univariable Kaplan-Meier survival analyses found that skewness$_{_ER}$ was a predictor of 8y-EFS whereas entropy$_{_HISTO}$ and uniformity$_{_HISTO}$ were not, although statistical significance was almost reached (**Figure 5**). On multivariable analysis including skewness$_{_ER}$ and other well-known prognosticators [age, clinical stage (I–II versus III)], all the statistics for the test of the null hypothesis are significant and we can conclude that considering explanatory variables provides significant additional information. There was no violation of the proportional hazards assumption. However, regression coefficients showed that none of the parameters were independent predictors of 8y-EFS (**Table 3**). Indeed, we found a significant negative correlation between skewness$_{_ER}$ and age ($\rho = -0.351$, $p = 0.031$) with skewness$_{_ER}$ values higher in youngest patients (**Figure 6A**). Moreover, skewness$_{_ER}$ was significantly higher in clinical stage III tumors ($p = 0.023$, **Figure 6B**). Of note, ER expression was scored + in 2 patients (5.3%), ++ in 6 patients (15.8%), and +++ in 30 patients (78.9%) by IHC analysis. Skewness$_{_ER}$ was not significantly different between patients scored +, ++, or +++ ($p = 0.508$, **Supplemental Figure 2**). A quantification of ER expression in percentage was also available for 35 patients with a mean value equal to 88.5% (± 15.5). It was not significantly correlated with skewness$_{_ER}$ ($p = 0.207$, $\rho = 0.048$).

DISCUSSION

The first and interesting finding of the present study is the quasi-absence of correlation between ER and PR descriptors of the distribution shape and Haralick texture parameters. This seems to indicate that their heterogeneity expressions are independent and could have different meanings and clinical consequences. Here, we decided to focus on EFS and it appeared that immunochemistry histogram parameters of estrogen receptors, and especially skewness, are predictors of 8y-EFS together with age and clinical stage, whereas none of the progesterone receptors were. Moreover, correlations of ER and PR parameters with PET histogram and textural parameters were clearly different. The ER immunochemistry heterogeneity was mainly correlated to PET histogram parameters, whereas PR immunochemistry heterogeneity was mainly correlated to second-order GLCM-derived PET textural features. Interestingly, skewness$_{_ER}$ was a significant predictor of 8y-EFS but not an independent one. Indeed, it was related to both the age of the patient at diagnosis and the clinical stage of the disease: estrogen receptors heterogeneity was higher in youngest patients and in higher-staged diseases. We can hypothesize that ER heterogeneity could be linked to more aggressive tumors. Returning to the PET methodology, the use of a HR PET acquisition to compute ^{18}F-FDG heterogeneity parameters (PSF algorithm and 1.3 × 1.3 × 1.9 mm voxels) is a strength. Indeed, it has been previously shown that the type of reconstruction as well as the voxel size, are important considerations when computing ^{18}F-FDG heterogeneity (19) especially in small lesions like those bearing breast cancer. However, even though high-resoluted histograms of PET parameters were significantly but fairly correlated to ER immunochemistry ones (especially skewness$_{_ER}$, kurtosis$_{_ER}$, entropy$_{_HISTO}$ and uniformity$_{_HISTO}$), PET parameters appeared to be less discriminant for 8y-EFS than immunochemistry ones. Nevertheless, we can notice that log-rank tests for entropy$_{_HISTO}$ and uniformity$_{_HISTO}$ almost reached statistical significance and that a larger study could have displayed more discriminant results.

Previously in the study of Antunovic et al. (13), using PET metabolic heterogeneity features, two clusters were obtained by the unsupervised hierarchical clustering analyses with different imaging signatures. Besides, these signatures were significantly associated with different molecular subtypes. Ha et al. (14) also performed an unsupervised tumor clustering using a radiomics pattern which resulted in 3 tumor clusters. The expression of histopathological factors between their clusters was different for Ki67. Of note, one cluster displayed higher estrogen and progesterone receptors (ER and PR) expression, but statistical significance was not reached. Lemarignier et al. (15) found a trend for lower local heterogeneity in hormone-positive breast cancer even though statistical significance was no longer observed after correction for multiple testing. Thus, all these results together with ours are first-evidences of a complementary role of imaging features, together with standard PET metrics for a clinically relevant *in vivo* characterization of breast cancer that could lead to a personalization of therapeutic management. The perspectives would be (i) to assess the clinical impact of these results, in particular by offering patients deemed to be at risk of recurrence a closer post-therapeutic monitoring and (ii) to test

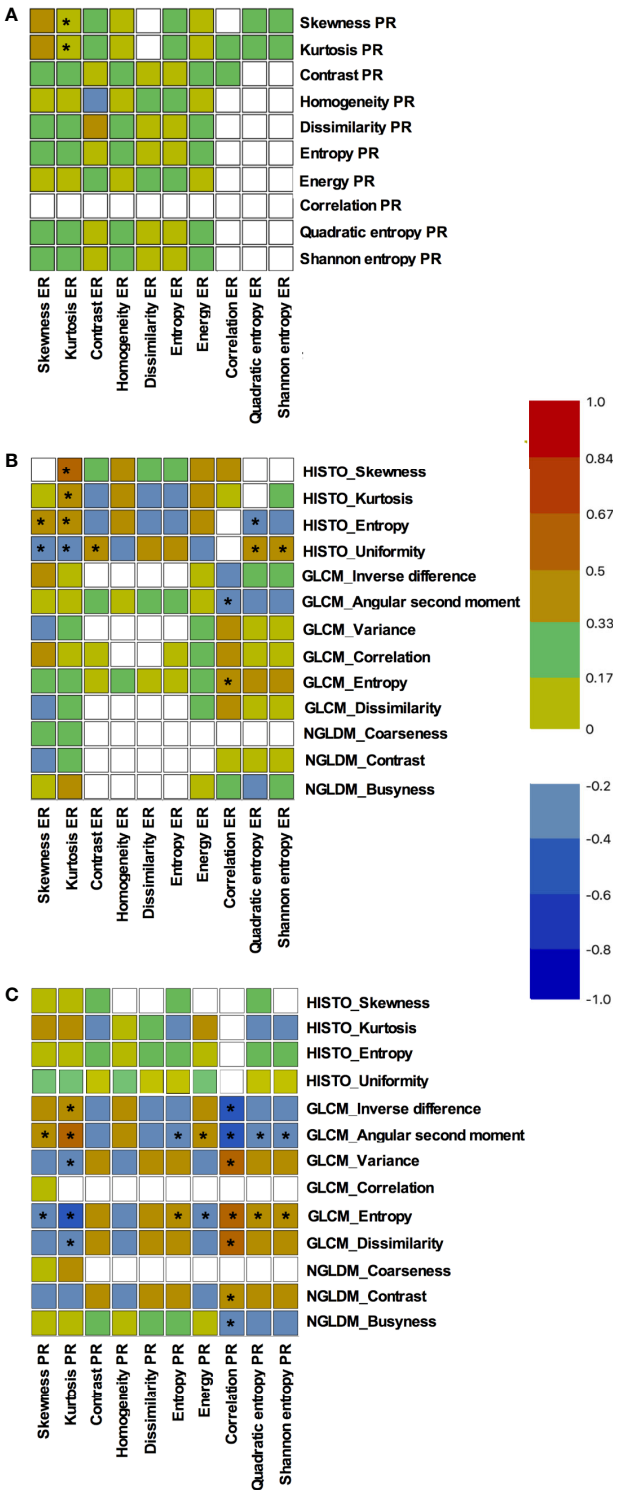


FIGURE 2 | Correlations among distribution descriptors and Haralick texture parameters of estrogen and progesterone receptors expression. Results are presented as Spearman correlations maps: **(A)** correlations between estrogen and progesterone receptors expression parameters, **(B)** correlations between ^{18}F-FDG textural parameters and estrogen receptors expression parameters, **(C)** correlations between ^{18}F-FDG textural parameters and progesterone receptors expression parameters. The blue color corresponds to a correlation close to −1 and the red color corresponds to a correlation close to 1. The green corresponds to a correlation close to 0. * represents significant correlations ($p < 0.05$).

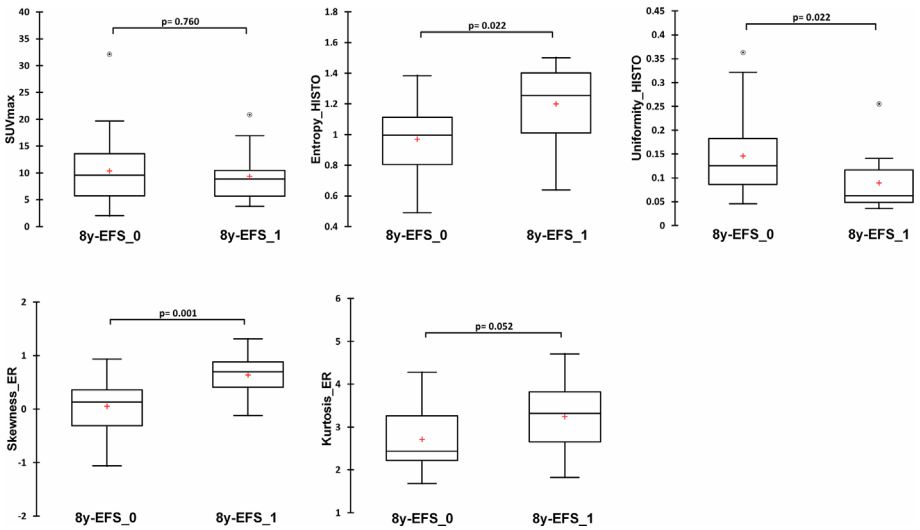


FIGURE 3 | Comparison of immunochemistry and PET variables of importance identified by random forest analysis between 8y-EFS_0 and 8y-EFS_1 patients (SUVmax, entropy$_{_HISTO}$, uniformity$_{_HISTO}$, skewness$_{_ER}$, and kurtosis$_{_ER}$). Data are shown as Tukey boxplots with (O) representing outliers.

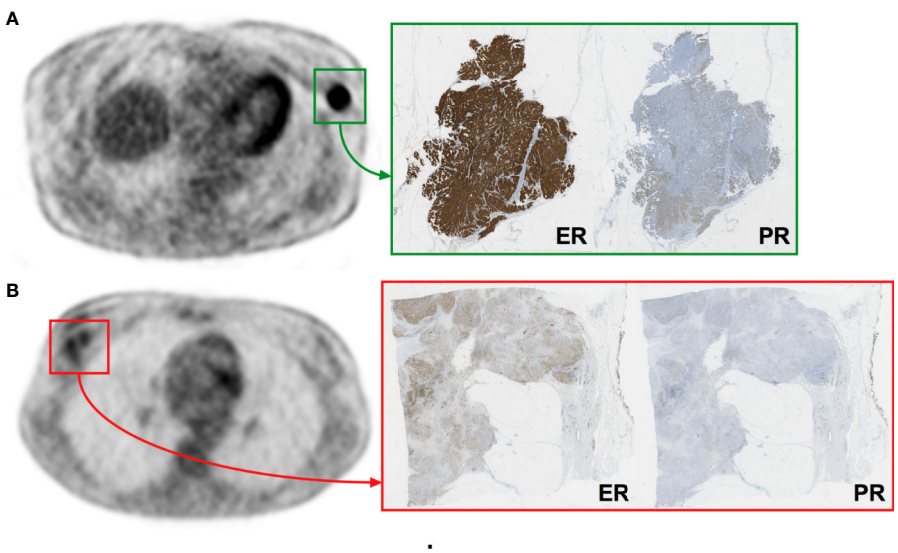


FIGURE 4 | Representative images of PET and digital-immunochemistry images. Patient **(A)** was a 74-year-old women with a luminal ER+/PR+ tumor staged II presenting homogeneous IHC and PET characteristics (skewness$_{_ER}$ = −1.06, entropy$_{_HISTO}$ = 0.66, uniformity$_{_HISTO}$ = 0.24) who experienced no event at 8 years (8y-EFS_0). Patient **(B)** was a 34-year-old women with a luminal ER+/PR+ tumor staged III presenting heterogeneous IHC and PET characteristics (skewness$_{_ER}$ = 1.31, entropy$_{_HISTO}$ = 1.50, uniformity$_{_HISTO}$ = 0.04) who experienced an event at 8 years (8y-EFS_1).

TABLE 2 | ROC analyses for 8-year event free survival for skewness$_{_ER}$, entropy$_{_HISTO}$, and uniformity$_{_HISTO}$.

Variable	AUC	Standard error	Lower bound (95%)	Upper bound (95%)	P	Cut-off value
Skewness$_{_ER}$	0.828	0.083	0.666	0.991	<0.0001	>0.163
Entropy$_{_HISTO}$	0.737	0.113	0.515	0.960	0.036	>1.230
Uniformity$_{_HISTO}$	0.741	0.116	0.514	0.968	0.038	<0.066

AUC, area under the curve.

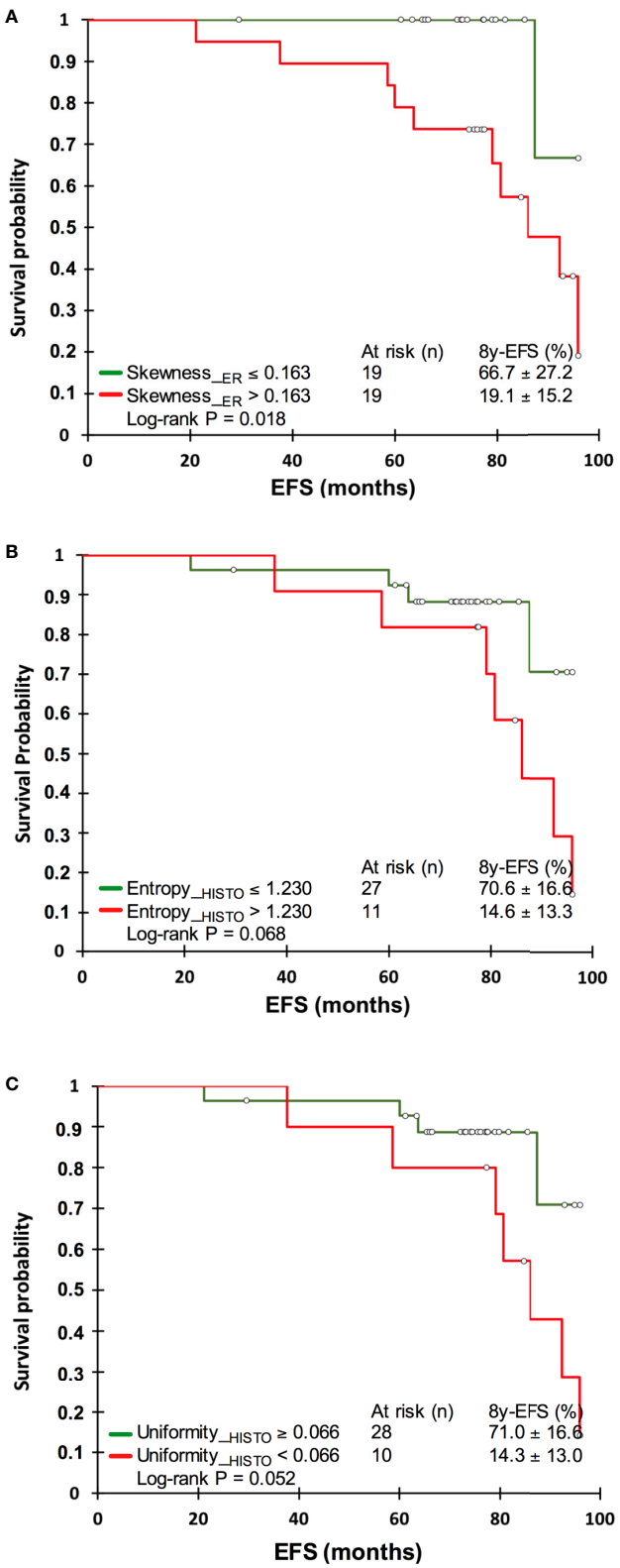


FIGURE 5 | Kaplan-Meyer analyses for skewness$_{_ER}$ **(A)**, entropy$_{_HISTO}$ **(B)**, and uniformity$_{_HISTO}$ **(C)**.

TABLE 3 | Cox regression analysis.

Test of the null hypothesis			
Statistic	**DF**	**Chi-square**	**P**
Likelihood ratio test	3	10.60	0.014
Score test	3	10.86	0.012
Wald test	3	9.24	0.026

Regression coefficients							
Variable	**Value**	**Standard error**	**Wald Chi-square**	**P**	**HR**	**HR lower bound (95%)**	**HR upper bound (95%)**
$Skewness_{\_ER}$	0.860	0.834	1.063	0.303	2.363	0.461	12.119
Age	–0.033	0.033	1.406	0.306	0.967	0.907	1.031
Clinical stage III	1.245	0.858	2.105	0.147	3.474	0.646	18.681

HR, Hazard ratio.

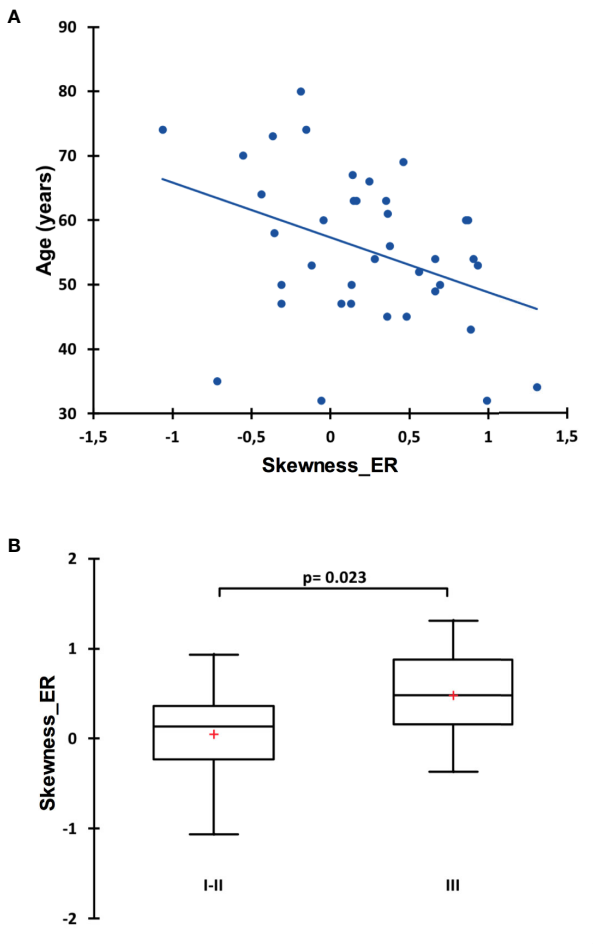


FIGURE 6 | Spearman correlation between $skewness_{\_ER}$ and the age at diagnosis **(A)** and comparison of $skewness_{\_ER}$ between clinical staged I–II versus staged III patients. Data is shown as Tukey boxplots **(B)**.

other innovative tracers such as ^{18}F-Fluoroestradiol. Data from a meta-analysis evaluating the ability of ^{18}F-Fluoroestradiol for the determination of tumor ER status (32) suggested acceptable diagnostic performance of this radiopharmaceutical despite a weakness in terms of sensitivity [pooled sensitivity = 82% (95% CI: 74–88%), pooled specificity = 95% (95% CI: 86–99%)]. However, to date, there is no data clearly documenting the clinical consequences of patient management following diagnosis with ^{18}F-Fluoroestradiol PET. Documenting the intra-tumoral heterogeneity of estrogen receptors using this tracer has not yet been investigated and could be of interest.

It is worth noticing that our findings, even if innovative, were observed in a small cohort and have to be validated by a larger clinical study. The lack of statistical significance might also be due to the limited spatial resolution of an analogic system and it could be wise to test innovative digital systems in future projects. Of note, PET third-order textural features were not considered in the present study because their computation was very far from that used for immunochemistry parameters. Indeed, immunochemistry parameters could only use histograms or co-occurrence matrixes. Also, inter-observer variability for the quantification of metabolic heterogeneity was not presently assessed. However, we have taken care to choose one of the most reproducible delineation methods, namely, a gradient-based method (33), thus limiting the variability linked to the operator. However, other sources of variability must be taken into account regarding the clinical export of such results: software, PET systems, reconstructions, etc. Therefore, we acknowledge that harmonization strategies will be necessary anyway. Finally concerning immunochemistry methodology, the age of the samples jeopardized the achievement of Ki67 expression heterogeneity exploration because of faint immunostaining, not enabling the digital-immunochemistry computation. For HER2 status, international standards require that it be tested at the time of diagnosis, therefore on biopsies. The recommendations say that it is not necessary to repeat it systematically on the piece of excision, because there is a good agreement between the HER2 status tested on the biopsy and remade on the piece, due to a usually homogeneous distribution when expressed (34–36).

To conclude, a heterogeneous distribution of estrogen receptors within the tumor in immunochemistry appeared as an event-free prognosticator in luminal non-metastatic breast cancers.

Furthermore, estrogen receptors heterogeneity is higher in youngest patients and the highest-graded tumors. Interestingly, this appeared to be correlated with PET histogram parameters which could therefore become potential tools to reflect the tumor estrogen receptors heterogeneity, provided these results are confirmed by further larger and prospective studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by CPP Nord Ouest III, reference 2009-10. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Study conception and design: NA and CLa. Screening and inclusion of patients: CLe. Data collection: TS and CLa. PET/CT analysis: TS. Immunochemistry: CB-F. Digital-immunochemistry computation: NE. Statistical analysis: CLa. Manuscript editing and reviewing: CLa, CB-F, NE, NA, and CLe. All authors contributed to the article and approved the submitted version.

FUNDING

This project was funded by an internal call for tenders from the Center François Baclesse.

ACKNOWLEDGMENTS

Helen Lapasset is thanked for English editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.599050/full#supplementary-material

SUPPLEMENTARY FIGURE 1 | Random forest variables of importance analysis displayed as mean decrease accuracy.

SUPPLEMENTARY FIGURE 2 | Comparison of skewness$_{-ER}$ by estrogen receptor expression score (+, ++ or +++). Data is shown as Tukey boxplots.

REFERENCES

1. Weigelt B, Baehner FL, Reis-Filho JS. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. *J Pathol* (2010) 220(2):263–80. doi: 10.1002/path.2648
2. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* (2001) 98(19):10869–74. doi: 10.1073/pnas.191367098
3. Prat A, Pineda E, Adamo B, Galvan P, Fernandez A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast (Edinburgh Scotland)* (2015) 24(Suppl 2):S26–35. doi: 10.1016/j.breast.2015.07.008
4. Provenzano E, Ulaner GA, Chin SF. Molecular Classification of Breast Cancer. *PET Clin* (2018) 13(3):325–38. doi: 10.1016/j.cpet.2018.02.004
5. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* (2011) 5(1):5–23. doi: 10.1016/j.molonc.2010.11.003
6. Groheux D, Martineau A, Teixeira L, Espie M, de Cremoux P, Bertheau P, et al. (18)FDG-PET/CT for predicting the outcome in ER+/HER2- breast cancer patients: comparison of clinicopathological parameters and PET image-derived indices including tumor texture analysis. *Breast Cancer Res BCR* (2017) 19(1):3. doi: 10.1186/s13058-016-0793-2
7. Groheux D, Hindie E. Breast Cancer Staging: To Which Women Should 18F-FDG PET/CT Be Offered? *J Nucl Med* (2015) 56(8):1293. doi: 10.2967/jnumed.115.160945
8. Groheux D, Cochet A, Humbert O, Alberini JL, Hindie E, Mankoff D. (1)(8)F-FDG PET/CT for Staging and Restaging of Breast Cancer. *J Nucl Med* (2016) 57(Suppl 1):17s–26s. doi: 10.2967/jnumed.115.157859
9. Groheux D. FDG-PET/CT for systemic staging of patients with newly diagnosed breast cancer. *Eur J Nucl Med Mol Imaging* (2017) 44(9):1417–9. doi: 10.1007/s00259-017-3731-3
10. Rosen EL, Eubank WB, Mankoff DA, FDG PET. PET/CT, and breast cancer imaging. *Radiographics* (2007) 27(Suppl 1):S215–29. doi: 10.1148/rg.27si075517
11. Avril N, Rose CA, Schelling M, Dose J, Kuhn W, Bense S, et al. Breast imaging with positron emission tomography and fluorine-18 fluorodeoxyglucose: use and limitations. *J Clin Oncol* (2000) 18(20):3495–502. doi: 10.1200/JCO.2000.18.20.3495
12. Acar E, Turgut B, Yigit S, Kaya G. Comparison of the volumetric and radiomics findings of 18F-FDG PET/CT images with immunohistochemical prognostic factors in local/locally advanced breast cancer. *Nuclear Med Commun* (2019) 40(7):764–72. doi: 10.1097/MNM.0000000000001019
13. Antunovic L, Gallivanone F, Sollini M, Sagona A, Invento A, Manfrinato G, et al. [(18)F]FDG PET/CT features for the molecular characterization of primary breast tumors. *Eur J Nucl Med Mol Imaging* (2017) 44(12):1945–54. doi: 10.1007/s00259-017-3770-9
14. Ha S, Park S, Bang JI, Kim EK, Lee HY. Metabolic Radiomics for Pretreatment (18)F-FDG PET/CT to Characterize Locally Advanced Breast Cancer: Histopathologic Characteristics, Response to Neoadjuvant Chemotherapy, and Prognosis. *Sci Rep* (2017) 7(1):1556. doi: 10.1038/s41598-017-01524-7
15. Lemarignier C, Martineau A, Teixeira L, Vercellino L, Espie M, Merlet P, et al. Correlation between tumour characteristics, SUV measurements, metabolic tumour volume, TLG and textural features assessed with (18)F-FDG PET in a large cohort of oestrogen receptor-positive breast cancer patients. *Eur J Nucl Med Mol Imaging* (2017) 44(7):1145–54. doi: 10.1007/s00259-017-3641-4
16. Laurinavicius A, Plancoulaine B, Herlin P, Laurinaviciene A. Comprehensive Immunohistochemistry: Digital, Analytical and Integrated. *Pathobiol J Immunopathol Mol Cell Biol* (2016) 83(2-3):156–63. doi: 10.1159/000442389
17. Plancoulaine B, Laurinaviciene A, Herlin P, Besusparis J, Meskauskas R, Baltrusaityte I, et al. A methodology for comprehensive breast cancer Ki67 labeling index with intra-tumor heterogeneity appraisal based on hexagonal

tiling of digital image analysis data. *Virchows Arch* (2015) 711–22. doi: 10.1007/s00428-015-1865-x
18. Bellevre D, Blanc Fournier C, Switsers O, Dugue AE, Levy C, Allouache D, et al. Staging the axilla in breast cancer patients with (1)(8)F-FDG PET: how small are the metastases that we can detect with new generation clinical PET systems? *Eur J Nucl Med Mol Imaging* (2014) 41(6):1103–12. doi: 10.1007/s00259-014-2689-7
19. Aide N, Salomon T, Blanc-Fournier C, Grellard JM, Levy C, Lasnon C. Implications of reconstruction protocol for histo-biological characterisation of breast cancers using FDG-PET radiomics. *EJNMMI Res* (2018) 8(1):114. doi: 10.1186/s13550-018-0466-5
20. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging* (2010) 37 (1):181–200. doi: 10.1007/s00259-010-1458-5
21. Nioche C, Orlhac F, Boughdad S, Reuze S, Goya-Outi J, Robert C, et al. LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res* (2018) 78(16):4786–9. doi: 10.1158/0008-5472.CAN-18-0125
22. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging* (2019) 46(13):2638–55. doi: 10.1007/s00259-019-04391-8
23. Orlhac F, Nioche C, Soussan M, Buvat I. Understanding Changes in Tumor Texture Indices in PET: A Comparison Between Visual Assessment and Index Values in Simulated and Patient Data. *J Nucl Med* (2017) 58(3):387–92. doi: 10.2967/jnumed.116.181859
24. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. 18F-FDG PET-Derived Textural Indices Reflect Tissue-Specific Uptake Pattern in Non-Small Cell Lung Cancer. *PLoS One* (2015) 10(12):e0145063. doi: 10.1371/journal.pone.0145063
25. Laurinavicius A, Plancoulaine B, Laurinaviciene A, Herlin P, Meskauskas R, Baltrusaityte I, et al. A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue. *Breast Cancer Res BCR* (2014) 16(2):R35. doi: 10.1186/bcr3639
26. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Trans Syst Man Cybern* (1973) SMC-3(6):610–21. doi: 10.1109/TSMC.1973.4309314
27. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thurlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* (2013) 24 (9):2206–23. doi: 10.1093/annonc/mdt303
28. Breiman L. *Classification and Regression Trees*. New York: Routledge (1984).
29. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer-Verlag (2009).
30. Breiman L. Bagging Predictors. *Mach Learn* (1996) 24:123–40. doi: 10.1007/BF00058655
31. Breiman L. *Manual on setting up, using and understanding random forests, v 3.1*. Berkeley, California: Statistical Department of the University of Berkeley (2002).
32. Evangelista L, Guarneri V, Conte PF. 18F-Fluoroestradiol Positron Emission Tomography in Breast Cancer Patients: Systematic Review of the Literature & Meta-Analysis. *Curr Radiopharm* (2016) 9(3):244–57. doi: 10.2174/1874471009666161019144950
33. Lasnon C, Enilorac B, Popotte H, Aide N. Impact of the EARL harmonization program on automatic delineation of metabolic active tumour volumes (MATVs). *EJNMMI Res* (2017) 7(1):30. doi: 10.1186/s13550-017-0279-y
34. Ramakrishna N, Temin S, Chandarlapaty S, Crews JR, Davidson NE, Esteva FJ, et al. Recommendations on Disease Management for Patients With Advanced Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer and Brain Metastases: ASCO Clinical Practice Guideline Update. *J Clin Oncol* (2018) 36 (27):2804–7. doi: 10.1200/JCO.2018.79.2713
35. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* (2013) 31 (31):3997–4013. doi: 10.1200/JCO.2013.50.9984
36. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *J Clin Oncol* (2018) 36 (20):2105–22. doi: 10.1200/JCO.2018.77.8738

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 20 January 2021
doi: 10.3389/fonc.2020.541663

Influence of Magnetic Field Strength on Magnetic Resonance Imaging Radiomics Features in Brain Imaging, an *In Vitro* and *In Vivo* Study

Samy Ammari[1,2], Stephanie Pitre-Champagnat[2], Laurent Dercle[1,3,4], Emilie Chouzenoux[5], Salma Moalla[1], Sylvain Reuze[6], Hugues Talbot[5], Tite Mokoyoko[1], Joya Hadchiti[1], Sebastien Diffetocq[1], Andreas Volk[2], Mickeal El Haik[1], Sara Lakiss[1], Corinne Balleyguier[1,2], Nathalie Lassau[1,2] and Francois Bidault[1,2]*

[1] Department of Radiology, Gustave Roussy Cancer Campus, Université Paris-Saclay, Villejuif, France, [2] BioMaps (UMR1281), Université Paris-Saclay, CNRS, INSERM, CEA, Orsay and Gustave Roussy, Villejuif, France, [3] Immunology of Tumours and Immunotherapy INSERM U1015, Gustave Roussy Cancer Campus, Université Paris Saclay, Villejuif, France, [4] Radiology Department, Columbia University Medical Center, New York Presbyterian Hospital, New York, NY, United States, [5] Center for Visual Computing, CentraleSupelec, Inria, Université Paris-Saclay, Gif-sur-Yvette, France, [6] Department of Radiotherapy - Medical Physics, Gustave Roussy, Université ParisSaclay, Villejuif, France

Background: The development and clinical adoption of quantitative imaging biomarkers (radiomics) has established the need for the identification of parameters altering radiomics reproducibility. The aim of this study was to assess the impact of magnetic field strength on magnetic resonance imaging (MRI) radiomics features in neuroradiology clinical practice.

Methods: T1 3D SPGR sequence was acquired on two phantoms and 10 healthy volunteers with two clinical MR devices from the same manufacturer using two different magnetic fields (1.5 and 3T). Phantoms varied in terms of gadolinium concentrations and textural heterogeneity. 27 regions of interest were segmented (phantom: 21, volunteers: 6) using the LIFEX software. 34 features were analyzed.

Results: In the phantom dataset, 10 (67%) out of 15 radiomics features were significantly different when measured at 1.5T or 3T (student's t-test, $p < 0.05$). Gray levels resampling, and pixel size also influence part of texture features. These findings were validated in healthy volunteers.

Conclusions: According to daily used protocols for clinical examinations, radiomic features extracted on 1.5T should not be used interchangeably with 3T when evaluating texture features. Such confounding factor should be adjusted when adapting the results of a study to a different platform, or when designing a multicentric trial.

Keywords: tissue features, heterogeneous phantom, homogeneous phantom, magnetic fields, texture, magnetic resonance imaging

OPEN ACCESS

Edited by:
Freimut Dankwart Juengling, Universität Bern, Switzerland

Reviewed by:
Prathyush Chirra, Case Western Reserve University, United States
Alberto Traverso, Maastro Clinic, Netherlands

****Correspondence:***
Samy Ammari
samy.ammari@gustaveroussy.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *10 March 2020*
Accepted: *23 November 2020*
Published: *20 January 2021*

Citation:
Ammari S, Pitre-Champagnat S, Dercle L, Chouzenoux E, Moalla S, Reuze S, Talbot H, Mokoyoko T, Hadchiti J, Diffetocq S, Volk A, El Haik M, Lakiss S, Balleyguier C, Lassau N and Bidault F (2021) Influence of Magnetic Field Strength on Magnetic Resonance Imaging Radiomics Features in Brain Imaging, an In Vitro and In Vivo Study. Front. Oncol. 10:541663. doi: 10.3389/fonc.2020.541663

HIGHLIGHTS

1. - Radiomic features at 1.5T are not interchangeable with 3T when evaluating tumor texture
2. - Field strength should be taken into account in the interpretation of the texture indices
3. - Signal to noise ratio should be taken into account in the interpretation of the texture indices

INTRODUCTION

Radiomics is a fast growing discipline, which is undergoing growing interest in computational medical imaging (1). This field of medical study aims at extracting a large amount of quantitative features from medical images using data-characterization algorithms. This research field is faced with multiple challenges (2). Radiomics is used in oncology to analyze features that are invisible to the naked eye and that may be associated with gene expression, tumor histology, treatment response and patient outcome (3).

MRI has several advantages and disadvantages for radiomics analysis (4–7). Among imaging modalities, it offers the best soft tissue contrast. Conversely, differences in MRI parameters (field strength, gradient characteristics), image acquisition protocols (8, 9), sequences, pixel size (10), and the signal-to-noise ratio (SNR) (11, 12) might impact radiomics features that are sensitive to image quality.

The impact of MRI acquisition and processing on radiomics reproducibility is scarcely reported. As in nuclear medicine, the validation of a biologically relevant and reproducible clinical biomarker based on radiomics implies standardization of protocols across several centers (13). For example, 3T compared with 1.5T MRI not only provides higher SNR, allows increased image resolution, and modifies relaxation times T1 and T2 but also induces some artifacts. Thus, there is a clear need to evaluate the influence of field strength and related settings like image resolution (pixel size, field of view [FOV], and matrix) on radiomic features.

Recent articles using radiomics as biomarkers consider that the major challenge is that grayscale MRI intensities, contrary to X-ray CT, are not standardized and are highly dependent on manufacturer, sequence type and acquisition parameters (14, 15).

To address this problem, authors focused on image pre-processing techniques that effectively minimize MR intensity inhomogeneity in a tissue region (16–18), spatial resampling (17–20) and brain arch extraction prior to image intensity normalisation (21, 22).

Although several studies have shown variability in texture analysis as a function of MRI acquisition parameters and gray level discretization steps, none of them have evaluated the combined impact of magnetic field strength, matrix size, pixel size, intensity normalization, and gray level discretization pre-processing methods on MRI radiomic feature values (9, 10, 23).

This study was designed to evaluate the impact of the field strength (1.5T vs. 3T) on radiomic features. Two clinical 1.5 and 3T MR devices from the same manufacturer were used to image the same phantoms, mimicking homogeneous and heterogeneous tissues. We also imaged healthy volunteers with the same sequences acquisition and image processing parameters.

It is believed that this step is crucial before proposing recommendations to standardize brain MRI pre-processing techniques, which is in turn essential for guaranteeing reliable radiomics-based models with big data and artificial intelligence.

MATERIAL AND METHODS

This prospective study was approved by the local institution review board.

Homogeneous Phantoms

Our homogeneous phantom was designed to mimic cerebrospinal fluid and opacified blood vessels. It was defined with different gadolinium concentrations: eight 30-ml tubes were filled with demineralized water mixed with increasing gadolinium chelate concentration (Gadoteric Acid; Dotarem®, Guerbet): 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, and 2 mmol/l (**Figures 1A, B**). Those eight homogeneous tubes were designated C1 to C8.

Heterogeneous Phantoms

Our heterogeneous phantom, designed to mimic the brain white matter, was composed of agarose gel coating polystyrene beads. The proton density of agarose has similar characteristics and relaxation times to biological tissues (10). We defined six heterogeneous tubes and two homogeneous tubes. The 30-ml tubes were filled with polystyrene beads of different diameters (1, 2 and 3 mm) in an agarose gel solution (2%), either pure or mixed with 0.25 mmol/l Gadolinium chelate (**Figures 1C, D**) (9). The tubes were displayed in the tube slots of the Eurospin phantom (A Eurospin II-(TO5) phantom; Diagnostic Sonar) glass cylinder, filled with 1% copper sulfate (24). The eight tubes were designated T1 to T8.

Healthy Volunteers

Clinical MRI sequences were performed in 10 healthy volunteers aged 21–26 years (six men, four women). Two MRI acquisitions were performed on each MRI device within a 40-min time interval.

MRI Devices and Protocols

The influence of field strength was tested on two different MRI devices from the same manufacturer (General Electric): an Optima MR450w 1.5T superconducting magnet MRI installed in 2016 with a 70 cm tunnel, 32 channels, 50 cm FOV (Z axis), and gradients 40 mT SR 200 mT/m/s, and a Discovery MR750w 3T superconducting magnet MRI installed in 2012, with a 70 cm tunnel, 32 channels, 50 cm FOV (Z axis), and gradients 44 mT/m SR 200 T/m/s.

Abbreviations: 3D SPGR, 3D rapid gradient echo sequence; EORTC, European Organization for Research and Treatment of Cancer; TR, Repeat time; MTX, Image matrix; CoMat, The co-occurrence matrix; GRLM, The gray-level run length matrix; GZLM, The gray-level zone length matrix; SBW, Sampling bandwidth.

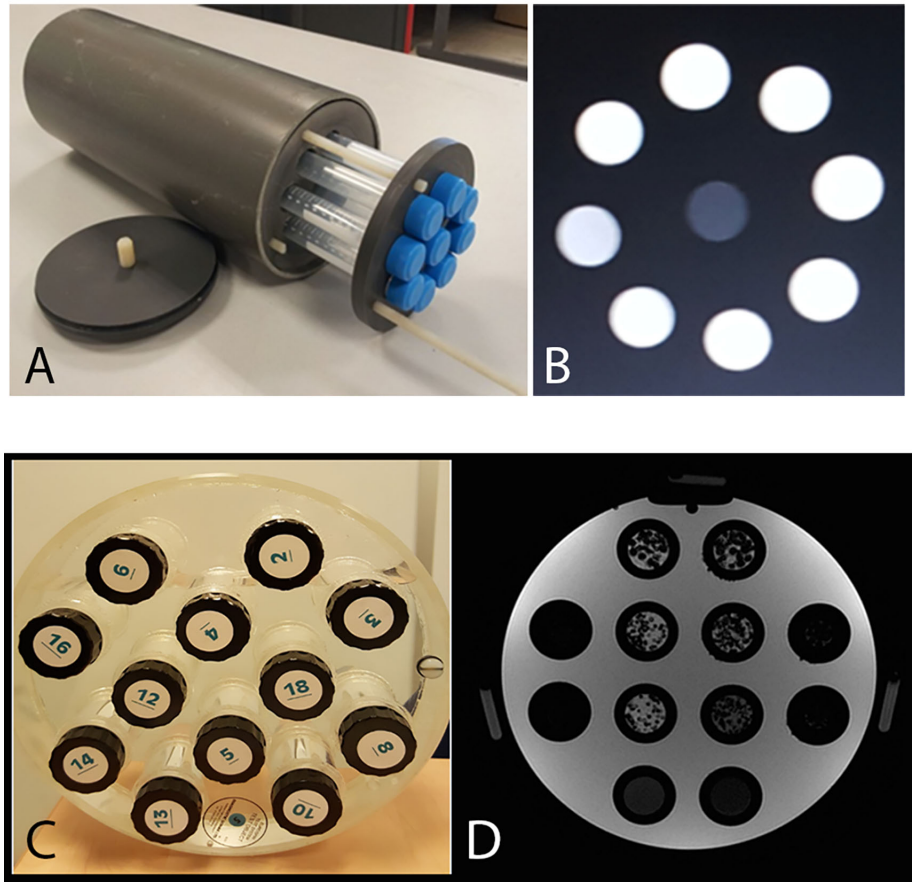


FIGURE 1 | **(A)** Homogeneous phantoms: eight 30-ml tubes filled with demineralized water mixed with increasing gadolinium chelate concentration (Gadoteric Acid; Dotarem®, Guerbet): 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75 and 2 mmol/l and a central demineralized water tube. **(B)** T1 weighted MRI image of the homogeneous phantom (256x256 matrix, 18 cm FOV). **(C, D)** T1 weighted MRI image of the heterogeneous phantom (256;256 matrix, 24 cm FOV): in the two central columns are six 30-ml tubes filled with polystyrene beads of different diameters (1, 2 and 3 mm) in an agarose gel solution (2%), either pure or mixed with 0.25 mmol/l gadolinium chelate, and two 30-ml tubes filled with agarose gel solution (2%) and two different concentration of gadolinium chelate.

We used 'head and neck' coils with 32 channels with a 35 cm diameter adapted to the frequency of each MRI.

Acquisition

MRI imaging acquired was based on a T1-weighted 3D rapid gradient echo sequence (3D SPGR). This sequence is used in clinical imaging for rapid volumetric imaging and can be acquired before or after contrast agent injection. The same sequences parameters (matrix, FOV, plane) were used for both MRIs. Parameters were chosen as recommended by the European Organization for Research and Treatment of Cancer (EORTC) to explore brain tumors: repeat time (TR) 6.1 ms; echo time (TE) inphase 1.2–2.1 ms; NEX1; thickness 1 mm in contiguous sections and bandwidth at 31 with 165 slices; 2–5-min acquisition depending on the FOV and matrices. The temperature was maintained between 19 and 21°C in each MRI during acquisition. For the healthy volunteers, the same 3DT1 SPGR sequence was used covering the whole encephalon.

Acquisition Parameters: Field of View and Matrix Size

For the phantom study, five couples of matrices and FOV, determining various pixel size, were applied on both machines (**Table 1**). The couples were chosen to be suitable with clinical acquisition. In healthy volunteers only one acquisition per device was performed, the FOV was fixed at 24 cm with an image matrix of 256 × 256 on both machines.

Texture Analysis

Segmentations and texture features calculations were performed with the freely available LIFEX software package (http://www.lifexsoft.org) (25). A 6 cm^3 VOI was placed in each homogeneous and heterogeneous phantom's tube. VOI position along Z axis of tubes was controlled according to the reference markers located on the phantom surface. For healthy volunteers, VOI were displayed in six bilateral areas: corpus callosum, central gray nuclei, and white matter of the centrum semiovale.

TABLE 1 | Matrices and FOVs studied at 1.5 and 3T with the corresponding pixel size.

MTX (pixels)	Field of view (cm)	Pixel size (mm)
256 x 256	24	0.938
256 x 256	18	0.703
256 x 256	12	0.468
256 x 128	24	0.937 x 1.875
128 x 128	24	1.875

In each VOI, 38 texture features (indices) were extracted from first order and second-order statistics, arising from the analysis of the intensity histogram and the calculation of three texture matrices: the co-occurrence matrix (CM), the gray-level run length matrix (GRLM) and the gray-level zone length matrix (GZLM) (26). The texture analysis software needs to resample the 16-bit gray levels of MRI images in order to calculate these features. Voxel intensities were resampled in three different ways, using 256, 128 and 64 discrete values.

Calculation of Signal-to-Noise Ratio

Mean SNR values were calculated on both MRIs by measuring the ratio between the mean signal intensity in each pure agarose sample and the standard deviation (SD) of the background noise selected in the frequency encoding direction. Calculation of SNR in healthy volunteers was also performed by measuring the mean signal intensity in each white substance and the SD of the background noise selected in the frequency encoding direction.

Statistical Methods

Two-by-two correlation texture features were calculated. The statistical significant differences between 1.5 and. 3T were determined by the Student's t-test using paired data. In the event of a lack of normal distribution (N < 30), results were also presented as boxplots to demonstrate the trend of texture values on 1.5T vs. 3T. The association between acquisition protocols (magnetic field, matrix, FOV), textural changes in the phantom (homogeneous vs. heterogeneous and polystyrene size), and imaging features, was calculated with Spearman's rho correlation coefficients with P-values corrected for multiple tests. Statistical analyses were performed using SPSS v24.0. The datasets generated and analyzed are available from the corresponding author.

RESULTS

Phantom Study

Influence of Software Resampling Step

The analysis software needs to resample the images gray scale by sub-sampling the number gray levels, in order to be able to calculate texture features. Thus, we firstly explored the influence of this sub-sampling step before exploring the influence of magnetic field strength. We did this through the two-by-two coefficient correlation calculation between the texture features (**Figure 2**). A colored (red/blue) hierarchical clustering is presented, showing the importance of the correlation beteween coefficients (**Figure 2**) for each level of resampling. A strong correlation was observed for a large number of texture features at the 256 gray level (only 9 features were totally independent with a Spearman's rho <0,70). The correlation value decreased with the sub-sampling in gray levels resolution. The gray level sub-sampling influence on texture features was not significant for parameters extracted from the histogram. In contrast, the gray levels sub-sampling had a significant impact on parameters extracted from matrices.

Influence of Field Strength on Homogeneous Phantom

The majority of texture features values were significantly different between the two magnetic fields (1.5T vs. 3T). For example the mean value is presented for comparison between the two fields strenght (**Figure 3**). The entropy mean value was higher on 3T versus 1.5T by a factor of four (**Figure 3**).

Influence of Field Strength on Heterogeneous Phantoms

The majority of texture features were significantly different between the two magnetic fields (1.5T vs. 3T) according to the Student's paired t-test (**Table 2**). Concerning features from histogram analysis, only Kurtosis, Entropy, and Energy did not significantly differ beteween 1.5 and 3T. For the matrix-based texture features, only LZHGE shows no significant difference.

Influence of Pixel Size

The pixels size, dependant on both matrix size and FOV size, altered the radiomics output in homogeneous phantom according to three behaviors. First of all, pixel size altered the values of the texture features on both the 1.5 and 3T magnetic field, for 15 indices: stdvalue, skewness, Homogeneity, Contrast, Correlation, Entropy, Dissimilarity, ZLNU, GLNU, RLNU, Coarseness, SZE, LZE, GLNU_1, and ZP. **Figure 4A** illustrates the "Correlation" named feature. Second, pixel size altered the value of texture features only on 1.5T, but not on 3T: Kurtosis, Entropy H, Energy H, Energy, LRE, RP. **Figure 4B** illustrates "Energy" named feature. Note that in contrast "Contrast_1" named feature value was isolatedly modified only on 3T. Thirdly, pixel size did not significantly alter the value of texture features for: minvalue, meanvalue, maxvalu, HGRE, SRHGE, LRHGE, HGZE, SZHGE, and LZHE.

The impact of matrix size and FOV as well as pixel size, on the radiomics output was also studied with the heterogeneous phantoms. Also three behaviors were observed. First, pixel size altered the values of the texture features on the two magnetic fields for: RLNU, Coarseness, SZLGE, GLNU and ZLNU. **Figure 5A** shows the example of "GLNU" named feature. Second, pixel size altered the value of texture features only on 1.5T, but not on 3T for: Correlation, LGRE, SRLGE, GLNU, and LGZE. **Figure 5B** shows an example with the "SRLGE" named feature. Third, pixel size did not significantly alter the value of texture features for the other 29 texture parameters.

Ability of texture features to identify a difference between phantom tubes, in both field sthrengh separatly, regardless of the pixel size, is observed through the absolute Spearman's rho correlation coefficient (**Table 3**). Eight imaging features

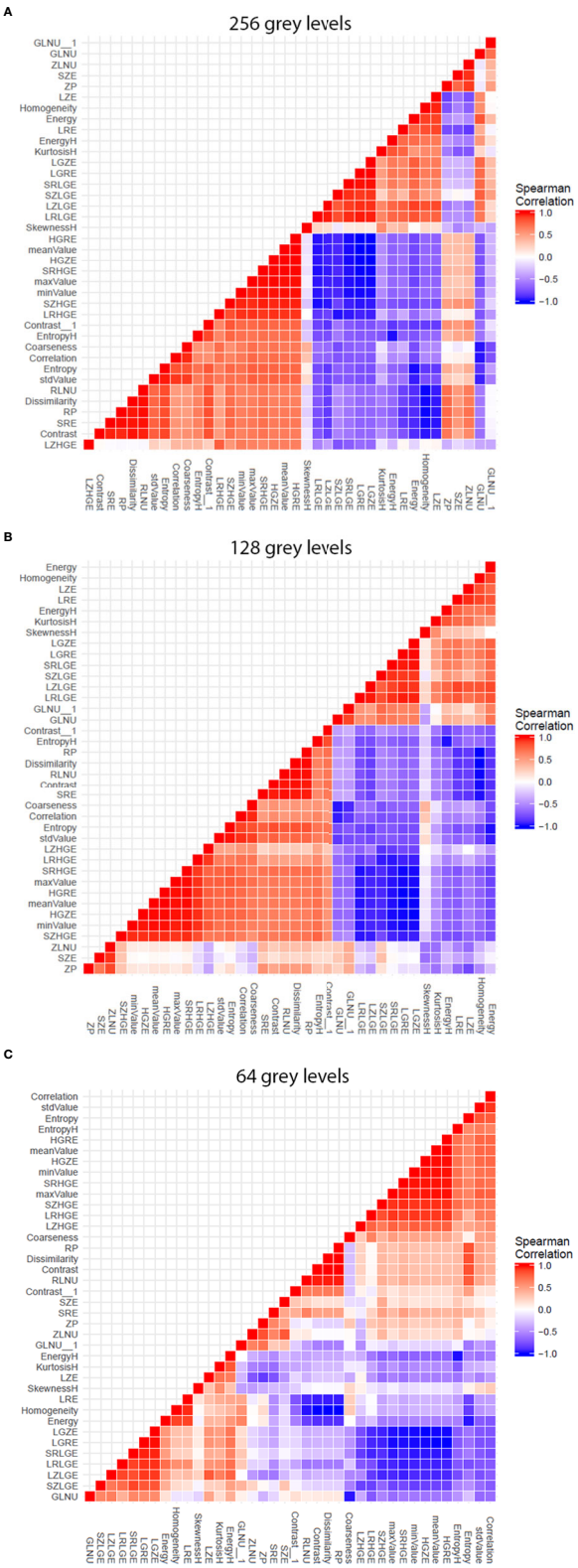


FIGURE 2 | Graphs showing influence of software resampling step on phantom study. Each graph representes the two-by-two coefficient correlation calculation between the texture parameters at one resampling level. The difference in resulting patterns shows the influence of the software resampling step on texture parameters calculation, particularly on parameters extracted from co occurrence matrix. **(A)** 256 gray levels resampling. **(B)** 128 gray levels resampling. **(C)** 64 gray levels resampling.

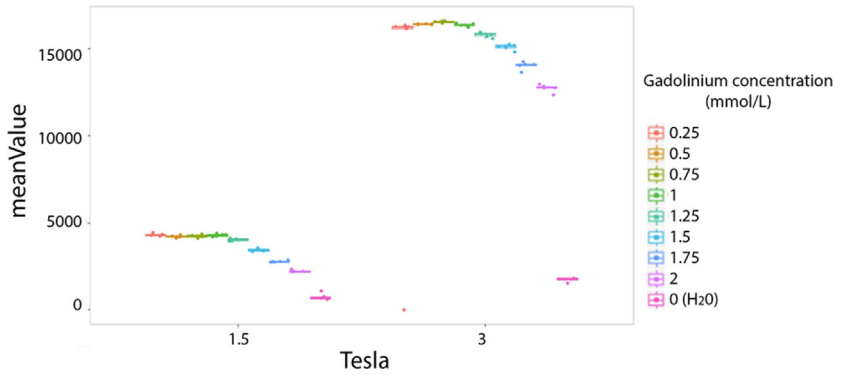


FIGURE 3 | Difference of phantoms textures features values at 1.5 and 3T. Example of the mean value. Each colored dot groups shows the mean value calculated from repeated phantom MRI acquisition (one dot per MRI acquisition) in different homogeneous tubes with different concentration of Gadolinium chelate (on color per tube), respectively at 1.5 and 3T.

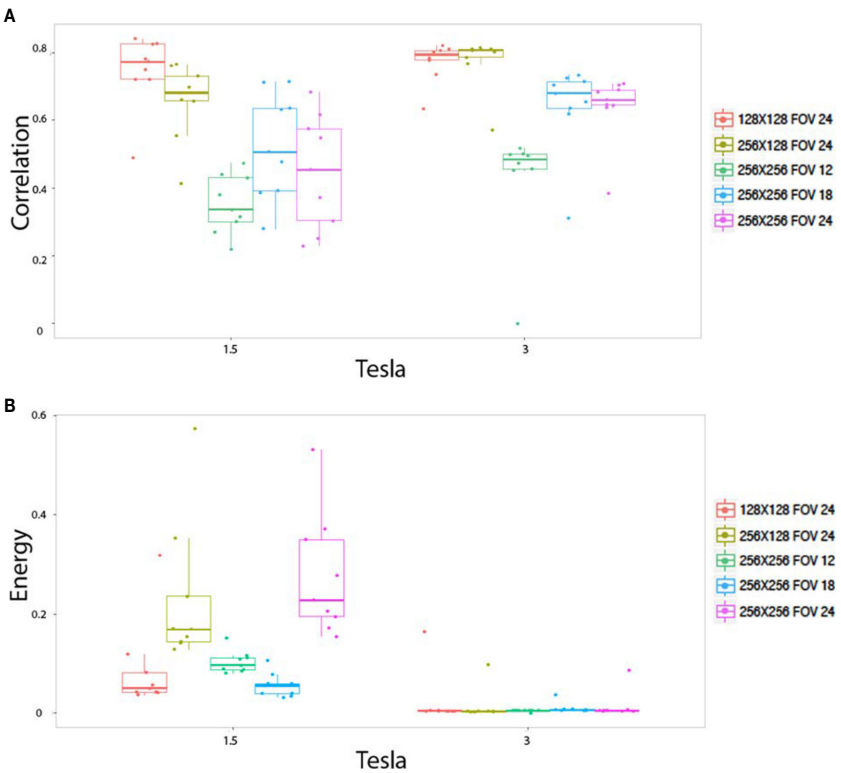


FIGURE 4 | *In vitro* parameters value according to magnetic field strengh and pixel size: For correlation **(A)** and energy **(B)**. Texture features extracted to the homogeneous phantoms. Each colored dot groups and related box plot shows the value of correlation **(A)** and energy **(B)** calculated from repeated phantom MRI acquisition (one dot per MRI acquisition) with different pixel size (one color per matrix/FOV couple), respectively, at 1.5 and 3T.

identified the (visually obvious) difference between homogeneous and heterogeneous phantoms, with an absolute Spearman's rho correlation coefficient above 0.5. The following texture features identified the difference between the different heterogeous phantoms (mainly different according to their polystyrene beads size and spatial distribution) in both field sthrengh separatly: dissimilarity, LZHGE, entropy, homogeneity, SRE, SZE, LRE, GLNU, RLNU, ZLNU, KurtosisH, LZE, entropyH, coarseness, SRLGE, SRLGE, SZLGE, and LGRE. Two texture features (coarseness, RLNU) identified the difference between homogeneous tubes (mainly different by their Gadolinium chelate concentration). As long as the pixel size remains sufficiently tiny

TABLE 2 | Texture features presenting significant or non-significant differences at 1.5 and 3T based on the heterogeneous phantoms examination (student's paired t-test).

Significant difference in texture features value at 1.5 and 3T (p-values < 10^{-5})	LGRE, Homogeneity, SZE, ZP, Contrast, LGZE, Correlation, RLNU Entropy, SRLGE, LRHGE, HGRE, mean value, SRHGE, LRE, GLNU, HGZE, GLNU, SRE, SZHGE, SZLGE, LRLGE, max value, Dissimilarity Contrast, RP, std value, LZE, ZLNU, skewness, H min, value Coarseness, LZLGE
Non-significant difference in texture features value at 1.5 and 3Tp-values > 0.005	Energy H, Entropy H, Kurtosis H, LZHGE, Energy H

small (1 mm or less in our study), 26 features were sensitive to texture alteration.

Effect of Field Strength on MRI Texture Features Values in Healthy Volunteers

Among the 38 parameters, significant differences were observed with 15 texture features measured in healthy volunteers between 1.5 and 3T MRIs for the same anatomical region (**Figure 6**). With constant fields (1.5T vs. 1.5T and 3T vs. 3T), those 15 texture features values appeared identical when measured in symmetrical anatomical structures (eg corpus calosum or caudate nucleus).

DISCUSSION

Our study demonstrates that field strength (1.5T vs. 3T) influences numerous texture features values for both phantom and human exploration. Our study also demonstrated that gray-level resampling and pixel size influence some texture features. Those results are of importance for clinical study design and for patient examination and follow-up. In addition to that, our study identifies texture parameters that are able to differentiate phantom textures in our setup. It is important to mention that these relevant textures parameters are significantly influenced by field strength.

An increasing number of radiomics or deep-learning studies use imaging based on MRI acquisitions (24–28). Confounding parameters that can alter radiomic output have been largely addressed for CT scans but less for MRIs (27, 29–31). Our study aims to complete knowledge in the field of radiomics applied to neuroradiology MRI acquisitions.

The presented study identifies the importance to consider the impact of magnetic field strength on radiomics features values in clinical practice. SNR is well-known to increase when the field strength increases. We hypothesize that our result may be partially due to this difference in SNR between the two MRI machines. Even given that field strength is a major factor

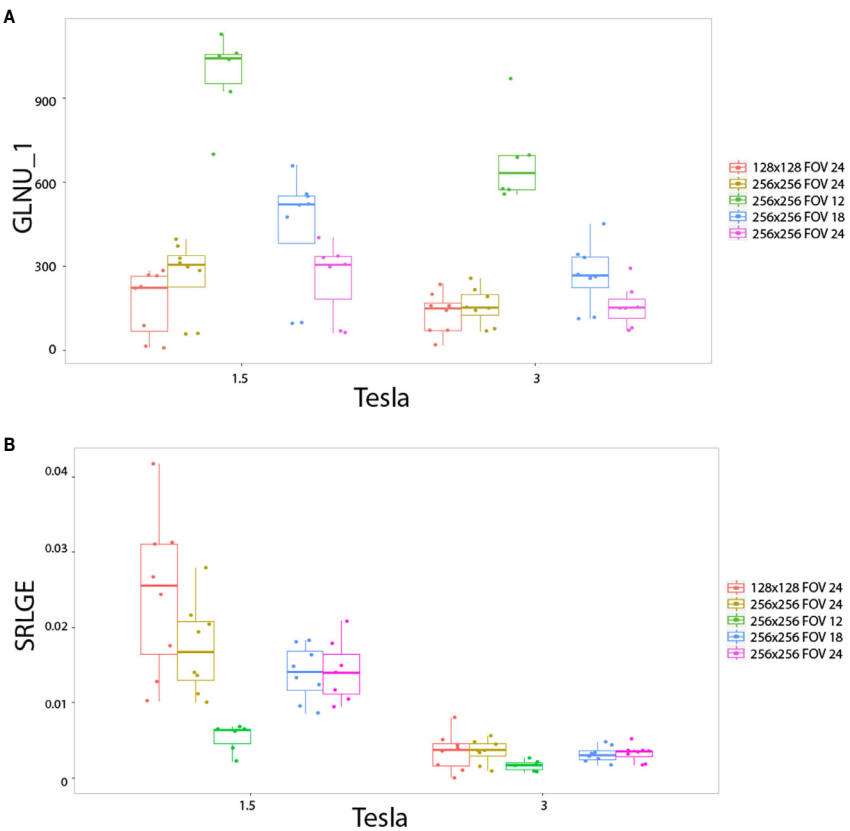


FIGURE 5 | *In vitro* parameters value according to magnetic field strengh and pixel size: for GLNU_1 **(A)** and SRLGE **(B)**. Texture features extracted to the heterogeneous phantoms. Each colored dot groups and related box plot shows the value of GLNU_1 **(A)** and SRLGE **(B)** calculated from repeated phantom MRI acquisition (one dot per MRI acquisition) with different pixel size (one color per matrix/FOV couple), respectively, at 1.5 and 3T.

TABLE 3 | Tabulation of textures features and their ability to dissociate the different phantom's tubes at both field strengths separately, regardless to pixel size.

Heterogeneous vs. Homogeneous media		Size of polystyrene		Gadolinium concentration	
RLNU	0.796	Contrast	0.801	Coarseness	0.577
Coarseness	0.795	Contrast_1	0.795	RLNU	0.513
minValue	0.683	Dissimilarity	0.794	minValue	0.482
GLNU 1	0.617	LZHGE	0.789	GLNU	0.414
SZLGE	0.552	stdValue	0.786	SZLGE	0.371
SRLGE	0.534	Entropy	0.783	SRLGE	0.354
Entropy	0.645	Homogeneity	0.773	GLNU	0.347
Homogeneity	0.622	ZP	0.77	EntropyH	0.33
ZLNU	0.491	SRE	0.768	meanValue	0.328
EntropyH	0.484	RP	0.762	LGZE	0.325
LGRE	0.481	Energy	0.756	LGRE	0.315
LGZE	0.477	SZE	0.756	EnergyH	0.308
GLNU	0.468	LRE	0.741	KurtosisH	0.302
Contrast	0.458	GLNU	0.738	SkewnessH	0.293
EnergyH	0.439	RLNU	0.732	Correlation	0.262
Contrast	0.429	minValue	0.727	maxValue	0.246
Dissimilarity	0.423	ZLNU	0.698	Contrast	0.246
stdValue	0.421	KurtosisH	0.677	ZLNU	0.237
Energy	0.418	LZE	0.646	LRLGE	0.228
meanValue	0.409	EntropyH	0.624	LRHGE	0.198
KurtosisH	0.405	Coarseness	0.617	Entropy	0.195
RP	0.378	SRLGE	0.583	Energy	0.182
SRE	0.373	EnergyH	0.554	Contrast	0.181
Correlation	0.363	SZLGE	0.554	Dissimilarity	0.18
SZE	0.346	LGRE	0.524	LZLGE	0.173
LRE	0.329	LGZE	0.46	Homogeneity	0.165
ZP	0.329	LRHGE	0.373	stdValue	0.155
LRHGE	0.316	LZLGE	0.321	HGZE	0.15
LRLGE	0.3	meanValue	0.275	HGRE	0.145
SkewnessH	0.229	SkewnessH	0.267	RP	0.141
maxValue	0.228	LRLGE	0.244	SZHGE	0.137
LZHGE	0.21	maxValue	0.112	SRE	0.132
HGZE	0.185	HGRE	0.101	SRHGE	0.132
HGRE	0.174	GLNU	0.092	SZE	0.132
SRHGE	0.14	HGZE	0.087	LRE	0.111
SZHGE	0.132	SRHGE	0.041	ZP	0.104
LZLGE	0.097	SZHGE	0.037	LZE	0.077
LZE	0.096	Correlation	0.017	LZHGE	0.01

Given are the names of the feature variables and the absolute Spearman's rho correlation coefficient between imaging features and the type of phantom tube. First column: ability to dissociate homogeneous versus heterogeneous tubes. Second column: ability to dissociate heterogeneous tubes according to their polystyrene beads sizes and their spatial distribution. Third column: ability to dissociate homogeneous tubes according to their Gadolinium chelate concentration.

influencing signal intensity, further factors modifying signal have to be considered, as discussed in a recent literature review. Mayerhoefer et al. previously demonstrated that MRI texture features are influenced by change in SNR (5, 9). The difference in SNR could be explained by magnetic field strength, but SNR is also linked to the entire signal acquisition system (coils, electronic device etc.). The European research project COST B11 (31, 32), aimed at developing quantitative methods for MRI imaging features extraction (between 1998 and 2002), demonstrated that signal and subsequently texture features were dependent on parameters such as configuration of the transmitting and/or receiving RF coil (antenna), and the number of active segments in the coil due to change in tilt angle.

Our results showed that most texture features are also sensitive to the variations in matrix and FOV and consequently spatial resolution, even if the clinical ranges in matrix size and FOV for a dedicated organ may be thinner than those investigated in literature physics studies. Of note, we acquired images at different pixel sizes whereas most studies changed the pixel size during post-processing. Our study showed that as long as the pixel size remains sufficiently small (1 mm in our experience), 26 features are sensitive to texture alteration. This is in accordance with Jirák study. which also compared MRI features measured on phantoms with nodular patterns in a multicenter settings and showed that texture classification was influenced by low resolution causing large errors (10).

For a defined magnetic field strength, our study showed that part of the texture features differentiated the heterogeneous phantoms from the homogeneous phantoms on all sequences, irrespective of the pixel size, and part of radiomics features differentiated phantoms with large and small spherical objects (polystyrene beads) scattered in agarose gel. These results are in agreement with literature (8, 10, 28, 29, 33, 34).

In our study, gray level resampling has a significant impact on radiomic parameters. This observation is in accordance with literature. Collewet et al. resampled gray scale using four methods: original gray levels, same maximum for all images, same average for all images, and limited dynamics. Results also

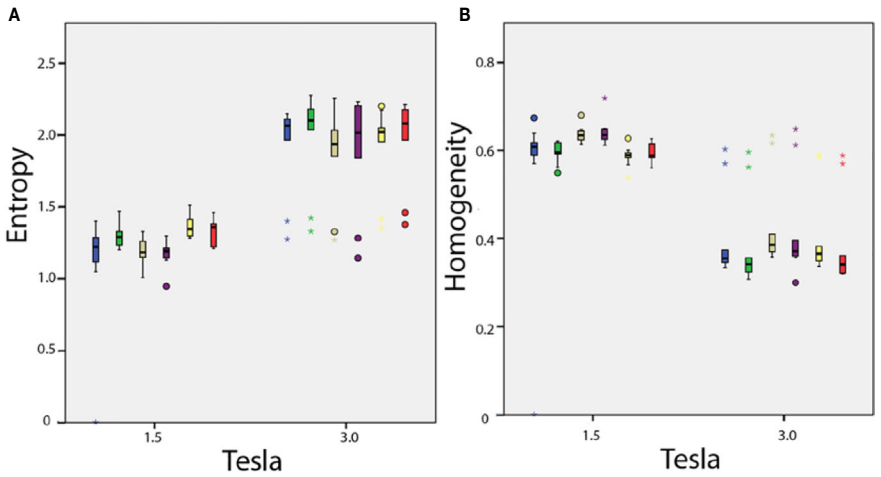


FIGURE 6 | *In vivo* parameters value difference according to magnetic field strength: **(A)** entropy and **(B)** homogeneity. Each box plot represents the values dispersion of the 10 healthy volunteers for one of the six anatomical regions.

showed the influence of the normalization method on the texture classification accuracy.

While our study focuses on field strength, pixel size and gray level resampling, previous studies have identified further MRI sequence parameters of interest, including repeat time (TR), echo time (TE) and receiver bandwidth (BW). Mayerhoefer et al. evaluated how texture features varied according to acquisition parameters (e.g., TR, TE, sampling bandwidth [SBW], and spatial resolution). Interestingly, variations in TR, TE and SBW had little effect on pattern discrimination results, as long as a high-spatial resolution was observed. In addition, both studies (Mayerhoefer's and ours) showed higher stability with the texture features derived from the co-occurrence matrix over first order texture features for characterizing patterns close to the resolution limits in varying conditions of TR/TE or pixel size.

Our approach to evaluate measures of radiomic features on standardized phantoms modeled to the organ under investigation is in line with the study by Bianchini et al. who explicitely designed and used an organ specific phantom for the optimization of radiomic studies of the female pelvis (35).

The question of the stability of our phantoms during the delay between the different acquisition/machine can be raised. Jirák et al. evaluated the long-term stability of the agarose phantoms, the optimal choice of texture parameters, and compared different MRI magnetic fields (3T, 4T, and 7T) (8, 10). They demonstrated that phantoms are stable over 12 months. In our study, all sequences were acquired on the same day within the same hour of examination, consequently we did not expect to incur phantom stability issue in our setup.

Optimal methodological guidelines are currently being defined to optimize data analysis strategies (4, 36–38) in the field of radiomics. Our results are critical since they demonstrate that standardization strategies for the use of radiomics in MRI should focus on addressing the challenge of heterogeneity in protocols since there is a dramatical difference between radiomics features extracted from the same patient or the same phantom at 1.5 or 3T.

Our study demonstrates that field strength had a strong influence on texture feature values, also spatial resolution and gray scale resampling. Quantification of the impact of the various parameters of imaging features is of major interest. Our findings are clinically relevant as image acquisition was performed according to daily used protocols for clinical examinations. Those confounding factors need to be adjusted when designing a multicentric trial and when adapting the results of a study to different platforms.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This study was approved by the Scientific Review Committee of Gustave Roussy.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.541663/full#supplementary-material

REFERENCES

1. Savadjiev P, Chong J, Dohan A, Vakalopoulou M, Reinhold C, Paragios N, et al. Demystification of AI-driven medical image interpretation: past, present and future. *Eur Radiol* (2019) 29:1616–24. doi: 10.1007/s00330-018-5674-x
2. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp* (2018) 2:36. doi: 10.1186/s41747-018-0068-z
3. Sun R, Orlhac F, Robert C, Reuzé S, Schernberg A, Buvat I, et al. In Regard to Mattonen et al. *Int J Radiat Oncol Biol Phys* (2016) 95:1544–5. doi: 10.1016/j.ijrobp.2016.03.038
4. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141
5. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* (2017) 28:1191–206. doi: 10.1093/annonc/mdx034
6. Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep* (2015) 5:11044. doi: 10.1038/srep11044
7. Sun R, Limkin EJ, Dercle L, Reuzé S, Zacharaki EI, Chargari C, et al. [Computational medical imaging (radiomics) and potential for immuno-oncology]. *Cancer Radiother* (2017) 21:648–54. doi: 10.1016/j.canrad.2017.07.035
8. Lerski RA, Schad LR, Luypaert R, Amorison A, Muller RN, Mascaro L, et al. Multicentre magnetic resonance texture analysis trial using reticulated foam test objects. *Magn Reson Imaging* (1999) 17:1025–31. doi: 10.1016/S0730-725X(99)00034-X
9. Mayerhoefer ME, Szomolanyi P, Jirak D, Materka A, Trattnig S. Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: an application-oriented study. *Med Phys* (2009) 36:1236–43. doi: 10.1118/1.3081408
10. Jirák D, Dezortová M, Hájek M. Phantoms for texture analysis of MR images. *Long-term Multi-center Study Med Phys* (2004) 31:616–22. doi: 10.1118/1.1646231
11. Hajek M, Dezortova M, Materka A, Lerski R. *Texture analysis for magnetic resonance imaging*. Prague: Med4publishing (2006).
12. Orlhac F, Nioche C, Soussan M, Buvat I. Understanding Changes in Tumor Texture Indices in PET: A Comparison Between Visual Assessment and Index Values in Simulated and Patient Data. *J Nucl Med* (2017) 58:387–92. doi: 10.2967/jnumed.116.181859
13. Shiri I, Rahmim A, Ghaffarian P, Geramifar P, Abdollahi H, Bitarafan-Rajabi A. The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies. *Eur Radiol* (2017) 27:4498–509. doi: 10.1007/s00330-017-4859-z
14. Carré A, Klausner G, Edjlali M, Lerousseau M, Briend-Diop J, Sun R, et al. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Sci Rep* (2020) 10:12340. doi: 10.1038/s41598-020-69298-z
15. Simmons A, Tofts PS, Barker GJ, Arridge SR. Sources of intensity nonuniformity in spin echo images at 1.5 T. *Magn Reson Med* (1994) 32:121–8. doi: 10.1002/mrm.1910320117
16. Antunes J, Viswanath S, Rusu M, Valls L, Hoimes C, Avril N, et al. Radiomics Analysis on FLT-PET/MRI for Characterization of Early Treatment Response in Renal Cell Carcinoma: A Proof-of-Concept Study. *Transl Oncol* (2016) 9:155–62. doi: 10.1016/j.tranon.2016.01.008
17. Chirra P, Leo P, Yim M, Bloch BN, Rastinehad AR, Purysko A, et al. Multisite evaluation of radiomic feature reproducibility and discriminability for identifying peripheral zone prostate tumors on MRI. *J Med Imaging (Bellingham)* (2019) 6:024502. doi: 10.1117/1.JMI.6.2.024502
18. Moradmand H, Aghamiri SMR, Ghaderi R. Impact of image preprocessing methods on reproducibility of radiomic features in multimodal magnetic resonance imaging in glioblastoma. *J Appl Clin Med Phys* (2020) 21:179–90. doi: 10.1002/acm2.12795
19. Bologna M, Corino V, Mainardi L. Technical Note: Virtual phantom analyses for preprocessing evaluation and detection of a robust feature set for MRI-radiomics of the brain. *Med Phys* (2019) 46:5116–23. doi: 10.1002/mp.13834
20. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* (2017) 44:1050–62. doi: 10.1002/mp.12123
21. Rathore S, Akbari H, Rozycki M, Abdullah KG, Nasrallah MP, Binder ZA, et al. Radiomic MRI signature reveals three distinct subtypes of glioblastoma with different clinical and molecular characteristics, offering prognostic value beyond IDH1. *Sci Rep* (2018) 8:5087. doi: 10.1038/s41598-018-22739-2
22. Zinn PO, Singh SK, Kotrotsou A, Hassan I, Thomas G, Luedi MM, et al. A Coclinical Radiogenomic Validation Study: Conserved Magnetic Resonance Radiomic Appearance of Periostin-Expressing Glioblastoma in Patients and Xenograft Models. *Precision Medicine and Imaging* (2018) 24:6288–99. doi: 10.1158/1078-0432.CCR-17-3420
23. Yip SS, Aerts HJ. Applications and limitations of radiomics. *Phys Med Biol* (2016) 61:R150–66. doi: 10.1088/0031-9155/61/13/R150
24. Lerski RA, de Certaines JD. Performance assessment and quality control in MRI by Eurospin test objects and protocols. *Magn Reson Imaging* (1993) 11:817–33. doi: 10.1016/0730-725X(93)90199-N
25. Buvat I, Orlhac F, Soussan M. Tumor Texture Analysis in PET: Where Do We Stand? *J Nucl Med* (2015) 56:1642–4. doi: 10.2967/jnumed.115.163469
26. Orlhac F, Soussan M, Maisonobe JA, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med* (2014) 55:414–22. doi: 10.2967/jnumed.113.129858
27. Shukla G, Alexander GS, Bakas S, Nikam R, Talekar K, Palmer JD, et al. Advanced magnetic resonance imaging in glioblastoma: a review. *Chin Clin Oncol* (2017) 6:40. doi: 10.21037/cco.2017.06.28
28. Xi YB, Guo F, Xu ZL, Li C, Wei W, Tian P, et al. Radiomics signature: A potential biomarker for the prediction of MGMT promoter methylation in glioblastoma. *J Magn Reson Imaging* (2018) 47:1380–7. doi: 10.1002/jmri.25860
29. Dercle L, Ammari S, Bateson M, Durand PB, Haspinger E, Massard C, et al. Limits of radiomic-based entropy as a surrogate of tumor heterogeneity: ROI-area, acquisition protocol and tissue site exert substantial influence. *Sci Rep* (2017) 7:7952. doi: 10.1038/s41598-017-08310-5
30. Lao J, Chen Y, Li ZC, Li Q, Zhang J, Liu J, et al. A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. *Sci Rep* (2017) 7:10353. doi: 10.1038/s41598-017-10649-8
31. Leger S, Zwanenburg A, Pilz K, Lohaus F, Linge A, Zöphel K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* (2017) 7:13206. doi: 10.1038/s41598-017-13448-3
32. *Quantification of magnetic resonance image texture*. European Commission (1998).
33. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging* (2004) 22:81–91. doi: 10.1016/j.mri.2003.09.001
34. Waugh SA, Lerski RA, Bidaut L, Thompson AM. The influence of field strength and different clinical breast MRI protocols on the outcome of texture analysis using foam phantoms. *Med Phys* (2011) 38:5058–66. doi: 10.1118/1.3622605
35. Bianchini L, Botta F, Origgi D, Rizzo S, Mariani M, Summers P, et al. PETER PHAN: An MRI phantom for the optimisation of radiomic studies of the female pelvis. *Phys Med* (2020) 71:71–81. doi: 10.1016/j.ejmp.2020.02.003
36. Mongan J, Moy L, Kahn CEJr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A guide for authors and reviewers. *Radiol Soc North Am* (2020) 2:200029. doi: 10.1148/ryai.2020200029
37. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging* (2019) 46:2638–55. doi: 10.1007/s00259-019-04391-8
38. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative—feature definitions. (2016). eprint, 2019.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

MINI REVIEW
published: 04 February 2021
doi: 10.3389/fonc.2020.614008

Trimodality PET/CT/MRI and Radiotherapy: A Mini-Review

Pierre Decazes[1,2]*, *Pauline Hinault*[2], *Ovidiu Veresezan*[3], *Sébastien Thureau*[1,2,3], *Pierrick Gouel*[1,2] *and Pierre Vera*[1,2]

[1] Nuclear Medicine Department, Henri Becquerel Cancer Center, Rouen, France, [2] QuantIF-LITIS EA4108, University of Rouen, Rouen, France, [3] Radiotherapy Department, Henri Becquerel Cancer Center, Rouen, France

Computed tomography (CT) has revolutionized external radiotherapy by making it possible to visualize and segment the tumors and the organs at risk in a three-dimensional way. However, if CT is a now a standard, it presents some limitations, notably concerning tumor characterization and delineation. Its association with functional and anatomical images, that are positron emission tomography (PET) and magnetic resonance imaging (MRI), surpasses its limits. This association can be in the form of a trimodality PET/CT/MRI. The objective of this mini-review is to describe the process of performing this PET/CT/MRI trimodality for radiotherapy and its potential clinical applications. Trimodality can be performed in two ways, either a PET/MRI fused to a planning CT (possibly with a pseudo-CT generated from the MRI for the planning), or a PET/CT fused to an MRI and then registered to a planning CT (possibly the CT of PET/CT if calibrated for radiotherapy). These examinations should be performed in the treatment position, and in the second case, a patient transfer system can be used between the PET/CT and MRI to limit movement. If trimodality requires adapted equipment, notably compatible MRI equipment with high-performance dedicated coils, it allows the advantages of the three techniques to be combined with a synergistic effect while limiting their disadvantages when carried out separately. Trimodality is already possible in clinical routine and can have a high clinical impact and good inter-observer agreement, notably for head and neck cancers, brain tumor, prostate cancer, cervical cancer.

Keywords: computed tomography, magnetic resonance imaging, positron emission tomography, radiotherapy, hybrid imaging

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Antoine Schernberg, Hôpital Tenon AP-HP, France
Hamid Mammar, Institut Curie, France
Alexandre Escande, Oscar Lambret Cancer Center, France

***Correspondence:**
Pierre Decazes
pierre.decazes@chb.unicancer.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *04 October 2020*
Accepted: *22 December 2020*
Published: *04 February 2021*

Citation:
Decazes P, Hinault P, Veresezan O, Thureau S, Gouel P and Vera P (2021) Trimodality PET/CT/MRI and Radiotherapy: A Mini-Review. Front. Oncol. 10:614008. doi: 10.3389/fonc.2020.614008

INTRODUCTION

External radiotherapy consists of treating an internal lesion, superficial and/or external lesion with an external source of radiation. In the nineties, computed tomography (CT) has revolutionized external radiotherapy by making it possible to visualize the tumor(s), corresponding to the target volume, and the organs at risk (OARs), which are normal tissues whose sensitivity to radiation can significantly influence treatment planning and/or prescribed dose. Because the treatment beams can be afterwards individually oriented on the tumor in a three-dimensional (3D) approach, this marked the beginning of 3D conformal radiotherapy (1).

A following improvement was the control of the intensity of the treatment beams which opened in the new millennium the era of the intensity-modulated radiotherapy (IMRT) whose aim is to deliver a high dose to the target volume while sparing the adjacent tissues, notably OARs (2). With this technique, a homogeneous dose is prescribed to the planning target volume (PTV) that considers uncertainty in treatment planning by encompassing the gross tumor volume (GTV, corresponding to the delineated macroscopic and radiologically measurable tumor) and the clinical target volume (CTV, which adds a margin to the GTV to cover nearby areas at risk of hosting microscopic disease) (3). The accuracy of anatomical localization is of particular importance for stereotactic radiotherapy (SRT), corresponding to an external beam radiotherapy used to deliver a high dose of radiation very precisely, as a single dose or a small number of fraction (4).

However, if CT imaging is a now a standard for radiotherapy, it presents some limitations, notably concerning tumor delineation which can be difficult, especially for soft tissues (3) or for the characterization of the lesions. Other 3D imaging modalities have therefore emerged for radiotherapy, in particular, positron emission tomography (PET) and magnetic resonance imaging (IRM). PET and MRI can notably visualize biological processes distinct and complementary to purely anatomical imaging. This led to the concept of biological target volume (BTV) focused on a metabolic function. For example, a boost radiotherapy can be performed on hypoxic tumors more resistant to radiation, identified by ^{18}F-fluoromisonidazole (FMISO) PET/CT (5). Finally, CT, PET, and MRI can be used to follow the patient during the radiotherapy, at the end of the treatment, or during the follow-up (6)

If these 3D imaging modalities (CT, PET, and MRI) can be considered separately, they can also be associated to form a hybrid imaging, two by two (PET/CT, CT/MRI, PET/MRI) but also as a trimodality PET/CT/MRI. Multimodality is already possible in clinical routine with a high clinical impact and good inter-observer agreement (7).

The aim of this mini-review is to present the concept of PET/CT/MRI trimodality, its rationale for radiotherapy, and its potential interest in characterizing tumor, performing treatment planning, and doing ART.

TECHNICAL PARTS

CT, MRI, and PET

CT is the reference imaging used by radiation oncologists for target volumes and organs at risk delineation for radiotherapy treatment planning. It is a high spatial resolution imaging modality that provides anatomical information with good spatial accuracy which is unaffected by geometric distortions. CT also provides a mapping of tissue electron density necessary for dosimetric calculations in radiotherapy. However, CT is an irradiating imaging modality with certain disadvantages such as lack of contrast in soft tissue and artifacts due to the presence of metal (8).

MRI is an anatomical and functional imaging modality that provides very good soft tissue contrast with millimetric spatial resolution. Although it has the advantage of being non-irradiating, the acquisition process is time consuming and this technique presents many contraindications (9). The possibility of using only MRI for radiotherapy treatment planning is however limited by the absence of information on tissue electronic density, a non-constant intensity of the images, and the presence of geometric distortions that deform images, including the volumes of interest.

CT and MRI are often associated with PET, a functional imaging modality. It provides a very good tumor/node contrast and the possibility to acquire large field of view. However, it is an irradiating examination with poor spatial resolution. In addition, the presence of partial volume artifacts creates blurred edges making more difficult the segmentation of volumes of interest.

Therefore, trimodality appears to be a technique of choice in the treatment of cancer in radiotherapy. It provides anatomical and functional information of high spatial resolution and allows improving the definition of target volumes in radiotherapy (10–12). A summary of the advantages and disadvantages of CT, MRI, and PET separately and combined in PET/CT/MRI is presented in **Table 1**.

Trimodality PET/CT/MRI

As trimodality allows obtaining additional information on disease and tumors; images of each modality must be performed in radiotherapy treatment position (13). Each machine is equipped with a rigid table that is positioned on

TABLE 1 | summary of the advantages and disadvantages of CT, MRI and PET separately and combined in PET/CT/MRI.

	CT	MRI	PET	Trimodality PET/CT/MRI
Advantages	Anatomy Spatial resolution (1 mm) Fast acquisition	Anatomy and function Spatial resolution (1 mm) Contrast (soft tissue) Non-irradiating	Function Tumor/Background Contrast Acquisition field	Anatomy and function Spatial resolution Tumor characterization Assessment of disease spread Dosimetry for RT
Limitations	Irradiating Contrast Artefacts (metal, teeth, *etc*.)	Long acquisition Compatible MRI equipment with high-performance dedicated coils Contraindications Artefacts (Distortions, no uniformity, *etc*.)	Irradiating Spatial resolution (>3–4 mm) Partial volume (blurred edges)	Irradiating Long acquisition Image registration Compatible MRI equipment with high-performance dedicated coils Contraindications MRI

the device table. Each of these rigid tables has markers and an indexing system that allows fixing radiotherapy immobilization solutions. In order to have exactly the same position for all acquisitions, it is essential that the equipment used is compatible with all the installations and in particular non-magnetic equipment for MR. Patient repositioning on each device is done using markers on the skin (*i.e.* the positioning referential) made during the planning CT and external positioning lasers.

Currently, no medical device allows simultaneous acquisition of all three imaging modalities. The solution is to use two separate imaging devices, a bimodal hybrid machine and an independent machine. Two trimodality systems are possible: a PET/MRI coupled with a CT or a PET/CT coupled with an MRI (14). For each of them, image registration will be indispensable to delineate volumes of interest and to perform radiotherapy treatment planning (15).

For the solution with PET/MR, precautions must be taken for data acquisition and processing. The patient is positioned on the device with MR coils compatible with radiotherapy immobilization fixations (16). To perform attenuation correction on the PET image, an attenuation mapping of MR coils must be performed before (17). The PET/MR images and the planning CT are then registered before volume delineation and dosimetric planning. An alternative to this solution is to replace the planning CT by a synthetic CT, commonly referred as pseudo-CT (18, 19). With the emergence of artificial intelligence, new robust algorithms such as GANs (Generative Adversial Networks) (20, 21) allow the creation of attenuation maps, synthetic CT, from the different MR images. Treatment planning can then be performed without proceeding to the image registration step.

The second PET/CT + MRI solution is performed following the same process; the patient has these two examinations one after the other in the radiotherapy treatment position. Two techniques can be used for this PET/CT + MRI workflow. The first is to use a transfer system compatible with both imaging devices (22). This consists of an air cushion bed with low attenuation and a non-magnetic stretcher that allows the bed to be moved from one device to the other without moving the patient. The air-cushioned bed is placed on the rigid tabletop of the first imaging device, and the patient is positioned in a position in agreement with the positioning referential realized for planning CT. At the end of the acquisition, the patient on the air-cushioned bed is moved to the stretcher with the help of a suction system and is then transferred to the second imaging unit from the stretcher to the examination table using the same suction system. In the end, the system allows the realization of multi-modal acquisitions while keeping the patient in the same position.

For the second technique, the PET/CT and MRI images are also performed in radiotherapy planning conditions, but the patient stands up between the two acquisitions. The patient is positioned on the first imager in radiotherapy treatment conditions using markers determined during planning CT acquisition and external lasers. He is then positioned in the same conditions to the second device. A summary of the methods of achieving PET/CT/MRI trimodality for radiotherapy is presented in **Figure 1**.

Image registration is the last step in the trimodality process. By placing the acquisitions in a common coordinate system, image registration allows correlating the information of each modality and thus improving clinical interpretation (see **Figure 2**). The image registration can be considered by two complementary approaches. The first one is material-based and consists in carrying out all the acquisitions under exactly the same conditions. First of all, the patient keeps the same position for each acquisition with the same radiotherapy immobilization fixations (23). The acquisition parameters will also allow obtaining the best possible alignment of the images by keeping the table height and choosing the same slice thickness, the same acquisition plane or 3D acquisitions with a large field of view (24).

This first approach facilitates the second software-based approach. The image registration is done manually by a physician or with the help of an automatic registration algorithm. In this second case, it is necessary to first evaluate the accuracy of the algorithm used. Several studies have evaluated CT-MRI or trimodality registration algorithms, either on phantom or from patient data (25–28). The average errors obtained are

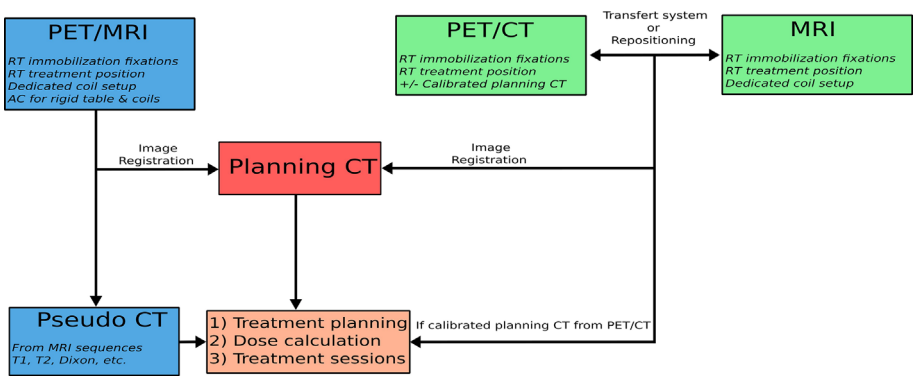


FIGURE 1 | Diagram of the process for performing the PET/CT/MRI trimodality in the radiotherapy treatment position.

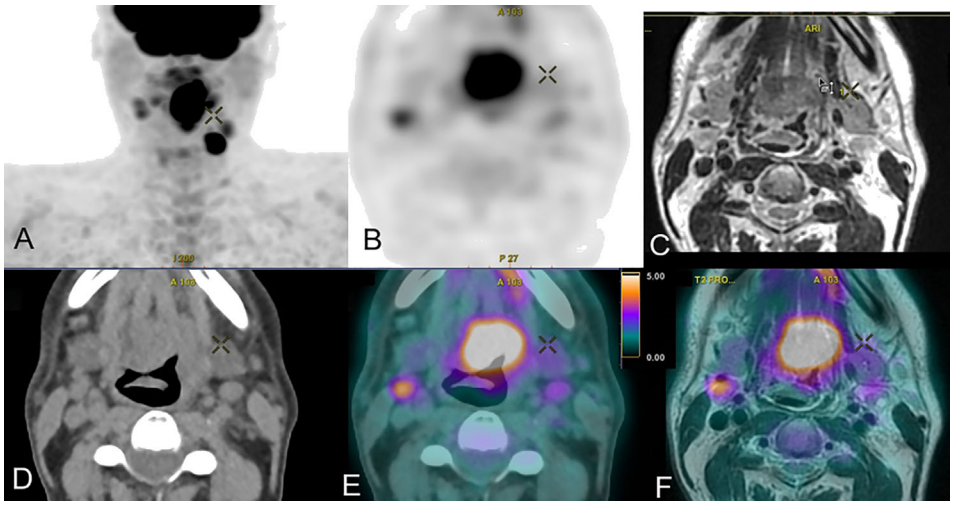


FIGURE 2 | Trimodal acquisition of a cancer of the base of the tongue with in **(A)** the frontal maximum intensity projection PET image, in **(B)** the axial PET FDG acquisition, in **(C)** the axial T2 MRI acquisition, in **(D)** the axial CT acquisition, in **(E)** the axial PET/CT fusion and in **(F)** the axial PET/MRI fusion.

between 0.4 and 2 mm. Eventually, a dedicated trimodality image fusion method can be used for better target delineation (29). A visual validation remains essential before proceeding with the planning of radiotherapy treatment, since humans are capable of detecting transformations of at least 1 mm and 1° (30). Image fusion is the last step in this process of trimodality; it allows correlating the information of each modality and improving the clinical interpretation (see **Figure 2**).

The registration can however be altered by the presence of artifacts, notably dental artifacts for CT (31, 32). Concerning MRI, the presence of geometric distortions, related to the system or the patient (33), can alter the registration. Algorithms can be used to correct distortions, but the presence of residual distortions has been mentioned (34). These distortions will affect the registration as well as the resulting treatment planning, in particular the target volume coverage in the case of stereotactic treatment (35, 36). In the context of multimodality for radiotherapy, the optimization of acquisition parameters is a crucial step to facilitate image registration.

CLINICAL APPLICATIONS

Characterization

Combining PET/CT/MRI could provide complementary information at baseline to assess the disease spread but also to guide biopsy and characterize the tumor to help decision making.

For glioma, many MRI sequences, notably perfusion and multiparametric and PET radiotracers, notably radiolabeled amino-acid like ^{18}F-FDOPA (FDOPA) and ^{18}F-FET (FET), are available (37). By their combination, it has been found that metabolic (FDOPA PET/CT) and anatomic (MRI) could aid in choosing the target to be biopsied under stereotactic conditions in tumors without MR enhancement (38). In glioblastoma, another study including multiparametric imaging with FET PET/CT and FDG PET/MRI (including diffusion and dynamic contrast enhanced perfusion) has shown that combining parameters in a multivariate model enabled patient-specific maps of recurrence probability, where FET was the most important parameter (39). Comparable results were observed in a study showing that combination of apparent diffusion coefficient (ADC) and FET was more accurate to detect glioma infiltration than standard MRI in enhancing gliomas (38); such an approach could allow risk-adapted radiotherapy planning (39). Finally, for brain metastasis, FDG PET associated with MRI could be interesting for lesion targeting of stereotactic radiation therapy in case of a previously irradiated recurrent tumor (40); place of FDOPA in this indication has yet to be evaluated.

For cervical cancer, low ADC value in MRI and high FDG SUV of the primary tumor are also predictive factors for identifying high-risk patients (41). More complex parameters with a radiomics analysis combining PET and MRI parameters have also been found interesting to improve the prognostic determination in locally advanced cervical cancer treated with chemoradiotherapy (42).

Finally, for lung cancers, combination of MRI and PET can help to determine the prognostic with poor issues when low ADC value in MRI and high FDG SUVmax in PET are associated before stereotactic body radiotherapy (43) or when parameters derived from DCE and FDG PET parameters are associated (44).

Planning

For Head and Neck Cancers, it has been known for a long time that image fusion between FDG PET and MRI/CT is useful with regard to the determination of GTV and CTV for 3D conformal radiotherapy with preservation of normal tissues for the majority of the cases (45), with an interest to do them in treatment position, notably in case of intra-cranial tumor extension, heavy metal dental work, or contraindication of contrast enhanced CT (46). A special irradiation setup including thermoplastic mask,

flat table, and head support can be used to allow a precise image co-registration of trimodality PET/CT MRI (47). As, according to pathological correlative series, no imaging modality completely encompasses the tumor, it can be useful to create a composite target volume derived from multi-modal imaging (48–51), notably for dose painting, consisting of delivering a heterogeneous irradiation within the tumor volume, as FDG PET and multiparametric MRI, in particular DWI, contain correlated but also independent information (47, 52–54). Concerning PET, other radiotracers than FDG can be used, such as FDOPA for skull base paraganglioma (55).

At the brain level, for gliomas treated with 3D conformal radiotherapy, or even by stereotactic radiotherapy by gamma-knife (9), the integration of both MRI and amino-acid PET/CT may help to improve GTV coverage by avoiding larger discrepancies between physical and biological imaging techniques (56–60), notably for the area of suspected non-enhancing tumor (61). Moreover for meningioma, ^{68}Ga-DOTATOC PET, exploring the expression of SST2 receptors, can be usefully associated with CT and MRI for the treatment planning (62), notably to detect and assess the extent of infracranial meningioma invasion (63) and with an impact on sparing normal tissues (64), although all locations do not benefit from this trimodality (65).

For lung cancer, improvements in radiotherapy techniques (IMRT) make it important to manage the definition of volume and its mobility (66). The delineation of tumor lesions in lung cancer patients based on PET/CT is advisable in radiotherapy treatment planning and for locally advanced non-small cell lung cancer treated with IMRT or SRT, PET/CT being regarded as an indispensable staging procedure. Respiratory gating techniques (4D PET/CT) optimize radiotherapy of lung cancer to reduce toxicities especially the pulmonary and cardiac late toxicities (67, 68). The place of trimodality PET/CT/MRI has however to be defined in this indication.

For cervical cancer, a study on a cohort of 134 patients has shown that dose delivered to the primary cervical tumor by the combination of MRI-guided high-dose-rate PET/CT-guided IMRT brachytherapy was highly correlated with local tumor control (69). However, the delineation stays difficult as tumor volume discrepancies are observed between MRI and PET/CT GTV (70) even if it could decrease inter-observer variability (71). If FDG PET/CT appears superior as a functional imaging modality when compared with DW MRI in tumor contouring (72), a threshold around 30% of the FDG SUVmax appears to provide the best segmentation for this cancer (73).

Concerning prostatic cancer, multimodality, in particular multiparametric MRI and ^{68}Ga-PSMA PET/CT, offers now large possibilities whatever the risk of the disease. In low-risk patients, selection of patients for active surveillance or treatment is improved; for intermediate-risk patients, it can help to select patients for supplemental brachytherapy; for high-risk patients it can help to guarantee adapted tumor volume segmentation, and finally, for recurrent or metastatic disease, it offers opportunities for more accurate assessment of tumor burden and treatment response (74). Therefore, it was found in a prospective study that combination of PSMA PET and multiparametric MRI provided a reliable TNM staging in patients with prostate cancer with a change in the therapeutic management for almost one third of the patients (75), PSMA PET being particularly interesting to delineate lymph node metastases (76). Moreover, for focal dose escalation to the dominant intraprostatic lesions, ^{68}Ga-PSMA PET/CT and multiparametric MRI provide concordant results for delineation in nearly 50% of the lesions, with a PET GTV significantly larger than MRI GTV and which could have a role in treatment planning with intraprostatic dose escalation (77), notably because dose distribution within dominant intraprostatic lesions defined by multiparametric MRI and/or PSMA PET imaging is an independent risk factor for biochemical failure after primary external beam radiation therapy (78). If further studies are needed to confirm the optimal imaging techniques (79), it is already possible to combine multiparametric MRI and PSMA PET with higher tumor control probability with minimal to no increase of normal tissue complication probability compared to dose escalation on GTV defined on only one imaging modality (80, 81). Moreover trimodality with PSMA PET/CT/MRI can be used to orient the therapy in case of biochemical relapse after treatment (82).

For rectal cancer, trimodality PET/CT/MRI has been known to be possible for a relatively long time, notably to allow dose escalation on primary tumor (83). As a mobile organ, non-rigid registration between PET/CT and MRI shows good results, but this must be considered for the treatment planning (83). FDG PET/CT adds therefore information to MRI, with potentially a larger GTV in total when using the union of MRI and PET, and new or differently evaluated lesions in as many as 15% of the patients, potentially changing the treatment (84); further studies are necessary to well define the place of FDG PET in this indication (85, 86).

Adaptative Radiotherapy

Functional and anatomical data can be used not only prior to treatment, but also during and after treatment to guide ART, by improving the tumor targeting while better sparing the OARs, as well as determine tumor response (87). If the ART approach has been based first historically on per-treatment CT and/or CBCT images, it is now possible with MRI-linear accelerator (MRI-linac), combining an MRI and a linear accelerator, allowing an MRI acquisition before each treatment delivery (88). PET-linac is also emerging even if it remains less mature than MRI-linac (89). While extremely promising, the utilization of functional adaptation in radiation therapy is only beginning and needs more prospective clinical validation (90).

PERSPECTIVES

If trimodality can already be used in some indications, its usefulness remains to be confirmed. To this goal, several clinical studies are in progress. Concerning prostate cancer, our team is exploring in the ongoing DEMETER study (NCT03734757), which will include 20 patients, the interest of the association of PET/CT with ^{18}F-choline and MRI compared

to standard initial staging (CT, MRI and bone scan) to determine radio-therapeutic volumes. Another recently opened phase 2 study (NCT04402151), which aims to include 50 patients, will explore the interest of combination of PSMA PET/MRI to radiation delivery with a MRI-Linac. This study is based on the principle that the combination of PET PSMA and MRI allows for better delineation of intraprostatic nodules and greater diagnostic accuracy for the detection of metastatic disease. Moreover, MRI-Linac also allows adaptive radiotherapy in addition to the planning. For cervical cancer, a prospective observational study including 237 patients (NCT01992861) is exploring the role of MRI, including DCE, DWI, and spectrometry, and FDG PET performed before, during and after radiotherapy and chemotherapy. These could help to predict patient's response to treatment and plan treatment. Finally, for head and neck cancers, our team is performing a prospective observational study with 60 inclusions planned called TRIMODAL (NCT03897166). Many questions will be explored in this study, including the comparison of the volumes determined on FDG PET/CT and on FDG PET/CT/MRI, the quality of image registration (in particular by using an air-cushion transfer system) and the use of algorithms for anthropometric measurements in MRI and CT scanners (with Dual x-ray absorptiometry as reference standard).

CONCLUSION

We have shown in this brief review the ways of carrying out a trimodality PET/CT/MRI for radiotherapy and potential clinical applications. Trimodality PET/CT/MRI, combining the strengths of the techniques and limiting the respective weaknesses, reinforces the role of imaging in guiding radiation therapy. Although this multimodality is recent, it is already possible in clinical routine with a high clinical impact and good inter-observer agreement. Clinical studies are still needed to confirm its role in clinical routine.

AUTHOR CONTRIBUTIONS

PD, PH, PV, OV, PG, ST helped in conceptualization and project administration. PD and PH contributed to data curation. PD and PH wrote the main manuscript text and prepared the figures. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Herrmann H, Seppenwoolde Y, Georg D, Widder J. Image guidance: past and future of radiotherapy. *Radiologe* (2019) 59:21–7. doi: 10.1007/s00117-019-0573-y
2. Bortfeld T. IMRT: a review and preview. *Phys Med Biol* (2006) 51:R363–79. doi: 10.1088/0031-9155/51/13/R21
3. Beaton L, Bandula S, Gaze MN, Sharma RA. How rapid advances in imaging are defining the future of precision radiation oncology. *Br J Cancer* (2019) 120:779–90. doi: 10.1038/s41416-019-0412-y
4. Tree AC, Khoo VS, Eeles RA, Ahmed M, Dearnaley DP, Hawkins MA, et al. Stereotactic body radiotherapy for oligometastases. *Lancet Oncol* (2013) 14: e28–37. doi: 10.1016/S1470-2045(12)70510-7
5. Vera P, Mihailescu S-D, Lequesne J, Modzelewski R, Bohn P, Hapdey S, et al. Radiotherapy boost in patients with hypoxic lesions identified by 18F-FMISO PET/CT in non-small-cell lung carcinoma: can we expect a better survival outcome without toxicity? [RTEP5 long-term follow-up]. *Eur J Nucl Med Mol Imaging* (2019) 46:1448–56. doi: 10.1007/s00259-019-04285-9
6. Decazes P, Thureau S, Dubray B, Vera P. How to use PET/CT in the evaluation of response to radiotherapy. *Q J Nucl Med Mol Imaging* (2018) 62:152–64. doi: 10.23736/S1824-4785.17.03033-3
7. Vogel WV, Lam MGEH, Pameijer FA, van der Heide UA, van de Kamer JB, Philippens ME, et al. Functional Imaging in Radiotherapy in the Netherlands: Availability and Impact on Clinical Practice. *Clin Oncol* (2016) 28:e206–15. doi: 10.1016/j.clon.2016.09.003
8. Richard P, Sandison G, Dang Q, Johnson B, Wong T, Parvathaneni U. Dental amalgam artifact: Adverse impact on tumor visualization and proton beam treatment planning in oral and oropharyngeal cancers. *Pract Radiat Oncol* (2015) 5:e583–8. doi: 10.1016/j.prro.2015.04.007
9. Rundo L, Stefano A, Militello C, Russo G, Sabini MG, D'Arrigo C, et al. A fully automatic approach for multimodal PET and MR image segmentation in gamma knife treatment planning. *Comput Methods Programs Biomed* (2017) 144:77–96. doi: 10.1016/j.cmpb.2017.03.011
10. Ak S, Kiliç C, Özlügedik S. Correlation of PET-CT, MRI and histopathology findings in the follow-up of patients with nasopharyngeal cancer. *Braz J Otorhinolaryngol* (2020) 20:30004–5. doi: 10.1016/j.bjorl.2019.12.004
11. Chopra S, Engineer R, Shah S, Shukla R, Dora T, Gupta P, et al. MRI- and PET-Guided Interstitial Brachytherapy for Postsurgical Vaginal Recurrences of Cervical Cancer: Results of Phase II Study. *Int J Radiat Oncol Biol Phys* (2020) 106:310–9. doi: 10.1016/j.ijrobp.2019.10.037
12. Yang J, Beadle BM, Garden AS, Schwartz DL, Aristophanous M. A multimodality segmentation framework for automatic target delineation in head and neck radiotherapy: Automatic target delineation in head and neck radiotherapy. *Med Phys* (2015) 42:5310–20. doi: 10.1118/1.4928485
13. Reynaert N. PET and MRI based RT treatment planning: Handling uncertainties. *Cancer Radiothér* (2019) 23:753–60. doi: 10.1016/j.canrad.2019.08.002
14. Paulus DH, Oehmigen M, Grüneisen J, Umutlu L, Quick HH. Whole-body hybrid imaging concept for the integration of PET/MR into radiation therapy treatment planning. *Phys Med Biol* (2016) 61:3504–20. doi: 10.1088/0031-9155/61/9/3504
15. Leibfarth S, Mönnich D, Welz S, Siegel C, Schwenzer N, Schmidt H, et al. A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning. *Acta Oncol* (2013) 52:1353–9. doi: 10.3109/0284186X.2013.813964
16. Winter RM, Leibfarth S, Schmidt H, Zwirner K, Mönnich D, Welz S, et al. Assessment of image quality of a radiotherapy-specific hardware solution for PET/MRI in head and neck cancer patients. *Radiother Oncol* (2018) 128:485–91. doi: 10.1016/j.radonc.2018.04.018
17. Paulus DH, Thorwath D, Schmidt H, Quick HH. Towards integration of PET/MR hybrid imaging into radiation therapy treatment planning. *Med Phys* (2014) 41:072505. doi: 10.1118/1.4881317
18. Wiesinger F, Bylund M, Yang J, Kaushik S, Shanbhag D, Ahn S, et al. Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning. *Magn Reson Med* (2018) 80:1440–51. doi: 10.1002/mrm.27134
19. Botman R, Tetar SU, Palacios MA, Slotman BJ, Lagerwaard FJ, Bruynzeel AME. The clinical introduction of MR-guided radiation therapy from a RTT perspective. *Clin Transl Radiat Oncol* (2019) 18:140–5. doi: 10.1016/j.ctro.2019.04.019
20. Peng Y, Chen S, Qin A, Chen M, Gao X, Liu Y, et al. Magnetic resonance-based synthetic computed tomography images generated using generative adversarial networks for nasopharyngeal carcinoma radiotherapy treatment

planning. *Radiother Oncol* (2020) 150:217–24. doi: 10.1016/j.radonc.2020.06.049
21. Maspero M, Savenije MHF, Dinkla AM, Seevinck PR, Intven MPW, Jurgenliemk-Schulz IM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Phys Med Biol* (2018) 63:185001. doi: 10.1088/1361-6560/aada6d
22. Veit-Haibach P, Kuhn FP, Wiesinger F, Delso G, von Schulthess G. PET-MR imaging using a tri-modality PET/CT-MR system with a dedicated shuttle in clinical routine. *MAGMA* (2013) 26:25–35. doi: 10.1007/s10334-012-0344-5
23. Samarin A, Kuhn FP, Brandsberg F, von Schulthess G, Burger IA. Image registration accuracy of an in-house developed patient transport system for PET/CT+MR and SPECT+CT imaging. *Nucl Med Commun* (2015) 36:194. doi: 10.1097/MNM.0000000000000229
24. Lotterie J-A, Duthil P, Januel A-C, Redon A, Menegalli D, Blond S, et al. [Stereotactic and diagnostic imaging in radiosurgery]. *Cancer Radiother* (2012) 16 Suppl:S10–25. doi: 10.1016/j.canrad.2011.09.005
25. Daisne J-F, Sibomana M, Bol A, Cosnard G, Lonneux M, Grégoire V. Evaluation of a multimodality image (CT, MRI and PET) coregistration procedure on phantom and head and neck cancer patients: accuracy, reproducibility and consistency. *Radiother Oncol* (2003) 69:237–45. doi: 10.1016/j.radonc.2003.10.009
26. Wang X, Li L, Hu C, Qiu J, Xu Z, Feng Y. A comparative study of three CT and MRI registration algorithms in nasopharyngeal carcinoma. *J Appl Clin Med Phys* (2009) 10:2906. doi: 10.1120/jacmp.v10i2.2906
27. Moore CS, Liney GP, Beavis AW. Quality assurance of registration of CT and MRI data sets for treatment planning of radiotherapy for head and neck cancers. *J Appl Clin Med Phys* (2004) 5:25–35. doi: 10.1120/jacmp.v5i1.1951
28. Roberson PL, McLaughlin PW, Narayana V, Troyer S, Hixson GV, Kessler ML. Use and uncertainties of mutual information for computed tomography/magnetic resonance (CT/MR) registration post permanent implant of the prostate. *Med Phys* (2005) 32:473–82. doi: 10.1118/1.1851920
29. Guo L, Shen S, Harris E, Wang Z, Jiang W, Guo Y, et al. A Tri-Modality Image Fusion Method for Target Delineation of Brain Tumors in Radiotherapy. *PloS One* (2014) 9:e112187. doi: 10.1371/journal.pone.0112187
30. Veninga T, Huisman H, van der Maazen RWM, Huizenga H. Clinical validation of the normalized mutual information method for registration of CT and MR images in radiotherapy of brain tumors. *J Appl Clin Med Phys* (2004) 5:66–79. doi: 10.1120/jacmp.v5i3.1959
31. Park HS, Choi JK, Park K-R, Kim KS, Lee S-H, Ye JC, et al. Metal artifact reduction in CT by identifying missing data hidden in metals. *J Xray Sci Technol* (2013) 21:357–72. doi: 10.3233/XST-130384
32. Lee M-Y, Song K-H, Lee J-W, Choe B-Y, Suh TS. Metal artifacts with dental implants: Evaluation using a dedicated CT/MR oral phantom with registration of the CT and MR images. *Sci Rep* (2019) 9:754. doi: 10.1038/s41598-018-36227-0
33. Fransson A, Andreo P, Pötter R. Aspects of MR image distortions in radiotherapy treatment planning. *Strahlenther Onkol* (2001) 177:59–73. doi: 10.1007/pl00002385
34. Weygand J, Fuller CD, Ibbott GS, Mohamed ASR, Ding Y, Yang J, et al. Spatial Precision in Magnetic Resonance Imaging-Guided Radiation Therapy: The Role of Geometric Distortion. *Int J Radiat Oncol Biol Phys* (2016) 95:1304–16. doi: 10.1016/j.ijrobp.2016.02.059
35. Putz F, Mengling V, Perrin R, Masitho S, Weissmann T, Rösch J, et al. Magnetic resonance imaging for brain stereotactic radiotherapy: A review of requirements and pitfalls. *Strahlenther Onkol* (2020) 196:444–56. doi: 10.1007/s00066-020-01604-0
36. Pappas EP, Alshanqity M, Moutsatsos A, Lababidi H, Alsafi K, Georgiou K, et al. MRI-Related Geometric Distortions in Stereotactic Radiotherapy Treatment Planning: Evaluation and Dosimetric Impact. *Technol Cancer Res Treat* (2017) 16:1120–9. doi: 10.1177/1533034617735454
37. Moreau A, Febvey O, Mognetti T, Frappaz D, Kryza D. Contribution of Different Positron Emission Tomography Tracers in Glioma Management: Focus on Glioblastoma. *Front Oncol* (2019) 9:1134. doi: 10.3389/fonc.2019.01134
38. Todeschi J, Bund C, Cebula H, Chibbaro S, Lhermitte B, Pin Y, et al. Diagnostic value of fusion of metabolic and structural images for stereotactic biopsy of brain tumors without enhancement after contrast medium injection. *Neurochirurgie* (2019) 65:357–64. doi: 10.1016/j.neuchi.2019.08.002
39. Lundemann M, Munck af Rosenschöld P, Muhic A, Larsen VA, Poulsen HS, Engelholm S-A, et al. Feasibility of multi-parametric PET and MRI for prediction of tumour recurrence in patients with glioblastoma. *Eur J Nucl Med Mol Imaging* (2019) 46:603–13. doi: 10.1007/s00259-018-4180-3
40. Scranton RA, Sadrameli S, Butler EB, Farach A, Wang H-C, Teh BS, et al. Coregistration of Magnetic Resonance and [18F] Fludeoxyglucose–Positron Emission Tomography Imaging for Stereotactic Radiation Therapy Planning: Case Report in a Previously Irradiated Brain Metastasis With Recurrent Tumor and Radiation Necrosis. *Pract Radiat Oncol* (2020) 10:133–7. doi: 10.1016/j.prro.2019.11.008
41. Akkus Yildirim B, Onal C, Erbay G, Cem Guler O, Karadeli E, Reyhan M, et al. Prognostic values of ADC $_{mean}$ and SUV $_{max}$ of the primary tumour in cervical cancer patients treated with definitive chemoradiotherapy. *J Obstet Gynaecol* (2019) 39:224–30. doi: 10.1080/01443615.2018.1492528
42. Lucia F, Visvikis D, Desseroit M-C, Miranda O, Malhaire J-P, Robin P, et al. Prediction of outcome using pretreatment 18F-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy. *Eur J Nucl Med Mol Imaging* (2018) 45:768–86. doi: 10.1007/s00259-017-3898-7
43. Iizuka Y, Matsuo Y, Umeoka S, Nakamoto Y, Ueki N, Mizowaki T, et al. Prediction of clinical outcome after stereotactic body radiotherapy for non-small cell lung cancer using diffusion-weighted MRI and 18F-FDG PET. *Eur J Radiol* (2014) 83:2087–92. doi: 10.1016/j.ejrad.2014.07.018
44. Huang Y-S, Chen JL-Y, Chen J-Y, Lee Y-F, Huang J-Y, Kuo S-H, et al. Predicting tumor responses and patient survival in chemoradiotherapy-treated patients with non-small-cell lung cancer using dynamic contrast-enhanced integrated magnetic resonance–positron-emission tomography. *Strahlenther Onkol* (2019) 195:707–18. doi: 10.1007/s00066-018-1418-8
45. Nishioka T, Shiga T, Shirato H, Tsukamoto E, Tsuchiya MDK, Kato T, et al. Image fusion between 18 FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* (2002) 53:1051–7. doi: 10.1016/S0360-3016(02)02854-7
46. Gardner M, Halimi P, Valinta D, Plantet M-M, Alberini J-L, Wartski M, et al. Use of single MRI and 18F-FDG PET-CT scans in both diagnosis and radiotherapy treatment planning in patients with head and neck cancer: Advantage on target volume and critical organ delineation. *Head Neck* (2009) 31:461–7. doi: 10.1002/hed.21005
47. Zwirner K, Thorwarth D, Winter RM, Welz S, Weiss J, Schwenzer NF, et al. Voxel-wise correlation of functional imaging parameters in HNSCC patients receiving PET/MRI in an irradiation setup. *Strahlenther Onkol* (2018) 194:719–26. doi: 10.1007/s00066-018-1292-4
48. Prestwich RJD, Sykes J, Carey B, Sen M, Dyker KE, Scarsbrook AF. Improving Target Definition for Head and Neck Radiotherapy: A Place for Magnetic Resonance Imaging and 18-Fluoride Fluorodeoxyglucose Positron Emission Tomography? *Clin Oncol* (2012) 24:577–89. doi: 10.1016/j.clon.2012.04.002
49. Bird D, Scarsbrook AF, Sykes J, Ramasamy S, Subesinghe M, Carey B, et al. Multimodality imaging with CT, MR and FDG-PET for radiotherapy target volume delineation in oropharyngeal squamous cell carcinoma. *BMC Cancer* (2015) 15:844. doi: 10.1186/s12885-015-1867-8
50. Ligtenberg H, Jager EA, Caldas-Magalhaes J, Schakel T, Pameijer FA, Kasperts N, et al. Modality-specific target definition for laryngeal and hypopharyngeal cancer on FDG-PET, CT and MRI. *Radiother Oncol* (2017) 123:63–70. doi: 10.1016/j.radonc.2017.02.005
51. Samołyk-Kogaczewska N, Sierko E, Zuzda K, Gugnacki P, Szumowski P, Mojsak M, et al. PET/MRI-guided GTV delineation during radiotherapy planning in patients with squamous cell carcinoma of the tongue. *Strahlenther Onkol* (2019) 195:780–91. doi: 10.1007/s00066-019-01480-3
52. Houweling AC, Wolf AL, Vogel WV, Hamming-Vrieze O, van Vliet-Vroegindeweij C, van de Kamer JB, et al. FDG-PET and diffusion-weighted MRI in head-and-neck cancer patients: Implications for dose painting. *Radiother Oncol* (2013) 106:250–4. doi: 10.1016/j.radonc.2013.01.003
53. Rasmussen JH, Nørgaard M, Hansen AE, Vogelius IR, Aznar MC, Johannesen HH, et al. Feasibility of Multiparametric Imaging with PET/MR in Head and Neck Squamous Cell Carcinoma. *J Nucl Med* (2017) 58:69–74. doi: 10.2967/jnumed.116.180091

54. Cardoso M, Min M, Jameson M, Tang S, Rumley C, Fowler A, et al. Evaluating diffusion-weighted magnetic resonance imaging for target volume delineation in head and neck radiotherapy. *J Med Imaging Radiat Oncol* (2019) 63:399–407. doi: 10.1111/1754-9485.12866
55. Helali M, Moreau M, Le Fèvre C, Heimburger C, Bund C, Goichot B, et al. 18F-FDOPA PET/CT Combined with MRI for Gross Tumor Volume Delineation in Patients with Skull Base Paraganglioma. *Cancers* (2019) 11:54. doi: 10.3390/cancers11010054
56. Rieken S, Habermehl D, Giesel FL, Hoffmann C, Burger U, Rief H, et al. Analysis of FET-PET imaging for target volume definition in patients with gliomas treated with conformal radiotherapy. *Radiother Oncol* (2013) 109:487–92. doi: 10.1016/j.radonc.2013.06.043
57. Kosztyla R, Chan EK, Hsu F, Wilson D, Ma R, Cheung A, et al. High-Grade Glioma Radiation Therapy Target Volumes and Patterns of Failure Obtained From Magnetic Resonance Imaging and 18F-FDOPA Positron Emission Tomography Delineations From Multiple Observers. *Int J Radiat Oncol Biol Phys* (2013) 87:1100–6. doi: 10.1016/j.ijrobp.2013.09.008
58. Harat M, Małkowski B, Makarewicz R. Pre-irradiation tumour volumes defined by MRI and dual time-point FET-PET for the prediction of glioblastoma multiforme recurrence: A prospective study. *Radiother Oncol* (2016) 120:241–7. doi: 10.1016/j.radonc.2016.06.004
59. Hirata T, Kinoshita M, Tamari K, Seo Y, Suzuki O, Wakai N, et al. 11C-methionine-18F-FDG dual-PET-tracer–based target delineation of malignant glioma: evaluation of its geometrical and clinical features for planning radiation therapy. *J Neurosurg* (2019) 131:676–86. doi: 10.3171/2018.4.JNS1859
60. Lohmann P, Stavrinou P, Lipke K, Bauer EK, Ceccon G, Werner J-M, et al. FET PET reveals considerable spatial differences in tumour burden compared to conventional MRI in newly diagnosed glioblastoma. *Eur J Nucl Med Mol Imaging* (2019) 46:591–602. doi: 10.1007/s00259-018-4188-8
61. Hayes AR, Jayamanne D, Hsiao E, Schembri GP, Bailey DL, Roach PJ, et al. Utilizing 18F-fluoroethyltyrosine (FET) positron emission tomography (PET) to define suspected nonenhancing tumor for radiation therapy planning of glioblastoma. *Pract Radiat Oncol* (2018) 8:230–8. doi: 10.1016/j.prro.2018.01.006
62. Thorwarth D, Müller A-C, Pfannenberg C, Beyer T. "Combined PET/MR Imaging Using 68Ga-DOTATOC for Radiotherapy Treatment Planning in Meningioma Patients,". In: RP Baum, F Rösch, editors. *Theranostics, Gallium-68, and Other Radionuclides Recent Results in Cancer Research*. Berlin, Heidelberg: Springer Berlin Heidelberg (2012). p. 425–39. doi: 10.1007/978-3-642-27994-2_23
63. Graf R, Plotkin M, Steffen IG, Wurm R, Wust P, Brenner W, et al. Magnetic resonance imaging, computed tomography, and 68Ga-DOTATOC positron emission tomography for imaging skull base meningiomas with infracranial extension treated with stereotactic radiotherapy - a case series. *Head Face Med* (2012) 8:1. doi: 10.1186/1746-160X-8-1
64. Stade F, Dittmar J-O, Jäkel O, Kratochwil C, Haberkorn U, Debus J, et al. Influence of 68Ga-DOTATOC on sparing of normal tissue for radiation therapy of skull base meningioma: differential impact of photon and proton radiotherapy. *Radiat Oncol* (2018) 13:58. doi: 10.1186/s13014-018-1008-z
65. Maclean J, Fersht N, Sullivan K, Kayani I, Bomanji J, Dickson J, et al. Simultaneous 68 Ga DOTATATE Positron Emission Tomography/Magnetic Resonance Imaging in Meningioma Target Contouring: Feasibility and Impact Upon Interobserver Variability Versus Positron Emission Tomography/Computed Tomography and Computed Tomography/Magnetic Resonance Imaging. *Clin Oncol* (2017) 29:448–58. doi: 10.1016/j.clon.2017.03.004
66. Chun SG, Hu C, Choy H, Komaki RU, Timmerman RD, Schild SE, et al. Impact of Intensity-Modulated Radiation Therapy Technique for Locally Advanced Non-Small-Cell Lung Cancer: A Secondary Analysis of the NRG Oncology RTOG 0617 Randomized Clinical Trial. *J Clin Oncol* (2017) 35:56–62. doi: 10.1200/JCO.2016.69.1378
67. Giraud P, Morvan E, Claude L, Mornex F, Le Pechoux C, Bachaud J-M, et al. Respiratory gating techniques for optimization of lung cancer radiotherapy. *J Thorac Oncol* (2011) 6:2058–68. doi: 10.1097/JTO.0b013e3182307ec2
68. Sindoni A, Minutoli F, Pontoriero A, Iatì G, Baldari S, Pergolizzi S. Usefulness of four dimensional (4D) PET/CT imaging in the evaluation of thoracic lesions and in radiotherapy planning: Review of the literature. *Lung Cancer* (2016) 96:78–86. doi: 10.1016/j.lungcan.2016.03.019
69. Dyk P, Jiang N, Sun B, DeWees TA, Fowler KJ, Narra V, et al. Cervical Gross Tumor Volume Dose Predicts Local Control Using Magnetic Resonance Imaging/Diffusion-Weighted Imaging—Guided High-Dose-Rate and Positron Emission Tomography/Computed Tomography—Guided Intensity Modulated Radiation Therapy. *Int J Radiat Oncol Biol Phys* (2014) 90:794–801. doi: 10.1016/j.ijrobp.2014.07.039
70. Zhang S, Xin J, Guo Q, Ma J, Ma Q, Sun H. Comparison of Tumor Volume Between PET and MRI in Cervical Cancer With Hybrid PET/MR. *Int J Gynecol Cancer* (2014) 24:744–50. doi: 10.1097/IGC.0000000000000097
71. Han K, Croke J, Foltz W, Metser U, Xie J, Shek T, et al. A prospective study of DWI, DCE-MRI and FDG PET imaging for target delineation in brachytherapy for cervical cancer. *Radiother Oncol* (2016) 120:519–25. doi: 10.1016/j.radonc.2016.08.002
72. Lai AYT, Perucho JAU, Xu X, Hui ES, Lee EYP. Concordance of FDG PET/CT metabolic tumour volume versus DW-MRI functional tumour volume with T2-weighted anatomical tumour volume in cervical cancer. *BMC Cancer* (2017) 17:825. doi: 10.1186/s12885-017-3800-9
73. Zhang S, Xin J, Guo Q, Ma J, Ma Q, Sun H, et al. Defining PET tumor volume in cervical cancer with hybrid PET/MRI: a comparative study. *Nucl Med Commun* (2014) 35:712–9. doi: 10.1097/MNM.0000000000000113
74. Moghanaki D, Turkbey B, Vapiwala N, Ehdaie B, Frank SJ, McLaughlin PW, et al. Advances in Prostate Cancer Magnetic Resonance Imaging and Positron Emission Tomography-Computed Tomography for Staging and Radiotherapy Treatment Planning. *Semin Radiat Oncol* (2017) 27:21–33. doi: 10.1016/j.semradonc.2016.08.008
75. Grubmüller B, Baltzer P, Hartenbach S, D'Andrea D, Helbich TH, Haug AR, et al. PSMA Ligand PET/MRI for Primary Prostate Cancer: Staging Performance and Clinical Impact. *Clin Cancer Res* (2018) 24:6300–7. doi: 10.1158/1078-0432.CCR-18-0768
76. Walacides D, Meier A, Knöchelmann AC, Meinecke D, Derlin T, Bengel FM, et al. Comparison of 68 Ga-PSMA ligand PET/CT versus conventional cross-sectional imaging for target volume delineation for metastasis-directed radiotherapy for metachronous lymph node metastases from prostate cancer. *Strahlenther Onkol* (2019) 195:420–9. doi: 10.1007/s00066-018-1417-9
77. Zamboglou C, Wieser G, Hennies S, Rempel I, Kirste S, Soschynski M, et al. MRI versus 68Ga-PSMA PET/CT for gross tumour volume delineation in radiation treatment planning of primary prostate cancer. *Eur J Nucl Med Mol Imaging* (2016) 43:889–97. doi: 10.1007/s00259-015-3257-5
78. Zamboglou C, Klein CM, Thomann B, Fassbender TF, Rischke HC, Kirste S, et al. The dose distribution in dominant intraprostatic tumour lesions defined by multiparametric MRI and PSMA PET/CT correlates with the outcome in patients treated with primary radiation therapy for prostate cancer. *Radiat Oncol* (2018) 13:65. doi: 10.1186/s13014-018-1014-1
79. Pathmanathan AU, Alexander EJ, Huddart RA, Tree AC. The delineation of intraprostatic boost regions for radiotherapy using multimodality imaging. *Future Oncol* (2016) 12:2495–511. doi: 10.2217/fon-2016-0129
80. Zamboglou C, Thomann B, Koubar K, Bronsert P, Krauss T, Rischke HC, et al. Focal dose escalation for prostate cancer using 68Ga-HBED-CC PSMA PET/CT and MRI: a planning study based on histology reference. *Radiat Oncol* (2018) 13:81. doi: 10.1186/s13014-018-1036-8
81. Bettermann AS, Zamboglou C, Kiefer S, Jilg CA, Spohn S, Kranz-Rudolph J, et al. [68Ga-]PSMA-11 PET/CT and multiparametric MRI for gross tumor volume delineation in a slice by slice analysis with whole mount histopathology as a reference standard – Implications for focal radiotherapy planning in primary prostate cancer. *Radiother Oncol* (2019) 141:214–9. doi: 10.1016/j.radonc.2019.07.005
82. Bhargava P, Ravizzini G, Chapin BF, Kundra V. Imaging Biochemical Recurrence After Prostatectomy: Where Are We Headed? *Am J Roentgenol* (2020) 214:1–11. doi: 10.2214/AJR.19.21905
83. Roels S, Slagmolen P, Nuyts J, Lee JA, Loeckx D, Maes F, et al. Biological Image-Guided Radiotherapy in Rectal Cancer: Challenges and Pitfalls. *Int J Radiat Oncol Biol Phys* (2009) 75:782–90. doi: 10.1016/j.ijrobp.2008.11.031
84. Brændengen M, Hansson K, Radu C, Siegbahn A, Jacobsson H, Glimelius B. Delineation of Gross Tumor Volume (GTV) for Radiation Treatment Planning of Locally Advanced Rectal Cancer Using Information From MRI or FDG-PET/CT: A Prospective Study. *Int J Radiat Oncol Biol Phys* (2011) 81:e439–45. doi: 10.1016/j.ijrobp.2011.03.031

85. Gwynne S, Mukherjee S, Webster R, Spezi E, Staffurth J, Coles B, et al. Imaging for Target Volume Delineation in Rectal Cancer Radiotherapy — A Systematic Review. *Clin Oncol* (2012) 24:52–63. doi: 10.1016/j.clon.2011.10.001
86. Wang YY, Zhe H. Clinical application of multimodality imaging in radiotherapy treatment planning for rectal cancer. *Cancer Imaging* (2013) 13:495–501. doi: 10.1102/1470-7330.2013.0046
87. Thureau S, Briens A, Decazes P, Castelli J, Barateau A, Garcia R, et al. PET and MRI guided adaptive radiotherapy: Rational, feasibility and benefit. *Cancer Radiother* (2020) 24:635–44. doi: 10.1016/j.canrad.2020.06.017
88. van Timmeren JE, Chamberlain M, Krayenbuehl J, Wilke L, Ehrbar S, Bogowicz M, et al. Treatment plan quality during online adaptive re-planning. *Radiat Oncol* (2020) 15:203. doi: 10.1186/s13014-020-01641-0
89. Fan Q, Nanduri A, Mazin S, Zhu L. Emission guided radiation therapy for lung and prostate cancers: a feasibility study on a digital patient. *Med Phys* (2012) 39:7140–52. doi: 10.1118/1.4761951
90. Matuszak MM, Kashani R, Green M, Owen D, Jolly S, Mierzwa M. Functional Adaptation in Radiation Therapy. *Semin Radiat Oncol* (2019) 29:236–44. doi: 10.1016/j.semradonc.2019.02.006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

REVIEW
published: 09 February 2021
doi: 10.3389/fonc.2020.604051

Artificial Intelligence for Personalized Medicine in Thyroid Cancer: Current Status and Future Perspectives

Ling-Rui Li[1], *Bo Du*[2,3], *Han-Qing Liu*[1] *and Chuang Chen*[1]*

[1] Department of Breast and Thyroid Surgery, Renmin Hospital of Wuhan University, Wuhan, China, [2] School of Computer Science, Wuhan University, Wuhan, China, [3] Institute of Artificial Intelligence, Wuhan University, Wuhan, China

Thyroid cancers (TC) have increasingly been detected following advances in diagnostic methods. Risk stratification guided by refined information becomes a crucial step toward the goal of personalized medicine. The diagnosis of TC mainly relies on imaging analysis, but visual examination may not reveal much information and not enable comprehensive analysis. Artificial intelligence (AI) is a technology used to extract and quantify key image information by simulating complex human functions. This latent, precise information contributes to stratify TC on the distinct risk and drives tailored management to transit from the surface (population-based) to a point (individual-based). In this review, we started with several challenges regarding personalized care in TC, for example, inconsistent rating ability of ultrasound physicians, uncertainty in cytopathological diagnosis, difficulty in discriminating follicular neoplasms, and inaccurate prognostication. We then analyzed and summarized the advances of AI to extract and analyze morphological, textural, and molecular features to reveal the ground truth of TC. Consequently, their combination with AI technology will make individual medical strategies possible.

Keywords: artificial intelligence, thyroid cancer, biomarker, personalized medicine, histopathology, fine-needle aspiration biopsy, ultrasound

OPEN ACCESS

Edited by:
Randy Yeh,
Memorial Sloan Kettering Cancer Center, United States

Reviewed by:
Vasyl Vasko,
Uniformed Services University of the Health Sciences, United States
Barbara Maria Jarzab,
Maria Skłodowska-Curie National Research Institute of Oncology, Gliwice Branch, Poland
Alessandro Antonelli,
University of Pisa, Italy

***Correspondence:**
Chuang Chen
chenc2469@163.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *08 September 2020*
Accepted: *21 December 2020*
Published: *09 February 2021*

Citation:
Li L-R, Du B, Liu H-Q and Chen C (2021) Artificial Intelligence for Personalized Medicine in Thyroid Cancer: Current Status and Future Perspectives. Front. Oncol. 10:604051. doi: 10.3389/fonc.2020.604051

INTRODUCTION

Thyroid cancers (TC) have emerged in popularity over the past decades, with indolent TC accounting for the majority (1–3). For advanced TC (1, 2) and aggressive papillary thyroid carcinomas (PTC) (4), the incidence and mortality rates are also steadily increasing, which makes it imperative to adopt more effective strategies for managing such changes. In the era of personalized medicine, precise and efficient risk stratification is important before, during, and after treatment, to choose and adjust its type and intensity. The foremost step is to discover key information that reveals the biological behavior of TC. There are abundant anatomical structures (texture, internal architecture, and spatial distribution) and molecular components (gene variation, protein expression, etc.) within TC. So far, TC's diagnosis mainly relies on image analysis (e.g., ultrasound images, cell smears, and tissue sections), but information obtained only by our naked eyes hardly enables a comprehensive analysis of the tumors (5). Given patients and their disease features, primary human cell cultures both from surgical biopsies and from fine-needle aspiration (FNA) samples foster the targeted therapies (6). However, many tough challenges still hinder a clear

break of personalized treatment such as inconsistent rating ability of ultrasound (US) physicians (7), uncertainty in cytopathological diagnosis (8), difficulty in discriminating follicular neoplasms (9, 10), and inaccurate prognostication.

Artificial intelligence (AI) is a series of technologies combined to mimic human interaction (**Figure 1**). In some tasks, it matches or exceeds human perception (11, 12). AI deals with various sorts of omics information in parallel, easily identifying and modeling a complicated nonlinear relationship in the image (13, 14). Several studies have demonstrated that AI classifier is comparable to radiologists while qualitatively analyzing thyroid nodules (TN) (15–18). Furthermore, AI can extract and quantify key image information, whereby image diagnosis converts from a subjective qualitative task to objective quantitative analysis. This more detailed and precise information is conducive to special risk stratification and propels tailored management to transit from the surface (population-based) to a point (individual-based).

In this review, we aimed to summarize the use of AI for extracting and analyzing morphological, textural, and molecular features to reveal detailed information and personalize therapies for TC patients (**Figure 2**).

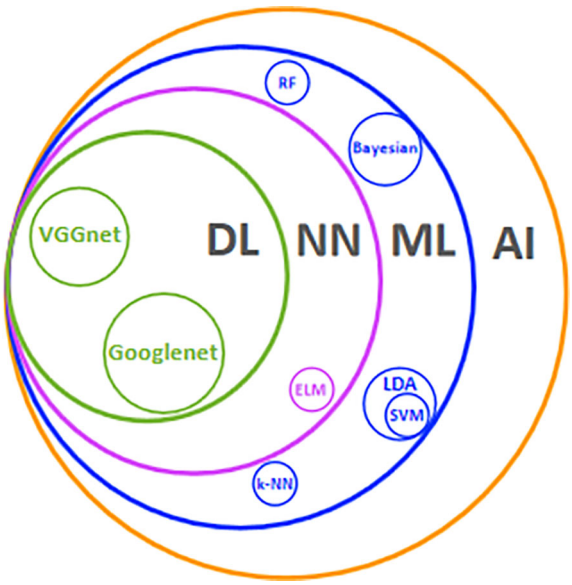


FIGURE 1 | Main AI technologies and their relationships. AI. Artificial intelligence; ML, machine learning; NN, neural network; DL, deep learning; LDA, linear discriminant analysis; ELM, extreme learning machine; RF, random forest; SVM, support vector machine; k-NN, k-nearest neighbor.

APPLICATIONS OF AI IN THE US DIAGNOSIS OF TN

TN with several typical ultrasound features implies an increased risk of malignancy, such as solid composition, hypoechogenicity, irregular margin, microcalcification, and taller-than-wide shape. However, these properties can neither confirm nor exclude the diagnosis of TC (19). The observer's agreement among multiple centers is poorly satisfactory in assessing these features (7). Thyroid Imaging Reporting and Data Systems (TI-RADS) are enormously valuable to PTC as risk stratification systems, while relatively less to FTC, MTC, and other malignancies (20). Interestingly, the AI model appears to be a promising tool to facilitate a better knowledge of TN via quantitative analysis of typical US features and introduction of texture features.

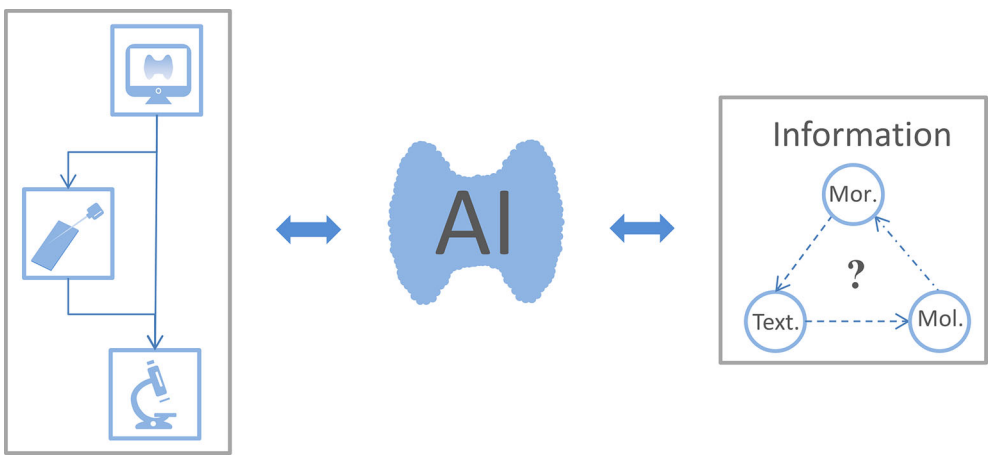


FIGURE 2 | The connection between the focus reviewed. Thyroid ultrasound is the preferred imaging examination for patients with thyroid nodules. When sonographers consider certain thyroid nodules as malignant, the patient could choose fine-needle aspiration biopsy or surgery for a confirmed diagnosis. Artificial intelligence (AI) uses ultrasound images, cell smears, and tissue slices to extract morphological, textural, and molecular features. This information is fed back into the AI classifier to improve its performance and thus optimize thyroid cancer diagnosis and treatment workflow. As expected, whether these morphological (Mor.), textural (Text.), and molecular (Mol.) features are related to each other warrants further study.

Performance of Typical US Features

Wildman-Tobriner et al. (21) developed an AI TI-RADS based on the American College of Radiology (ARC) TI-RADS. This system optimized the evaluation task through reassigned values for eight ultrasound features, highlighting the status of hypoechogenicity or marked hypoechogenicity. The novel AI TI-RADS had better accuracy than ARC TI-RADS when performed by inexperienced radiologists (55% vs. 48%) and experts (65% vs. 47%). Similar to other studies, ARC TI-RADS-based classifiers had higher sensitivity and slightly lower specificity (21–24). Wu et al. (25) evaluated quantitative echoic indexes for detecting malignant TN, which showed higher accuracy than typical ultrasound hypoechogenicity (>60% vs. 54.01%). We summarized the outcomes of the ultrasound features employed by AI for classification in **Table 1** and found the most widely used features were shape, margin, echogenicity, calcification, composition, and size. In other words, these discriminative features seem to be the focus for the AI model to learn (31, 37). Particularly, Choi et al. (30) demonstrated several new calcification features associated with TN malignancy, including shorter calcification distance ratio, smaller amounts of calcification, and dimmer calcification. Chen et al. (28) quantified TN malignant risk through the calcification index. These new features boosted diagnostic accuracy by combining qualitative and quantitative methods (30, 38). Current AI classifiers focus on benign and malignant TN dichotomy, and certain of them like the S-Detect series have already become commercially available (32, 34). Furthermore, they are expected to predict more tumor-biological behaviors such as lymph node metastasis (39, 40) and pathological subtypes (41).

Performance of Texture Features

A meta-analysis suggested that a taller-than-wide shape displays TN's variation in space and orientation growth, and it is defined as the most suggestive feature for malignancy (42). Texture features refer to the characterization of spatial distribution and surface orientation with numerical features (43). Thus, texture analysis as a powerful alternative will make it possible for radiologists to comprehend the TN in depth and gain a correct diagnosis. Raghavendra et al. (44) integrated spatial and fractal texture features and screened two features with an excellent area under the curve in diagnostic practice (94.45%). Prochazka et al. (45) used AI to extract texture features from US images independent of the direction of the US probe and achieved better accuracy (94.64%). Yu et al. (46) performed a numerical transformation of two US features, unregulated shape and long/short-axis ratio into the perimeter2/area and the angle between the long axis and the horizontal axis. These new features showed excellent sensitivity and specificity (100% and 87.88%, respectively) combined with 65 texture features. Collectively, AI mode has a role in integrating typical ultrasonic and texture features, and this fusion might sharply reduce the differences in judgments among US professionals. Despite the mounting advantages of the AI model in optimizing and even creating workflows, many remarkable factors hold its ultimate practice back in the real world. The three main factors are

TABLE 1 | Summary of key studies on the outcome of ultrasound features in artificial intelligence classifier identifying benign and malignant thyroid nodules.

Study	Patients	Features	Classifiers	Accuracy, %	Sensitivity, %	Specificity, %	AUC
Lim et al. (26)	96	Size, margin, cystic change, echogenicity, and macrocalcification	ANN	93.78	NA	NA	0.949
Savelonas et al. (27)	387	Boundary features	SVM	NA	NA	NA	0.95
Chen et al. (28)	256	Calcification index	AmCAD-UT	NA	NA	NA	0.746
Zhu et al. (29)	618	Not well-circumscribed, solid, hypoechogenicity, microcalcification, taller than wide, absent peripheral halo	ANN	83.10	83.80	81.80	0.828
Choi et al. (30)	85	Quantitative calcification	NN	82.80	83.00	82.40	0.83
Wu et al. (25)	333	Quantitative echogenetic values	AmCAD-UT	70.32	33.12	93.31	NA
Xia et al. (31)	187	Margin, shape, composition, echogenicity, and calcification	ELM	87.72	78.89	94.55	0.867
Choi et al. (32)	89	Size, margins, shape, composition, echogenicity, orientation, and spongiform	S-Detect 1 (SVM)	81.40	90.70	74.60	0.83
Ouyang et al. (33)	1036	Size, margins, shape, composition, echogenicity, calcification, aspect ratio, capsule, hypoechoic halo, vascularity, and cervical lymph node status	RF + k-SVM	NA	NA	NA	0.954
Kim et al. (34)	106	Size, margins, shape, composition, echogenicity, calcification, orientation, and spongiform	S-Detect 2 (CNN)	73.40	81.40	68.20	NA
Liu et al. (35)	4655	Shape, context, and margin	CNN	94.90	97.20	89.10	NA
Wildman-Tobrineret et al. (21)	1264	ACR TIRADS	Genetic Algorithm	65.00	93.30	64.70	0.93
Guan et al. (36)[a]	2235	Margin size	Inception-v3	90.50	93.30	87.40	0.956
Zhao et al. (22)	822	Size + ACR TIRADS	ML	82.10	90.90	78.10	0.917
Jin et al. (23)	695	ACR TIRADS	CNN	80.35	80.64	80.13	0.87
Bai et al. (24)	13984[b]	ACR TIRADS	CNN	88.00	98.10	79.10	NA

ANN, artificial neural network; SVM, support vector machine; AmCAD-UT, a software integrating AI technology and clinicians' expertise; NN, neural network; ELM, extreme learning machine; RF, random forest; k-SVM, kernel support vector machine; Inception-v3, a kind of Googlenet; CNN, convolutional neural network; ML, machine learning; NA, not available; ACR TRIADS, American College of Radiology Thyroid Imaging Reporting and Data System, the features included in this system is composition, echogenicity, shape, margin, and echogenic foci.

[a]*This study focused on the classification between papillary thyroid carcinomas and benign nodules.*

[b]*Nodules, not patients.*

as follows: (i) poor availability of large-high-quality datasets to guarantee great robustness (17); (ii) lack of explainability for conclusions from a black-box algorithm to solidify the trust between physicians and patients (47, 48); (iii) financial burden from specific equipment and research costs (48).

APPLICATIONS OF AI IN CYTOPATHOLOGICAL EVALUATION FROM FNA

FNA is a primary preoperative examination to evaluate TN. Its report system, the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC), is a state-of-the-art and category-based method for clinicians' decision-making. While TBSRTC includes six diagnostic categories on the estimated risk of malignancy (ROM) (**Table 2**), 15%–30% of TN continues to be classified as indeterminate TN (ITN), most frequently TBSRTC categories III, IV, and V (8). Recent studies showed excellent consistency between machine learning (ML) models and cytologists in malignancy prediction (49–51), in which the ROM of TBSRTC III determined by the ML model was considerably lower than by manual classification (4.2% vs. 18.8%) (51). It's worth noting that morphological and genetic classifications assisted by the AI model are fairly accurate at distinguishing malignancy from benign TN (52–54) (**Table 3**).

TABLE 2 | The 2017 TBSRTC categories and their own risk of malignancy.

Diagnostic category	Risk of malignancy if NIFTP≠ CA (%)
I. nondiagnostic or unsatisfactory	5–10
Cyst fluid only	
Virtually acellular specimen	
Other (obscuring blood, clotting artifact, etc.)	
II. Benign	0–3
Consistent with a benign follicular nodule (includes adenomatoid nodule, colloid nodule, etc.)	
Consistent with lymphocytic (Hashimoto) thyroiditis in the proper clinical context	
Consistent with granulomatous (subacute) thyroiditis	
Other	
III. atypia of undetermined significance or follicular lesion of undetermined significance	6–18
III. follicular neoplasm or suspicious for a follicular neoplasm	10–40
Specify if Hürthle cell (oncocytic) type	
IV. suspicious for malignancy	45–60
Suspicious for papillary carcinoma	
Suspicious for medullary carcinoma	
Suspicious for metastatic carcinoma	
Suspicious for lymphoma Other	
V. Malignant	94–96
Papillary thyroid carcinoma	
Poorly differentiated carcinoma	
Medullary thyroid carcinoma	
Undifferentiated (anaplastic) carcinoma	
Squamous-cell carcinoma	
Carcinoma with mixed features (specify)	
Metastatic carcinoma Non-Hodgkin lymphoma	
Other	

This is an integrated table from reference (8).

Performance of Morphological Features

PTC, the most common TC (>80%), arises from abnormal growth of thyroid epithelial cells (28, 38). In recent years, AI models with quantitative morphological features have tried to improve follicular lesions' recognition capacity (55–57). Sanyal et al. (55) obtained the nuclear morphology and papillary structure of PTC under two magnifications (×10 and ×40). CNN model selected PTC from colloid goiter, follicular neoplasms, and lymphocytic thyroiditis by right of these features. Guan et al. (56) developed a new AI cytological classification based on nuclear size and staining information (the contours, perimeter, area, and means of pixel intensity), whose results showed high accuracy (97.66%) to differentiate PTC from benign nodules. Another research group also confirmed this performance (57). They first derived nuclear pleomorphism and area information and then reported the weight of 17 cytological and morphological features. Finally, their model successfully discriminated follicular carcinoma (FC) from follicular adenoma (FA) (57) (**Table 3**).

The major difference between FC and FA is the occurrence of capsular or vascular invasion (67). Preoperative examinations of both US and FNA have difficulty in making a reliable diagnosis. A highly vascularized tumor protrusion on the US strongly indicates FC, which is rather rare yet (68). Seo et al. (69) took full advantage of this difference by collecting information about the tumor edge in the US images. The overall accuracy was 89.51% for distinguishing FC and FA. Yang et al. (70) segmented the whole lesions of follicular neoplasms; as a result, the classification accuracy was significantly improved to 96%. This clarified the importance of internal information and affirmed the study's reliability by Savala et al. (57). Similarly, the diagnosis of MTC and ATC is histology dependent (71, 72), yet now no studies to our knowledge have answered the hope of AI in their ultrasound and cytopathological diagnosis.

Performance of Biomolecules

For patients with ITN, repeat FNA or lobectomy might be performed because management guidelines are more flexible (8, 73). Fortunately, molecular tests provide a noninvasive and accurate option to reduce clinical and healthy uncertainty (8, 67). Each genome contains as much information as 100,000 photographs (74). Next-generation sequencing (NGS) can perform high-speed analysis of multiple genes parallelly in a single operation, producing billions of molecular fragments (74, 75). It has always been a crucial component of big data due to its large volume of data, the astonishing velocity of the sequencing methods, and the result output's veracity. Traditional information systems are less competent to analyze large and complex datasets (76, 77). AI as a big data algorithm can integrate multi-omic data in a different learning task, and automatically realize high-level features' detection or classification (77). Some genetic classifiers have played their strengths in TN such as the Afirma gene expression classifier (GEC) (58), gene sequence classifier (GSC) (59), gene mutation-based classifier (ThyroSeq) (60, 78), and microRNA-based classifier (RosettaGX Reveal) (61, 79). The GEC involved 167 genes that displayed high sensitivity (92%)

TABLE 3 | The main performance of artificial intelligence using pathological information in different task.

Study	Subject	Test	Feature	Task	Classifier	Accuracy, %
Cochand-Priollet et al. (54)	157	FNA	Nuclear size, shape, and texture	Classification of malignant and benign TN	FNN	89.00
Daskalakis et al. (53)	115	FNA	Nuclear morphology and texture	Classification of malignant and benign TN	k-NN + PNN + Bayesian	95.70
Tomei et al. (52)	93	FNA	mRNA expression	Classification of malignant and benign TN	BNN	88.80
Sanyal et al. (55)	544	FNA	Nuclear morphology and papillary structure	Classification of PTC and non-PTC	ANN	85.06
Guan et al. (56)	279	FNA	Nuclear contour	Classification of PTC and benign TN	VGG-16	97.66
Savala et al. (57)	57	FNA	Cellular and nuclear morphology	Classification of FC and FA	ANN	100.00
Alexander et al. (58)	249[a]	FNA	RNA expression	Classification of malignant and benign ITN	SVM	65.00
Patel et al. (59)	183[a]	FNA	RNA sequencing	Classification of malignant and benign ITN	SVM	74.00
Nikiforova et al. (60)	175[ab]	FNA	Genetic alterations	Classification of malignant and benign ITN	Torrent Suite software	90.90
Lithwick-Yanai et al. (61)	150[ab]	FNA	MicroRNA expression	Classification of malignant and benign ITN	LDA + k-NN	83.65
Sun et al. (62)	64[a]	FNA	Protein	Classification of malignant and benign TN	ANN	87.53
Wang et al. (63)	10	Histo.	Nuclear size and chromatin concentration	Classification of FC, FA, and normal thyroid	SVM	100.00
Ozolek et al. (64)	94	Histo.	Nuclear morphology	Classification of five follicular lesions	LDA +k-NN	100.00[c]
Zhao et al. (65)	800	Histo.	Gene variant pathways	TC risk stratification	ANN	77.50/ 86.00[d]
Ruiz et al. (66)	495	Histo.	Gene signature	Prediction of lymph-node metastasis and disease-free survival	LDA	82.63

FNA, fine-needle aspiration biopsy; Histo., Histopathology; TN, thyroid nodules; FNN, feedforward neural network; PNN, probabilistic neural network; BNN, Bayesian neural network; PTC, papillary thyroid carcinoma; FC, follicular carcinoma; FA, follicular adenoma; ANN, artificial neural network; SVM, support vector machine; LDA, linear discriminant analysis; k-NN, k-nearest neighbor; NA, not available.

[a]*Only the validation cohort is included, which in the study by Lithwick-Yanai et al. was specifically the set agreed upon by the three pathologists.*

[b]*FNA smears, not patients.*

[c]*The accuracy in the group of FA vs. FC, FA vs. NG, FC vs. NG, FA vs. FV-PTC, and FC vs. WIFC.*

[d]*The accuracy of recognizing the different- risk cases was 77.50% (low-risk) and 86.00% (high-risk) respectively.*

and positive predictive value (PPV: 93%) but limited by its relatively low specificity (52%) and negative predictive value (NPV: 47%) (58). GSC expanded the gene spectrum to 10,196 genes by RNA-enhanced NGS. Compared to GEC in the same samples, it made progress in screening for benign nodules (sensitivity: 91.1% and specificity: 68.3%) (59) (**Table 3**). These two classifiers are the most broadly accepted methods to rule out malignant nodules. In general, ThyroSeq and RosettaGX Reveal are more like rule-in entities. Nikiforova et al. (60) achieved a robust sensitivity of 98% and a hopeful specificity of 81.8% by employing the latest version of ThyroSeq (ThyroSeq V3) to recognize a few cancers from most benign tumors. Steward et al. (78) drew similar conclusions in a prospective blinded multicenter study and reported 94% sensitivity and 82% specificity. RosettaGX Reveal showed 98% sensitivity and 78% specificity when validated in independent cases with all three pathologists' agreement on the histopathological diagnosis (61) (**Table 3**). However, whether the mentioned classifiers could consolidate and complement each other remains so ambiguous that we need to further investigate the precise application strategy.

The multi-gene analysis is able to enhance diagnostic performance, but it may be limited due to key genes' deletion or their reduced expression. Of note, the number of thyroglobulins has been considered as a predictor of postoperative disease progression (67). Therefore, the key proteins might provide some added information for personalized therapy. Recent research has confirmed that proteins are more stable than RNA in clinical tissues (80). Sun et al. (62) completed a 14 protein-based ANN classifier for TN classification. This model realized the accuracy of 90.62% and 87.53% in multicenter retrospective and prospective samples respectively (**Table 3**). Some molecular alterations such as *BRAF* mutations (81) are diagnostic of cancer, but most of the other alterations (82, 83) show overlap in both benign and malignant lesions. Therefore, assessing the risk of malignancy by molecular testing should depend on knowledge of the prior cytological appearance.

APPLICATIONS OF AI IN HISTOPATHOLOGICAL ANALYSIS

Upon reliable evidence obtained by the US and FNA examination, tumor information from the resected specimens is significant for pathologists to diagnosis TC such as tumor size, pathologic types, and degree of malignancy. Molecular patterns in the tumor microenvironment like cytokines, chemokines, and adipocytokines interconnect the units of immune-inflammatory responses (e.g., macrophages, neutrophils, lymphocytes) and tumor nest (e.g., epithelial cancer cells, fibroblasts, endothelial cells) (84). The more detailed information the pathologists provide, the more precise the treatment strategies physicians take. The combination of AI, morphology, and molecular markers is expected to provide more information for TC management at a patient's level.

Performance of Morphological Features

The morphological feature is the final station of biological behavior and genetic variation of TN. The morphological performance supported by AI might be beneficial for the accurate diagnosis of TN. Wang et al. (63) successfully classified FA, FC, and normal tissues according to nuclear size and chromatin concentration. Ozolek et al. (64) achieved nearly perfect accuracy based on nine nuclear morphological features for discriminating five thyroid follicular lesions: FA, FC, follicular variant of PTC, nodular goiter, and the widely invasive FC (**Table 3**). However, further validation of these models is required due to tumor complicated heterogeneity, which was also turned out in a recent study for classifying TC, normal tissues, nodular goiter, and adenomas using a deep learning model (85).

Morphologically, FV-PTC is a mixed entity for typical PTC nuclear features and entirely or almost entirely follicular growth patterns. FV-PTC includes two major subtypes: encapsulated (EFV-PTC) and non-encapsulated or infiltrative variants (IFV-PTC) (86). The former generally have RAS mutations like follicular tumors, the latter often presents extrathyroidal extension (ETE), lymphatic metastasis, and BRAF mutations like classical PTC (cPTC) (87). Likewise, EFV-PTC usually appears invasive or non-invasive, and the noninvasive encapsulate tumor was redefined from carcinoma to borderline tumor, noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) (86). Up to a point, the invasive EFV-PTC behaves more aggressively like FC, whereas NIFTP is with indolent clinical behaviors like FA (87). It is believed that invasive EFV-PTC might develop from NIFTP (88). Borrelli et al. (89) revealed a significant difference in miRNA expression of FA, NIFTP, and IFV-PTC. In particular, just two miRNA (miR-10a-5p and miR-320e) enable us to differentiate NIFTP from IFV-PTC. In another study by Selvaggi (90), none of the multinucleated giant cells (MGCs) were observed in 20 NIFTP cases, while the amount of MGCs varied from 1 to 4 in 88% of the FVPTC cases (both IFV-PTC and invasive EFV-PTC). When utilizing computer quantitative analysis to classify FV-PTC, Chain et al. (91) demonstrated the NIFTP nuclear area (mean, 54.8 μm2) and elongation was smaller than PTC (mean, 77.2 μm2); Hsieh et al. (92) addressed PD-L1 expression in NIFTP was lower than in invasive EFV-PTC. These quantitative morphological characteristics and definite molecular alterations contribute to FV-PTC classification.

As FV-PTC's definition stated, the coexistence of papillary and variable follicular structures is so common in cancer nests that we hold a positive view about more transitional or intermediate categories between the cPTC and FV-PTC. Undoubtedly, the clearer the learning exemplars, the easier it is to learn for the AI model because it receives fewer error messages (13). For greater efficiency, it's essential to accurately classify the training set and refine the output target.

Performance of Genetic Parameters

The American Thyroid Association risk stratification system and the American Joint Committee on Cancer TNM staging system are used to guide postoperative treatment and predict post-treatment outcomes, which incorporate several parameters including age, ETE, anatomic location, number, and size of metastatic lymph nodes, aggressive variants, vascular invasion, and distant metastasis. Nonetheless, these systems fail to routinely recommend a genetic determination to guide individual management (67, 93). Zhao et al. (65) selected 10 gene variant pathways that involved inflammatory and immune responses to determine the TC patients' risk level. Based on these pathways, the patients were divided into the high-risk and low-risk groups whose survival time was significantly better than the former. Ruiz et al. (66) demonstrated a 25-gene panel related to molecular pathways, cell structure, and function was an independent prognostic factor for lymphatic metastasis and disease-free survival (**Table 3**). Further evidence is still warranted to address the value of this genetic information to TN's triage and biological behaviors. As AI and gene testing technology upgrade, the cooperation of traditional clinic-pathological parameters and gene molecules might yield more precise therapeutic implications.

CONCLUSION

The future development of personalized medicine in TC still faces several challenges like inconsistent rating ability of US physicians, uncertainty in cytopathological diagnosis, difficulty in discriminating follicular lesions, and inaccurate prognostication. AI's application has improved the efficiency and accuracy of diagnosis and treatment in other tumors (94–96). A growing amount of medical information can be extracted and analyzed through AI technology. This review has innovatively offered ideas for the ultrasonic and pathological testing out of these dilemmas in terms of morphological, textural, and molecular features. As more key parameters are explored from the tumor and its microenvironment, the AI-aided combination of morphological and molecular features will pave the way for TC's protocol at the individual level.

AUTHOR CONTRIBUTIONS

L-RL: study design, literature review, article writing, and revision. BD: literature review. H-QL: literature review. CC: study design, article revision. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the grants from the Fundamental Research Funds for the Central Universities (2042019kf0229) and the Science and Technology Major Project of Hubei Province (Next-Generation AI Technologies) (2019AEA170).

ACKNOWLEDGMENTS

Sincere thanks to the teachers, classmates, and editors who worked together on this research.

REFERENCES

1. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. *JAMA* (2017) 317 (13):1338–48. doi: 10.1001/jama.2017.2719
2. Jegerlehner S, Bulliard JL, Aujesky D, Rodondi N, Germann S, Konzelmann I, et al. Overdiagnosis and overtreatment of thyroid cancer: A population-based temporal trend study. *PloS One* (2017) 12(6):e0179387. doi: 10.1371/journal.pone.0179387
3. La Vecchia C, Malvezzi M, Bosetti C, Garavello W, Bertuccio P, Levi F, et al. Thyroid cancer mortality and incidence: a global overview. *Int J Cancer* (2015) 136(9):2187–95. doi: 10.1002/ijc.29251
4. Ho AS, Luu M, Barrios L, Chen I, Melany M, Ali N, et al. Incidence and Mortality risk spectrum across aggressive variants of papillary thyroid Carcinoma. *JAMA Oncol* (2020) 6(5):706–13. doi: 10.1001/jamaoncol.2019.6851
5. Koelzer VH, Sirinukunwattana K, Rittscher J, Mertz KD. Precision immunoprofiling by image analysis and artificial intelligence. *Virchows Archiv* (2019) 474(4):511–22. doi: 10.1007/s00428-018-2485-z
6. Fallahi P, Ferrari SM, Elia G, Ragusa F, Patrizio A, Paparo SR, et al. Primary cell cultures for the personalized therapy in aggressive thyroid cancer of follicular origin. *Semin Cancer Biol* (2020). doi: 10.1016/j.semcancer.2020.06.013
7. Persichetti A, Di Stasio E, Coccaro C, Graziano F, Bianchini A, Di Donna V, et al. Inter- and intraobserver agreement in the assessment of thyroid nodule ultrasound features and classification systems: a blinded multicenter study. *Thyroid* (2020) 30(2):237–42. doi: 10.1089/thy.2019.0360
8. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid* (2017) 27(11):1341–6. doi: 10.1089/thy.2017.0500
9. Baloch ZW, Livolsi VA. Follicular-patterned lesions of the thyroid: the bane of the pathologist. *Am J Clin Pathol* (2002) 117(1):143–50. doi: 10.1309/8vl9-ecxy-nvmx-2rqf
10. LiVolsi VA, Baloch ZW. Follicular neoplasms of the thyroid: view, biases, and experiences. *Adv Anat Pathol* (2004) 11(6):279–87. doi: 10.1097/01.pap.0000138143.34505.02
11. Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. *J Intern Med* (2020) 288(1):62–81. doi: 10.1111/joim.13030
12. Cao C, Liu F, Tan H, Song D, Shu W, Li W, et al. Deep learning and its applications in biomedicine. *Genomics Proteomics Bioinf* (2018) 16(1):17–32. doi: 10.1016/j.gpb.2017.07.003
13. Manning T, Sleator RD, Walsh P. Biologically inspired intelligent decision making: a commentary on the use of artificial neural networks in bioinformatics. *Bioengineered* (2014) 5(2):80–95. doi: 10.4161/bioe.26997
14. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol* (1996) 49(11):1225–31. doi: 10.1016/s0895-4356(96)00002-9
15. Sun C, Zhan Y, Chang Q, Liu T, Zhang S, Wang X, et al. Evaluation of a deep learning-based computer-aided diagnosis system for distinguishing benign from malignant thyroid nodules in ultrasound images. *Med Phys* (2020) 47 (9):3952–60. doi: 10.1002/mp.14301
16. Wang L, Yang S, Yang S, Zhao C, Tian G, Gao Y, et al. Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network. *World J Surg Oncol* (2019) 17(1):12. doi: 10.1186/s12957-019-1558-z
17. Li X, Zhang S, Zhang Q, Wei X, Pan Y, Zhao J, et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study. *Lancet Oncol* (2019) 20(2):193–201. doi: 10.1016/s1470-2045(18)30762-9
18. Buda M, Wildman-Tobriner B, Hoang JK, Thayer D, Tessler FN, Middleton WD, et al. Management of Thyroid Nodules Seen on US Images: Deep Learning May Match Performance of Radiologists. *Radiology* (2019) 292 (3):695–701. doi: 10.1148/radiol.2019181343
19. Remonti LR, Kramer CK, Leitao CB, Pinto LC, Gross JL. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. *Thyroid* (2015) 25(5):538–50. doi: 10.1089/thy.2014.0353
20. Trimboli P, Castellana M, Piccardo A, Romanelli F, Grani G, Giovanella L, et al. The ultrasound risk stratification systems for thyroid nodule have been evaluated against papillary carcinoma. A meta-analysis. *Rev Endocr Metab Disord* (2020). doi: 10.1007/s11154-020-09592-3
21. Wildman-Tobriner B, Buda M, Hoang JK, Middleton WD, Thayer D, Short RG, et al. Using artificial intelligence to revise ACR TI-RADS risk stratification of thyroid nodules: diagnostic accuracy and utility. *Radiology* (2019) 292(1):112–9. doi: 10.1148/radiol.2019182128
22. Zhao CK, Ren TT, Yin YF, Shi H, Wang HX, Zhou BY, et al. A comparative analysis of two machine learning-based diagnostic patterns with thyroid imaging reporting and data system for thyroid nodules: diagnostic performance and unnecessary biopsy rate. *Thyroid* (2020). doi: 10.1089/thy.2020.0305
23. Jin Z, Zhu Y, Zhang S, Xie F, Zhang M, Zhang Y, et al. Ultrasound Computer-aided diagnosis (CAD) based on the thyroid imaging reporting and data system (TI-RADS) to distinguish benign from malignant thyroid nodules and the diagnostic performance of radiologists with different diagnostic experience. *Med Sci Monit* (2020) 26:e918452. doi: 10.12659/msm.918452
24. Bai Z, Chang L, Yu R, Li X, Wei X, Yu M, et al. Thyroid nodules risk stratification through deep learning based on ultrasound images. *Med Phys* (2020) 47(12):6355–65. doi: 10.1002/mp.14543
25. Wu MH, Chen CN, Chen KY, Ho MC, Tai HC, Wang YH, et al. Quantitative analysis of echogenicity for patients with thyroid nodules. *Sci Rep* (2016) 6:35632. doi: 10.1038/srep35632
26. Lim KJ, Choi CS, Yoon DY, Chang SK, Kim KK, Han H, et al. Computer-aided diagnosis for the differentiation of malignant from benign thyroid nodules on ultrasonography. *Acad Radiol* (2008) 15(7):853–8. doi: 10.1016/j.acra.2007.12.022
27. Savelonas M, Maroulis D. Sangriotis M. A computer-aided system for malignancy risk assessment of nodules in thyroid US images based on boundary features. *Comput Methods Prog BioMed* (2009) 96(1):25–32. doi: 10.1016/j.cmpb.2009.04.001
28. Chen KY, Chen CN, Wu MH, Ho MC, Tai HC, Huang WC, et al. Computerized detection and quantification of microcalcifications in thyroid nodules. *Ultrasound Med Biol* (2011) 37(6):870–8. doi: 10.1016/j.ultrasmedbio.2011.03.002
29. Zhu LC, Ye YL, Luo WH, Su M, Wei HP, Zhang XB, et al. A model to discriminate malignant from benign thyroid nodules using artificial neural network. *PloS One* (2013) 8(12):e82211. doi: 10.1371/journal.pone.0082211
30. Choi WJ, Park JS, Kim KG, Kim SY, Koo HR, Lee YJ. Computerized analysis of calcification of thyroid nodules as visualized by ultrasonography. *Eur J Radiol* (2015) 84(10):1949–53. doi: 10.1016/j.ejrad.2015.06.021
31. Xia J, Chen H, Li Q, Zhou M, Chen L, Cai Z, et al. Ultrasound-based differentiation of malignant and benign thyroid Nodules: An extreme learning machine approach. *Comput Methods Prog BioMed* (2017) 147:37–49. doi: 10.1016/j.cmpb.2017.06.005
32. Choi YJ, Baek JH, Park HS, Shim WH, Kim TY, Shong YK, et al. A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: initial clinical assessment. *Thyroid* (2017) 27(4):546–52. doi: 10.1089/thy.2016.0372
33. Ouyang FS, Guo BL, Ouyang LZ, Liu ZW, Lin SJ, Meng W, et al. Comparison between linear and nonlinear machine-learning algorithms for the classification of thyroid nodules. *Eur J Radiol* (2019) 113:251–7. doi: 10.1016/j.ejrad.2019.02.029
34. Kim HL, Ha EJ, Han M. Real-world performance of computer-aided diagnosis system for thyroid nodules using ultrasonography. *Ultrasound Med Biol* (2019) 45(10):2672–8. doi: 10.1016/j.ultrasmedbio.2019.05.032
35. Liu T, Guo Q, Lian C, Ren X, Liang S, Yu J, et al. Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. *Med Image Anal* (2019) 58:101555. doi: 10.1016/j.media.2019.101555
36. Guan Q, Wang Y, Du J, Qin Y, Lu H, Xiang J, et al. Deep learning based classification of ultrasound images for thyroid nodules: a large scale of pilot study. *Ann Transl Med* (2019) 7(7):137. doi: 10.21037/atm.2019.04.34
37. Wei X, Zhu J, Zhang H, Gao H, Yu R, Liu Z, et al. Visual interpretability in computer-assisted diagnosis of thyroid nodules using ultrasound images. *Med Sci Monit* (2020) 26:e927007. doi: 10.12659/msm.927007

38. Chambara N, Ying M. The Diagnostic Efficiency of ultrasound computer-aided diagnosis in differentiating thyroid nodules: a systematic review and narrative synthesis. *Cancers (Basel)* (2019) 11(11):1759. doi: 10.3390/cancers11111759
39. Yu J, Deng Y, Liu T, Zhou J, Jia X, Xiao T, et al. Lymph node metastasis prediction of papillary thyroid carcinoma based on transfer learning radiomics. *Nat Commun* (2020) 11(1):4807. doi: 10.1038/s41467-020-18497-3
40. Lee JH, Baek JH, Kim JH, Shim WH, Chung SR, Choi YJ, et al. Deep Learning-Based Computer-Aided Diagnosis System for Localization and Diagnosis of Metastatic Lymph Nodes on Ultrasound: A Pilot Study. *Thyroid* (2018) 28 (10):1332–8. doi: 10.1089/thy.2018.0082
41. Wu T, Sultan LR, Tian J, Cary TW, Sehgal CM. Machine learning for diagnostic ultrasound of triple-negative breast cancer. *Breast Cancer Res Treat* (2019) 173(2):365–73. doi: 10.1007/s10549-018-4984-7
42. Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, et al. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *J Clin Endocrinol Metab* (2014) 99 (4):1253–63. doi: 10.1210/jc.2013-2928
43. Sollini M, Cozzi L, Chiti A, Kirienko M. Texture analysis and machine learning to characterize suspected thyroid nodules and differentiated thyroid cancer: Where do we stand? *Eur J Radiol* (2018) 99:1–8. doi: 10.1016/j.ejrad.2017.12.004
44. Raghavendra U, Rajendra Acharya U, Gudigar A, Hong Tan J, Fujita H, Hagiwara Y, et al. Fusion of spatial gray level dependency and fractal texture features for the characterization of thyroid lesions. *Ultrasonics* (2017) 77:110–20. doi: 10.1016/j.ultras.2017.02.003
45. Prochazka A, Gulati S, Holinka S, Smutek D. Classification of thyroid nodules in ultrasound images using direction-independent features extracted by two-threshold binary decomposition. *Technol Cancer Res Treat* (2019) 18:1533033819830748. doi: 10.1177/1533033819830748
46. Yu Q, Jiang T, Zhou A, Zhang L, Zhang C, Xu P. Computer-aided diagnosis of malignant or benign thyroid nodes based on ultrasound images. *Eur Arch Otorhinolaryngol* (2017) 274(7):2891–7. doi: 10.1007/s00405-017-4562-3
47. Thomas J, Haertling T. AIBx, artificial intelligence model to risk stratify thyroid nodules. *Thyroid* (2020) 30(6):878–84. doi: 10.1089/thy.2019.0752
48. Martín Noguerol T, Paulano-Godino F, Martín-Valdivia MT, Menias CO, Luna A. Strengths, Weaknesses, Opportunities, and Threats Analysis of Artificial Intelligence and Machine Learning Applications in Radiology. *J Am Coll Radiol* (2019) 16(9 Pt B):1239–47. doi: 10.1016/j.jacr.2019.05.047
49. Zoulias EA, Asvestas PA, Matsopoulos GK, Tseleni-Balafouta S. A decision support system for assisting fine needle aspiration diagnosis of thyroid malignancy. *Anal Quant Cytol Histol* (2011) 33(4):215–22.
50. Margari N, Mastorakis E, Pouliakis A, Gouloumi AR, Asimis E, Konstantoudakis S, et al. Classification and regression trees for the evaluation of thyroid cytomorphological characteristics: A study based on liquid based cytology specimens from thyroid fine needle aspirations. *Diagn Cytopathol* (2018) 46(8):670–81. doi: 10.1002/dc.23977
51. Elliott Range DD, Dov D, Kovalsky SZ, Henao R, Carin L, Cohen J. Application of a machine learning algorithm to predict malignancy in thyroid cytopathology. *Cancer Cytopathol* (2020) 128(4):287–95. doi: 10.1002/cncy.22238
52. Tomei S, Marchetti I, Zavaglia K, Lessi F, Apollo A, Aretini P, et al. A molecular computational model improves the preoperative diagnosis of thyroid nodules. *BMC Cancer* (2012) 12:396. doi: 10.1186/1471-2407-12-396
53. Daskalakis A, Kostopoulos S, Spyridonos P, Glotsos D, Ravazoula P, Kardari M, et al. Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological images. *Comput Biol Med* (2008) 38(2):196–203. doi: 10.1016/j.compbiomed.2007.09.005
54. Cochand-Priollet B, Koutroumbas K, Megalopoulou TM, Pouliakis A, Sivolapenko G, Karakitsos P. Discriminating benign from malignant thyroid lesions using artificial intelligence and statistical selection of morphometric features. *Oncol Rep* (2006) 15 Spec no:1023–6. doi: 10.3892/or.15.4.1023
55. Sanyal P, Mukherjee T, Barui S, Das A, Gangopadhyay P. Artificial intelligence in cytopathology: A neural network to identify papillary carcinoma on thyroid fine-needle aspiration cytology smears. *J Pathol Inform* (2018) 9:43. doi: 10.4103/jpi.jpi_43_18
56. Guan Q, Wang Y, Ping B, Li D, Du J, Qin Y, et al. Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. *J Cancer* (2019) 10 (20):4876–82. doi: 10.7150/jca.28769
57. Savala R, Dey P, Gupta N. Artificial neural network model to distinguish follicular adenoma from follicular carcinoma on fine needle aspiration of thyroid. *Diagn Cytopathol* (2018) 46(3):244–9. doi: 10.1002/dc.23880
58. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *New Engl J Med* (2012) 367(8):705–15. doi: 10.1056/NEJMoa1203208
59. Patel KN, Angell TE, Babiarz J, Barth NM, Blevins T, Duh QY, et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg* (2018) 153(9):817–24. doi: 10.1001/jamasurg.2018.1153
60. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer* (2018) 124(8):1682–90. doi: 10.1002/cncr.31245
61. Lithwick-Yanai G, Dromi N, Shtabsky A, Morgenstern S, Strenov Y, Feinmesser M, et al. Multicentre validation of a microRNA-based assay for diagnosing indeterminate thyroid nodules utilising fine needle aspirate smears. *J Clin Pathol* (2017) 70(6):500–7. doi: 10.1136/jclinpath-2016-204089
62. Sun Y, Selvarajan S, Zang Z, Liu W, Zhu YJ, Zhang H, et al. Protein classifier for thyroid nodules learned from rapidly acquired proteotypes. *medRxiv* (2020). doi: 10.1101/2020.04.09.20059741
63. Wang W, Ozolek JA, Rohde GK. Detection and classification of thyroid follicular lesions based on nuclear structure from histopathology images. *Cytometry A* (2010) 77(5):485–94. doi: 10.1002/cyto.a.20853
64. Ozolek JA, Tosun AB, Wang W, Chen C, Kolouri S, Basu S, et al. Accurate diagnosis of thyroid follicular lesions from nuclear morphology using supervised learning. *Med Image Anal* (2014) 18(5):772–80. doi: 10.1016/j.media.2014.04.004
65. Zhao Y, Zhao L, Mao T, Zhong L. Assessment of risk based on variant pathways and establishment of an artificial neural network model of thyroid cancer. *BMC Med Genet* (2019) 20(1):92. doi: 10.1186/s12881-019-0829-4
66. Ruiz EML, Niu T, Zerfaoui M, Kunnimalaiyaan M, Friedlander PL, Abdel-Mageed AB, et al. A novel gene panel for prediction of lymph-node metastasis and recurrence in patients with thyroid cancer. *Surgery* (2020) 167(1):73–9. doi: 10.1016/j.surg.2019.06.058
67. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The american thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* (2016) 26(1):1–133. doi: 10.1089/thy.2015.0020
68. Kobayashi K, Hirokawa M, Yabuta T, Masuoka H, Fukushima M, Kihara M, et al. Tumor protrusion with intensive blood signals on ultrasonography is a strongly suggestive finding of follicular thyroid carcinoma. *Med Ultrason* (2016) 18(1):25–9. doi: 10.11152/mu.2013.2066.181.kok
69. Seo JK, Kim YJ, Kim KG, Shin I, Shin JH, Kwak JY. Differentiation of the Follicular Neoplasm on the Gray-Scale US by Image Selection Subsampling along with the Marginal Outline Using Convolutional Neural Network. *BioMed Res Int* (2017) 2017:3098293. doi: 10.1155/2017/3098293
70. Yang B, Yan M, Yan Z, Zhu C, Xu D, Dong F. Segmentation and classification of thyroid follicular neoplasm using cascaded convolutional neural network. *Phys Med Biol* (2020). doi: 10.1088/1361-6560/abc6f2
71. Smallridge RC, Ain KB, Asa SL, Bible KC, Brierley JD, Burman KD, et al. American Thyroid Association guidelines for management of patients with anaplastic thyroid cancer. *Thyroid* (2012) 22(11):1104–39. doi: 10.1089/thy.2012.0302
72. Wells SAJr., Asa SL, Dralle H, Elisei R, Evans DB, Gagel RF, et al. Revised American Thyroid Association guidelines for the management of medullary thyroid carcinoma. *Thyroid* (2015) 25(6):567–610. doi: 10.1089/thy.2014.0335
73. Misiakos EP, Margari N, Meristoudis C, Machairas N, Schizas D, Petropoulos K, et al. Cytopathologic diagnosis of fine needle aspiration biopsies of thyroid nodules. *World J Clin Cases* (2016) 4(2):38–48. doi: 10.12998/wjcc.v4.i2.38

74. Phillips KA, Trosman JR, Kelley RK, Pletcher MJ, Douglas MP, Weldon CB. Genomic sequencing: assessing the health care system, policy, and big-data implications. *Health Aff* (2014) 33(7):1246–53. doi: 10.1377/hlthaff.2014.0020
75. Schmidt B, Hildebrandt A. Next-generation sequencing: big data meets high performance computing. *Drug Discovery Today* (2017) 22(4):712–7. doi: 10.1016/j.drudis.2017.01.014
76. Mayer-Schönberger V, Ingelsson E. Big Data and medicine: a big deal? *J Intern Med* (2018) 283(5):418–29. doi: 10.1111/joim.12721
77. Hamamoto R, Komatsu M, Takasawa K, Asada K, Kaneko S. Epigenetics analysis and integrated analysis of multiomics data, including epigenetic data, using artificial intelligence in the era of precision medicine. *Biomolecules* (2019) 10(1):62. doi: 10.3390/biom10010062
78. Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, Sipos JA, et al. Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. *JAMA Oncol* (2019) 5(2):204–12. doi: 10.1001/jamaoncol.2018.4616
79. Benjamin H, Schnitzer-Perlman T, Shtabsky A, VandenBussche CJ, Ali SZ, Kolar Z, et al. Analytical validity of a microRNA-based assay for diagnosing indeterminate thyroid FNA smears from routinely prepared cytology slides. *Cancer Cytopathol* (2016) 124(10):711–21. doi: 10.1002/cncy.21731
80. Shao W, Guo T, Toussaint NC, Xue P, Wagner U, Li L, et al. Comparative analysis of mRNA and protein degradation in prostate tissues indicates high stability of proteins. *Nat Commun* (2019) 10(1):2524. doi: 10.1038/s41467-019-10513-5
81. Xu X, Quiros RM, Gattuso P, Ain KB, Prinz RA. High prevalence of BRAF gene mutation in papillary thyroid carcinomas and thyroid tumor cell lines. *Cancer Res* (2003) 63(15):4561–7.
82. Zou M, Shi Y, Farid NR, al-Sedairy ST, Paterson MC. FHIT gene abnormalities in both benign and malignant thyroid tumours. *Eur J Cancer* (1999) 35(3):467–72. doi: 10.1016/s0959-8049(98)00370-0
83. Angell TE. RAS-positive thyroid nodules. *Curr Opin Endocrinol Diabetes Obes* (2017) 24(5):372–6. doi: 10.1097/med.0000000000000354
84. Ferrari SM, Fallahi P, Galdiero MR, Ruffilli I, Elia G, Ragusa F, et al. Immune and inflammatory cells in thyroid cancer microenvironment. *Int J Mol Sci* (2019) 20(18):4413. doi: 10.3390/ijms20184413
85. Wang Y, Guan Q, Lao I, Wang L, Wu Y, Li D, et al. Using deep convolutional neural networks for multi-classification of thyroid tumor by histopathology: a large-scale pilot study. *Ann Transl Med* (2019) 7(18):468. doi: 10.21037/atm.2019.08.54
86. Kakudo K, Bychkov A, Bai Y, Li Y, Liu Z, Jung CK. The new 4th edition World Health Organization classification for thyroid tumors, Asian perspectives. *Pathol Int* (2018) 68(12):641–64. doi: 10.1111/pin.12737
87. Daniels GH. Follicular variant of papillary thyroid carcinoma: hybrid or mixture? *Thyroid* (2016) 26(7):872–4. doi: 10.1089/thy.2016.0244
88. Hodak S, Tuttle RM, Maytal G, Nikiforov YE, Randolph G. Changing the cancer diagnosis: the case of follicular variant of papillary thyroid cancer-primum non nocere and NIFTP. *Thyroid* (2016) 26(7):869–71. doi: 10.1089/thy.2016.0205
89. Borrelli N, Denaro M, Ugolini C, Poma AM, Miccoli M, Vitti P, et al. miRNA expression profiling of 'noninvasive follicular thyroid neoplasms with papillary-like nuclear features' compared with adenomas and infiltrative follicular variants of papillary thyroid carcinomas. *Mod Pathol* (2017) 30(1):39–51. doi: 10.1038/modpathol.2016.157
90. Selvaggi SM. The presence of multinucleated giant cells: Noninvasive follicular thyroid neoplasm with papillary-like nuclear features vs the follicular variant of papillary thyroid carcinoma. *Diagn Cytopathol* (2019) 47(10):1007–10. doi: 10.1002/dc.24259
91. Chain K, Legesse T, Heath JE, Staats PN. Digital image-assisted quantitative nuclear analysis improves diagnostic accuracy of thyroid fine-needle aspiration cytology. *Cancer Cytopathol* (2019) 127(8):501–13. doi: 10.1002/cncy.22120
92. Hsieh AM, Polyakova O, Fu G, Chazen RS, MacMillan C, Witterick IJ, et al. Programmed death-ligand 1 expression by digital image analysis advances thyroid cancer diagnosis among encapsulated follicular lesions. *Oncotarget* (2018) 9(28):19767–82. doi: 10.18632/oncotarget.24833
93. Perrier ND, Brierley JD, Tuttle RM. Differentiated and anaplastic thyroid carcinoma: Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin* (2018) 68(1):55–63. doi: 10.3322/caac.21439
94. Gardezi SJS, Elazab A, Lei B, Wang T. Breast cancer detection and diagnosis using mammographic data: systematic review. *J Med Internet Res* (2019) 21(7):e14464. doi: 10.2196/14464
95. Lui TKL, Guo CG, Leung WK. Accuracy of artificial intelligence on histology prediction and detection of colorectal polyps: a systematic review and meta-analysis. *Gastrointest Endosc* (2020) 92(1):11–22. doi: 10.1016/j.gie.2020.02.033
96. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *Bmj* (2020) 368:m689. doi: 10.1136/bmj.m689

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 25 March 2021
doi: 10.3389/fonc.2021.631649

Predicting Masaoka-Koga Clinical Stage of Thymic Epithelial Tumors Using Preoperative Spectral Computed Tomography Imaging

Qing Zhou[1,2,3], *Xiaoai Ke*[1,3], *Jiangwei Man*[1], *Bin Zhang*[1,2,3], *Furong Wang*[1] *and Junlin Zhou*[1,3]*

[1] Lanzhou University Second Hospital, Lanzhou, China, [2] Second Clinical School, Lanzhou University, Lanzhou, China, [3] Key Laboratory of Medical Imaging of Gansu Province, Lanzhou, China

OPEN ACCESS

Edited by:
Fatima-Zohra Mokrane, Université Toulouse III Paul Sabatier, France

Reviewed by:
Giovannella Palmieri, Federico II University Hospital, Italy
Yuchuan Hu, Tangdu Hospital, China
Alexandre Nguyen, CHU Toulouse Rangueil, France

***Correspondence:**
Junlin Zhou
lzuzjl601@163.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 20 November 2020
Accepted: 03 March 2021
Published: 25 March 2021

Citation:
Zhou Q, Ke X, Man J, Zhang B, Wang F and Zhou J (2021) Predicting Masaoka-Koga Clinical Stage of Thymic Epithelial Tumors Using Preoperative Spectral Computed Tomography Imaging. Front. Oncol. 11:631649. doi: 10.3389/fonc.2021.631649

Objectives: To investigate the utility of spectral computed tomography (CT) parameters for the prediction of the preoperative Masaoka-Koga stage of thymic epithelial tumors (TETs).

Materials and Methods: Fifty-four patients with TETs, aged from 37 to 73 years old, an average age of 55.56 ± 9.79 years, were included in the study.According to the Masaoka-Koga staging method, there were 19 cases of stage I, 15 cases of stage II, 8 cases of stage III, and 12 cases of stage IV disease. All patients underwent dual-phase enhanced energy spectral CT scans. Regions of interest (ROIs) were defined in sections of the lesion with homogeneous density, the thoracic aorta at the same level as the lesion, the outer fat layer of the lesion, and the anterior chest wall fat layer. The single-energy CT value at 40-140 keV, iodine concentration, and energy spectrum curve of all lesion and thoracic aorta were obtained. The energy spectrum CT parameters of the lesions, extracapsular fat of the lesions, and anterior chest wall fat in stage I and stage II were obtained. The energy spectrum CT parameters of the lesions, enlarged lymph nodes and intravascular emboli in the 3 groups were obtained. The slope of the energy spectrum curve and the normalized iodine concentration were calculated.

Results: In stage I lesions, there was a statistically significant difference between the slope of the energy spectrum curve for the lesion and those of the fat outside the lesion and the anterior chest wall in the arteriovenous phase ($P<0.001$, $P<0.001$). The energy spectrum curve of the tumor parenchyma was the opposite of that of the extracapsular fat. In stage II lesions, there was a statistically significant difference between the slope of the energy spectrum curve for the anterior chest wall and those of the lesion and the fat outside the lesion in the arteriovenous phase($P<0.001$, $P<0.001$). The energy spectrum curve of the tumor parenchyma was consistent with that of the extracapsular fat. Distinction between stage I and II tumors be evaluated by comparing the energy spectrum curves of the mass and the extracapsular fat of the mass. The accuracy rate of is 79.4%. For stages III and IV, there was no significant difference in the slope of the energy spectrum curve of the tumor parenchyma, metastatic lymph node, and intravascular embolism ($P>0.05$). The energy

spectrum curve of the tumor parenchyma was consistent with that of the enlarged lymph nodes and intravascular emboli. The two radiologists have strong consistency in evaluating TETs Masaoka-Koga staging, The Kappa coefficient is 0.873,(95%CI:0.768-0.978).

Conclusion: Spectral CT parameters, especially the energy spectrum curve and slope, are valuable for preoperative TET and can be used in preoperative staging prediction.

Keywords: thymic epithelial tumor, spectral CT, imaging, Masaoka-Koga clinical staging, predicts

INTRODUCTION

Thymic epithelial tumors (TETs), which include thymoma and thymic cancer, originate from thymic epithelial cells and are composed of different proportions of epithelial cells and lymphocytes. TETs are the most common tumors in the anterior mediastinum, accounting for 47% of tumors in this location (1). However, their overall incidence is not high, accounting for 1.3–2.2/10^6 cases (2, 3). TETs are common in elderly people but rare in children and have been reported in patients aged 10 to 80 years, with an average age of 50 to 60 years among men and women (4). Although some patients present with paraneoplastic autoimmune diseases, most patients have no clinical symptoms. Among these diseases, myasthenia gravis is the most common, affecting approximately one-third of patients. Surgery remains the most common treatment for TETs. Thoroughness of tumor resection is an important factor that influences the prognosis of patients, as the prognosis is significantly better after complete resection than after incomplete resection (5). Tumor diameter, tumor resection style, World Health Organization (WHO) classification, Masaoka-Koga stage, and postoperative radiotherapy and chemotherapy are independent factors that affect the prognosis of patients with thymoma. As a result of the correlation between the CT imaging characteristics of TETs and clinical stage (6, 7), tumor stage is evaluated before surgery based on imaging parameters and appropriate treatment strategies are applied. Computed tomography (CT) is currently the preferred method for identifying and characterizing thymic tumors and provides higher sensitivity and specificity (8). The current CT-based staging approach for TETs is based on the appearance of tumor images, and they remain difficult to accurately stage; magnetic resonance imaging (MRI) combined with diffusion-weighted imaging and apparent diffusion coefficient values can also be used in quantitative analyses (9). The staging of TETs using MRI is more accurate than that when using CT; however, MRI scanning requires higher, for example, longer inspection time and restrictions on metal foreign bodies in the body. More convenient, non-invasive, and quantifiable indicators are necessary for preoperative staging.

There are many staging systems for TETs (10), and the most commonly used is the Masaoka-Koga staging system. The Masaoka-Koga staging system is based on the tumor invasion range, surrounding structure invasion and implantation, and distant metastasis, among other factors (8, 11, 12). The latest staging system is the eighth edition of the TNM staging system (13). A recent study (14) collected 217 responses from 37 countries in four continents; 78% of scholars thought the TNM classification was useful (N=169), and 87% of scholars still use the Masaoka-Koga staging system (N=189). Staging is related to the prognosis of TETs (9, 15), and the treatment of different stages of disease are also different. Surgery is generally considered an effective treatment method (16, 17); however, some researchers have pointed out that for some patients with thymoma, surgical treatment may not be the best treatment method (18). It has been reported that for patients with stage I thymoma according to the Masaoka-Koga staging, open surgical resection carries the risk of potential surgical damage and other complications, and CT-guided percutaneous radiofrequency treatment of stage I thymoma is related to minor trauma, fewer complications, and good treatment results (19). In patients with Masaoka-Koga stage I and II thymoma, after appropriate surgical resection, auxiliary radiation treatment has no obvious effect on outcomes (20). Platinum-based chemotherapy remains the standard first-line treatment for patients with advanced or metastatic TETs (21).

Based on the correlation between stage and tumor treatment decisions and prognosis, it is necessary to use non-invasive imaging methods to predict the stage before surgery; CT is currently the preferred method for identifying and characterizing thymic tumors, providing higher sensitivity and specificity (8, 22). Preoperative CT analysis and understanding of the tumor margin, capsule, surrounding invasion levels, pleural implants, lymph nodes, and distant metastases can predict the stage of TETs and have a positive effect on patient treatment and prognosis (23). However, conventional CT is subjective and is based on characterization of the lesion, multiple spectral CT variables can be used to quantify tumor indicators and more accurate preoperative staging.

MATERIALS AND METHODS

This study was approved by the institutional review board, and the requirement for written informed consent was waived due to the retrospective nature of this study.

General Information

Sixty-two patients with surgically and pathologically confirmed TETs treated in our hospital from October 1, 2014 to December

31, 2020 were enrolled; among these, eight patients with insufficient image quality for regions of interest (ROIs) to be drawn were excluded. Finally, 54 patients were included, comprising 27 men and 27 women aged from 37 to 73 years old, with an average age of 55.56 ± 9.79 years. According to the Masaoka-Koga staging system, 19 patients had stage I, 15 patients had stage II, 8 patients had stage III, and 12 patients had stage IV disease. The 54 patients were divided into three groups: non-invasive group (stage I), surrounding fat group (stage II), and surrounding structures and distant metastasis group (stages III and IV). The most common clinical manifestations were intermittent cough, chest tightness, chest pain, and shortness of breath. Eight patients had myasthenia gravis, which manifested as drooping eyelids, limb weakness, and difficulty chewing. Seven of these patients were identified during physical examination, and three patients had recurrence. All patients were treated surgically, and the specimens were sent for pathological examination postoperatively.

Instruments and Methods

All patients underwent plain and dual-phase-enhanced energy spectrum CT (Discovery CT 750 HD; GE Healthcare, Waukesha, WI, USA). Patients were supine during the scan, which was conducted from the apex to the bottom of the lung. The scanning parameters were as follows: flat scan tube at a voltage of 120 kV, enhanced scan tube at a voltage of 80/140 kV with fast switching, rack speed of 0.6 s/r, tube current of 325 mA, pitch of 0.983:1, detector coverage of 40 mm, collimator width of 1.25 mm, and reconstruction layer thickness and layer spacing of 1.25 mm. An iodinated contrast agent (Ultravist 370, Bayer Schering Pharma, Berlin, Germany) was used at a flow rate of 3–4 mL/s and a dose of 1.0 mL/kg injected *via* the anterior cubital vein using a high-pressure syringe (Ulrich Medical, Ulm, Germany). The trigger threshold for thoracic aorta monitoring was 80 Hounsfield units. The arterial phase scan was performed 8 s after the trigger, while the venous phase scan was performed 30 s after the trigger.

CT Image Analysis

Two radiologists used the AW 4.6 workstation GSI Viewer software to perform CT scans. ROIs 5×5 mm in size were delineated in homogeneous lesions while avoiding necrotic cysts, calcification, and vascular shadows. ROIs in the thoracic aorta, the fat layer within 10 mm outside the lesion capsule, and the fat layer on the anterior chest wall were delineated with 5×5-mm ROIs. Enlarged lymph nodes and intravascular emboli were outlined by ROIs of the same size. Each ROI was drawn three times to obtain an average value. The single-energy CT value, energy spectrum curve, and iodine concentration in the ROI was recorded. The standardized iodine concentration (NIC) was determined using the IC_{lesion}/IC_{chest} equation, where IC_{lesion} is the iodine concentration in the lesion and IC_{chest} is the iodine concentration in the thoracic aorta. The slope was calculated as $(CT_{40}–CT_{100})/(40–100)$, where CT_{40} is the CT value at 40 keV and CT_{100} is the CT value at 100 keV.

Pathological Images

Pathological images were first observed under low and high magnification by a pathologist with 3–5 years of work experience. Diagnosis was based on the WHO histological classification. Microscopic observation was performed to assess whether the tumor had invaded the surrounding fat and connective tissue, and whether it had invaded the mediastinal pleura, pericardium, and lungs. It was also clarified whether an intravascular embolus was a thrombus or a tumor, and whether there was metastasis to the lymph nodes. The diagnosis was confirmed by a pathologist with more than 10 years of work experience.

Masaoka-Koga Stage

In combination with the postoperative pathology, tumor staging was performed based on whether the tumor capsule was intact during surgery, the surrounding fatty layer was clear, and the mediastinal pleura, pericardium, lungs, and blood vessels were adherent and extensively invaded (8, 11, 12). Stage I tumors were defined as those where the capsule was intact and no extracapsular invasion was observed under the microscope. Stage II tumors were defined as those where macroscopic and microscopic invasion of the mediastinal adipose tissue could be observed, with no invasion of the mediastinal pleura or pericardium. Stage III tumors were defined as those with visible invasion of adjacent structures (such as the pericardium, large blood vessels, or lungs) were observed. Stage IV-A tumors were defined as those with pleural dissemination (pleural or pericardial metastasis) (was observed) and stage IV-B tumors were defined as those with lymph or blood-based metastasis to a location outside the thoracic region was observed. Based on the extent of tumor invasion, the tumors were divided into a non-invasive group (stage I), invasion of the surrounding fat group (stage II), and invasion of surrounding structures (pleura, pericardium, lung, blood vessels) and distant metastasis group (stages III and IV).We performed TNM staging based on all data and performed a consistency analysis with Masaoka-Koga staging.

Statistical Methods

SPSS 23.0 software was used for analysis; t-test was used for quantitative comparisons between two groups. Multiple sets of quantitative data were analyzed by one-way analysis of variance, and variances were analyzed using the Kruskal-Wallis test. Pairwise comparisons between groups were performed using the Bonferroni method. The Kappa value was used to evaluate the agreement between two observers. Calculate the accuracy of the energy spectrum CT parameters for TETs staging. $P < 0.05$ indicating that the differences were statistically significant.

RESULTS

For all 54 cases, there was a strong consistency between TNM staging and Masaoka-Koga staging (ICC=0.852). According to TNM staging, there were more stage I cases than according to the Masaoka-Koga staging. In the two staging methods, there was no significant change in the data for stages III and IV. The two radiologists have strong consistency in evaluating TETs Masaoka-Koga staging, The Kappa coefficient is 0.873 (95% CI: 0.768-0.978).

Energy Spectrum CT Results for Three Groups of Lesions

The single-energy CT values at 40-140 keV, NIC, and slope of the energy curve in the non-invasive group (stage I), surrounding fat group (stage II), and surrounding structures (pleura, pericardium, lung, blood vessels) and distant metastases group (stages III and IV) were not significantly different (**Table 1**).

CT Results for Stage I and Stage II Lesions, Lesion Extracapsular Fat, and Anterior Chest Wall Fat

In stage I, the difference in the slope of the energy spectrum curve of tumor parenchyma, tumor extracapsular fat, and anterior chest wall fat was statistically significant ($P < 0.001$, $P < 0.001$) (**Table 2**) (**Figure 1**). In stage II, the difference between the tumor parenchyma, the fat outside the tumor, and the slope of the spectrum curve of the anterior chest wall was statistically significant ($P < 0.001$, $P < 0.001$) (**Table 3**) (**Figure 2**). According to the slope of the energy spectrum curve of the outer fat layer of the lesion, combined with the energy spectrum curve graph, the accuracy of spectral CT for predicting the Masaoka-Koga stage I and II lesions was 79.4%.

CT Results for Stage IV Lesions, Intravascular Tumor Thrombi, and Metastatic Lymph Nodes

There were no significant differences in the slopes of the energy spectrum curves for stage IV lesions, metastatic lymph nodes, and intravascular emboli ($P > 0.05$) (**Table 4**) (**Figures 3** and **4**).The slope of the energy spectrum curve combined with the energy spectrum curve graph, the accuracy of assessing lymph node metastasis and embolic properties is 80.0%.

DISCUSSION

Three parameters were evaluated in this study, the single energy CT value of the lesion, the NIC, and the slope of the energy spectrum curve. The single-energy imaging mode of energy spectrum CT simulates the image of an object in the presence of a monochromatic X-ray source. The attenuation of different tissues changes in accordance with changes in the X-ray beam energy. A lower single-energy beam can increase the density resolution of the image, which helps to show lesions. The slope of the energy spectrum curve reflects the magnitude of the enhancement. The larger the slope is, the greater the magnitude of the enhancement will be, and the energy spectrum curve of the tissue changes greatly at low energy. Therefore, low energy (40-100 keV) was selected to determine the slope of the energy spectrum curve for each group. The shape-like energy spectrum curve may reflect the homology of the region of interest. The iodine (water) concentration of the lesion was directly obtained from the material separation map for the energy spectrum CT, which quantitatively reflects the iodine uptake and distribution in the lesion. To eliminate individual differences, the NIC was used to compare the iodine (water) base values. The NIC was more reliable than the iodine concentration (24). NIC is an indirect reflection of the iodine content of a lesion, that is, the degree of strengthening.

The 54 patients evaluated in this study included 19 with stage I, 15 with stage II, 8 with stage III lesions, and 12 with stage IV lesions according to the Masaoka-Koga staging. The 54 patients were divided into three groups: non-invasive group (stage I),

TABLE 1 | Spectral CT multi-parameter results of the three groups in stages (Arterial phase).

Parameters	Non-invasive group (stage I)	Invasion of the surrounding fat group (stage II)	Invasion of surrounding structures and distant metastasis groups (stage III和IV)	F/H	P
40 keV	179.51 ± 48.56	177.28 ± 43.52	169.24 ± 39.12	0.291	0.749
50 keV	130.22 ± 31.71	129.98 ± 28.93	123.04 ± 25.68	0.373	0.691
60 keV	98.86 ± 21.42	99.53 ± 21.25	92.15 ± 18.06	0.744	0.481
70 keV	79.87 ± 16.32	80.90 ± 15.92	74.10 ± 13.39	1.044	0.360
80 keV	68.60 ± 14.05	70.30 ± 12.96	64.67 ± 10.92	0.892	0.416
90 keV	61.11 ± 12.61	63.63 ± 11.55	58.23 ± 10.35	0.889	0.418
100 keV	55.56 ± 12.00	58.28 ± 10.84	53.06 ± 9.81	0.912	0.408
110 keV	51.65 ± 11.76	54.49 ± 10.44	49.38 ± 9.59	0.915	0.408
120 keV	48.89 ± 11.71	51.88 ± 10.23	46.80 ± 9.51	0.916	0.407
130 keV	46.77 ± 11.72	49.85 ± 10.15	44.81 ± 9.48	0.906	0.411
140 keV	45.13 ± 11.75	48.26 ± 10.07	43.27 ± 9.49	0.891	0.417
NIC	0.15 ± 0.05	0.12 ± 0.04	0.13 ± 0.04	1.191	0.313
Curve slope	-2.06 ± 0.74	-1.98 ± 0.63	-1.93 ± 0.58	0.191	0.827

TABLE 2 | Slopes of energy spectrum curves, extracapsular fat and anterior chest wall fat of TETs stage I.

	Slope of the energy spectrum curve of the lesion	Slope of energy spectrum curve of extracapsular fat in lesions	Slope of energy spectrum curve of anterior chest wall fat	F/H	*P*
Arterial phase	-2.00 ± 0.76	1.94 ± 0.69*	2.32 ± 0.33	293.52	0.000
Venous phase	-2.10 ± 0.51	1.72 ± 0.48*	2.08 ± 0.33	500.89	0.000

*Compared with the slope of the lesion energy spectrum curve, *$P < 0.05$; Compared with the slope of the anterior chest wall energy spectrum curve, #$P < 0.05$.*

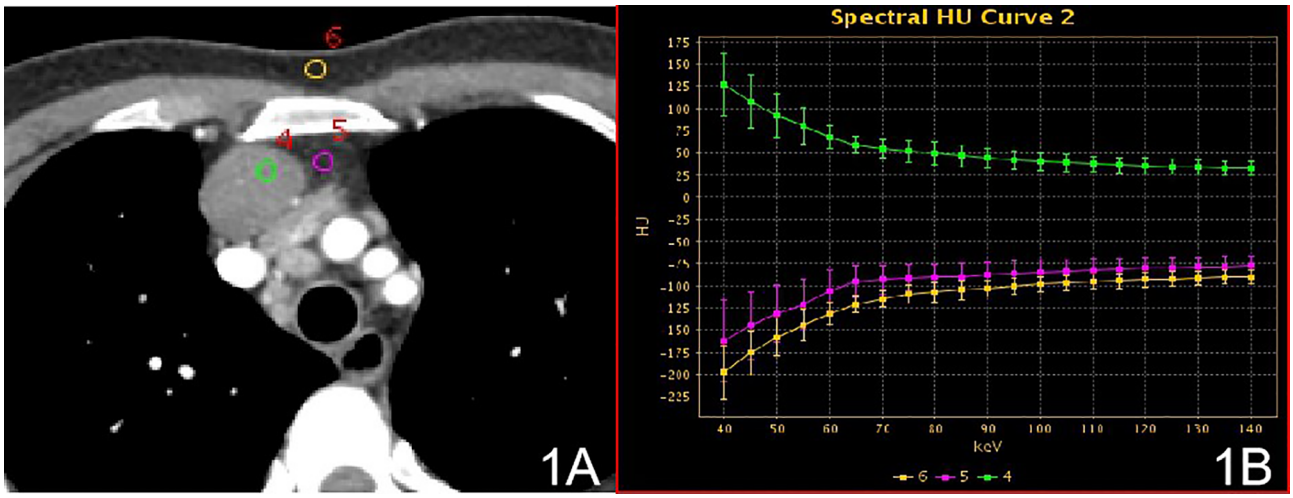


FIGURE 1 | Stage I, **(A)** 70 keV single-energy CT value of the arterial phase is used to outline the ROI of the lesion; **(B)** spectrum of arterial lesions (green), spectrum of fat outside the lesion (purple) and spectrum of anterior chest wall fat (yellow).

TABLE 3 | Slopes of energy spectrum curves of, extracapsular fat and anterior chest wall fat TETs stage II.

	Slope of the energy spectrum curve of the lesion	Slope of energy spectrum curve of extracapsular fat in lesions	Slope of energy spectrum curve of anterior chest wall fat	F/H	P
Arterial phase	-1.91 ± 0.66	-1.27 ± 0.57*	2.18 ± 0.27	223.54	0.000
Venous phase	-2.27 ± 0.48	-1.46 ± 0.53*#	2.05 ± 0.28	409.71	0.000

*Compared with the slope of the lesion energy spectrum curve, *P < 0.05; Compared with the slope of the anterior chest wall energy spectrum curve, #P < 0.05.*

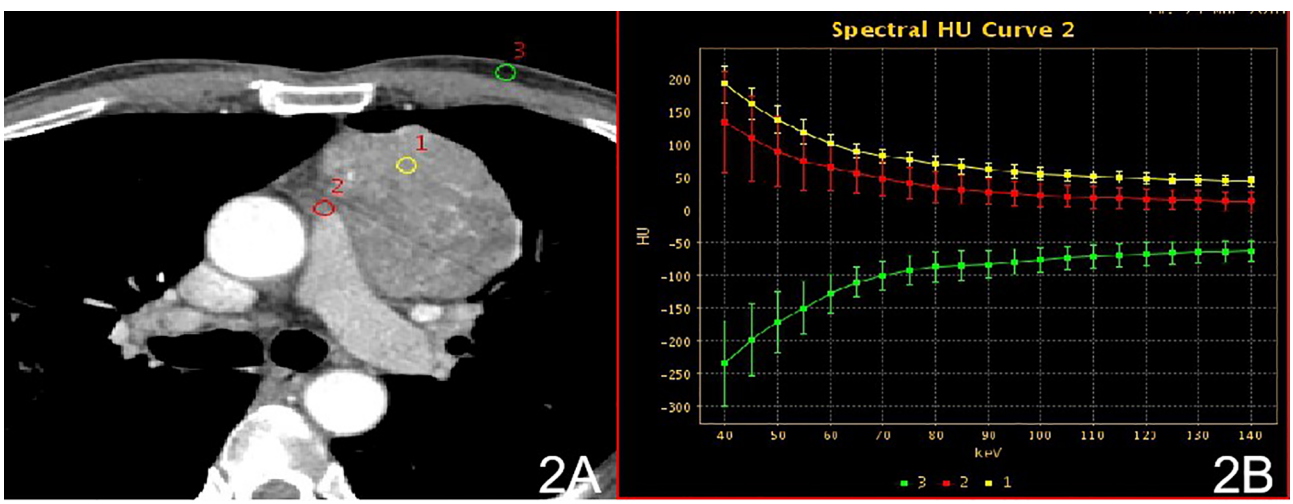


FIGURE 2 | Stage II, **(A)** 70 keV single-energy CT value of the arterial phase is used to outline the ROI of the lesion; **(B)** arterial lesion spectrum (yellow), extra-focal fat spectrum (red) and anterior chest wall fat spectrum (green).

invasive surrounding fat group (stage II), and invasive surrounding structures (pleura, pericardium, lung, blood vessels) and distant metastasis group (stages III and IV). In the qualitative diagnosis, irregular or lobed edges lesions on conventional CT. It may indicate that the tumor capsule and extracapsular fat have been invaded (25). In this study, ROIs were delineated in stage I and stage II solid lesions, extracapsular fat, and anterior chest wall fat. The energy spectrum curves for

TABLE 4 | Slope analysis of spectrum curves of lesions, metastatic lymph nodes, and intravascular emboli in group 3 (stages III and IV).

	Slope of the energy spectrum curve of the lesion	Slope of the energy spectrum curve of metastatic lymph nodes	Slope of the energy spectrum curve of intravascular emboli	F/H	P
Arterial phase	-1.83 ± 0.64	-1.53 ± 0.68	-1.45 ± 0.65	0.964	0.393
Venous phase	-1.72 ± 0.59	-1.69 ± 0.57	-1.37 ± 0.77	2.482	0.101

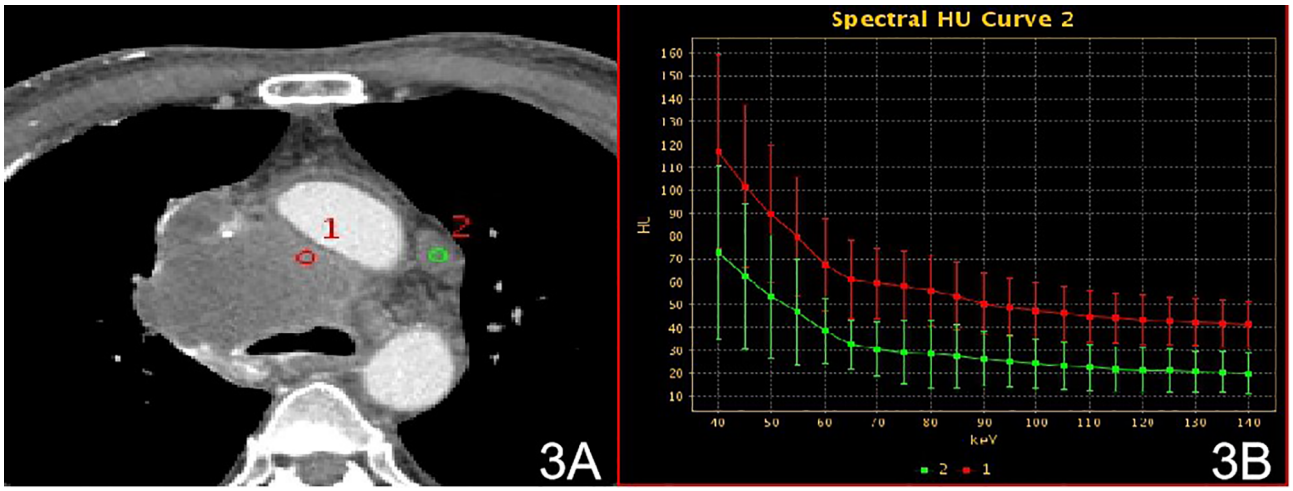


FIGURE 3 | Stage IV, **(A)** 70 keV single-energy CT value of the arterial phase is used to outline the ROI of the lesion; **(B)** arterial lesion spectrum (red), lymph node spectrum (green).

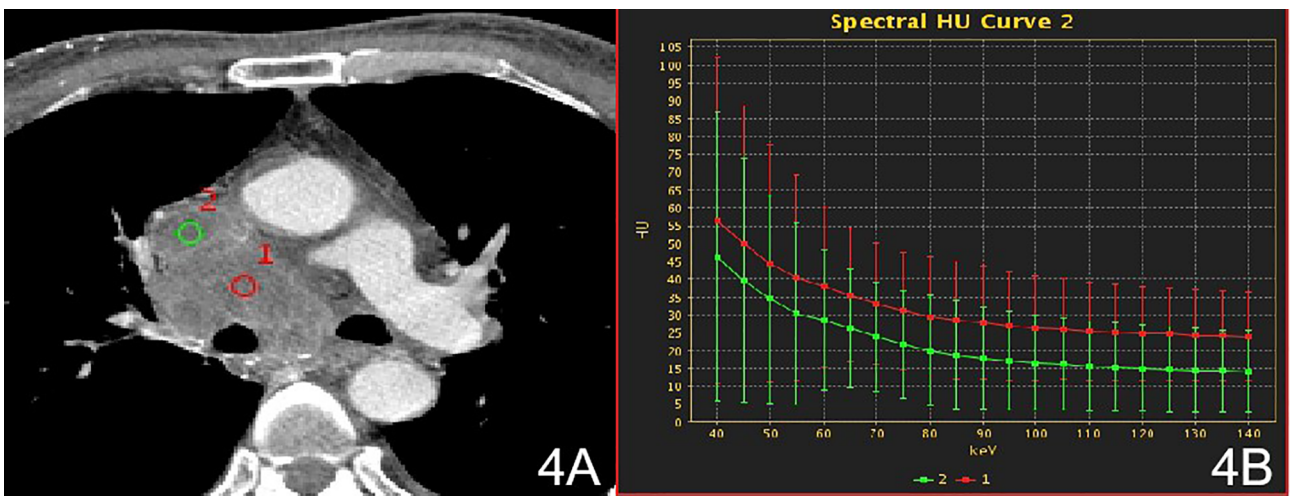


FIGURE 4 | Stage IV, **(A)** 70 keV single-energy CT value of the arterial phase is used to outline the ROI of the lesion; **(B)** Arterial lesion spectrum (red), intravascular emboli spectrum (green).

the lesion, extracapsular fat, and anterior chest wall fat were obtained. In stage I lesions, the energy spectrum curve of the tumor parenchyma and the energy spectrum curve of the extracapsular fat have the opposite shape, and the energy spectrum curve of the extracapsular fat and the front chest wall fat have the same shape, indicating that the lesion did not extend outside the capsule. The slope of the spectrum curve for the anterior chest wall fat in stage II lesions showed statistically

significant differences with those of the spectra for the lesion and the lesion's extracapsular fat, indicating that the lesion had invaded the extracapsular fat. According to the spectrum curve, the lesion and the fat in the outer capsule of the lesion were homologous. In this study, ROIs were delineated in pathologically confirmed mediastinal lymph node with metastasis, intravascular tumor emboli, and solid lesions. The spectral curves for the metastatic lymph nodes, intravascular emboli, and lesions were consistent and the differences in the slopes were not statistically significant, indicating that the lesions were homologous to lymph nodes and intravascular emboli (26). Among these parameters, the energy spectrum curve objectively reflects whether the tumor parenchyma and extracapsular fat are homologous, and whether the enlarged lymph nodes and emboli are metastatic.

In conclusion, distinction between stage I and II tumors may be evaluated by comparing the energy spectrum curves of the mass and the extracapsular fat of the mass. Areas of interest should be outlined in suspicious enlarged lymph nodes and intravascular emboli before surgery to determine whether the lymph nodes have metastasized and the nature of the emboli in the vessel.

Limitations

The limitations of our study are as follows: This study was a retrospective and single-center study, lacking multi-center data verification and the sample size was small. The accuracy of preoperative predicting staging needs further large sample size research. Measurement of the ROI of fat outside the capsule stays a difficult points, even though we used the average value of 3 ROIs of the same size drawn in the fat layer 10 mm outside the capsule of the lesion to reflect as much as possible whether the lesion has extracapsular fat invasion.

DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this published article.

AUTHOR CONTRIBUTIONS

QZ designed the study, performed the literature search and data analysis, and drafted the manuscript. JZ and QZ participated in the conception and design of the study, performed the quality assessment and helped to revise the manuscript. JM, FW, BZ and XK performed the study selection and statistical analysis and helped to draft the manuscript. QZ, JM, XK, BZ, and FW performed the data extraction and analysis, modified the language. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported the National Natural Science Foundations of China(81772006).

ACKNOWLEDGMENTS

Thanks to all the partners who contributed to this research, including Qing Zhou, Jiangwei Man, Furong Wang, Xiaoai Ke, Junlin Zhou. Special thanks to Ms. Furong Wang from the Pathology Department and Mr. Jiangwei Man from the Department of Surgery.

REFERENCES

1. Chen JL, Weisbrod GL, Herman SJ. Computed tomography and pathologic correlations of thymic lesions. *J Thorac Imaging* (1988) 3(1):61–5. doi: 10.1097/00005382-198801000-00010
2. Scorsetti M, Leo F, Trama A, D'Angelillo R, Serpico D, Macerelli M, et al. Thymoma and thymic carcinomas. *Crit Rev Oncol Hematol* (2016) 99:332–50. doi: 10.1016/j.critrevonc.2016.01.012
3. Engels EA. Epidemiology of thymoma and associated malignancies. *J Thorac Oncol* (2010) 5(10 Suppl 4):S260–5. doi: 10.1097/JTO.0b013e3181f1f62d
4. den Bakker MA, Roden AC, Marx A, Marino M. Histologic classification of thymoma: a practical guide for routine cases. *J Thorac Oncol* (2014) 9(9 Suppl 2): S125–30. doi: 10.1097/JTO.0000000000000297
5. Banna GL, Sheel A, Sheel V, Bille A, Routledge T, Fernando S, et al. Treatment and prognostic factors of patients with thymic epithelial tumors at first recurrence or progression. *Future Oncol* (2017) 13(27):2429–39. doi: 10.2217/fon-2017-0236
6. Iannarelli A, Sacconi B, Tomei F, Anile M, Longo F, Bezzi M, et al. Analysis of CT features and quantitative texture analysis in patients with thymic tumors: correlation with grading and staging. *Radiol Med* (2018) 123(5):345–50. doi: 10.1007/s11547-017-0845-4
7. Qu YJ, Liu GB, Shi HS, Liao MY, Yang GF, Tian ZX. Preoperative CT findings of thymoma are correlated with postoperative Masaoka clinical stage. *Acad Radiol* (2013) 20(1):66–72. doi: 10.1016/j.acra.2012.08.002
8. Masaoka A, Monden Y, Nakahara K, Tanioka T. Follow-up study of thymomas with special reference to their clinical stages. *Cancer* (1981) 48 (11):2485–92. doi: 10.1002/1097-0142(19811201)48:11<2485::AID-CNCR2820481123>3.0.CO;2-R
9. Kong LY, Zhang W, Zhou Y, Xu H, Shi HB, Feng Q, et al. Histogram analysis of apparent diffusion coefficient maps for assessing thymic epithelial tumours: correlation with world health organization classification and clinical staging. *Br J Radiol 2018* (1084) 91:20170580. doi: 10.1259/bjr.20170580
10. Weissferdt A, Moran CA. Staging of thymic epithelial neoplasms: thymoma and thymic carcinoma. *Pathol Res Pract* (2015) 211(1):2–11. doi: 10.1016/j.prp.2014.06.007
11. Koga K, Matsuno Y, Noguchi M, Mukai K, Asamura H, Goya T, et al. A review of 79 thymomas: modification of staging system and reappraisal of conventional division into invasive and non-invasive thymoma. *Pathol Int* (1994) 44(5):359–67. doi: 10.1111/j.1440-1827.1994.tb02936.x
12. Detterbeck FC, Nicholson AG, Kondo K, Van Schil P, Moran C. The Masaoka-Koga stage classification for thymic malignancies: clarification and definition of terms. *J Thorac Oncol* (2011) 6(7 Suppl 3):S1710–6. doi: 10.1097/JTO.0b013e31821e8cff
13. Detterbeck FC, Stratton K, Giroux D, Asamura H, Crowley J, Falkson C, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposal for an evidence-based stage classification system for the forthcoming (8th) edition of the TNM classification of malignant tumors. *J Thorac Oncol* (2014) 9:S65–72. doi: 10.1097/JTO.0000000000000290
14. Ruffini E, Fang W, Guerrera F, Huang J, Okumura M, Kim DK, et al. The IASLC Thymic Tumors Staging Project. The Impact of the 8th Edition of the UICC/AJCC TNM Stage Classification of Thymic Tumors: Results of a Survey. *J Thorac Oncol* (2019) 15(3):436–47. doi: 10.1016/j.jtho.2019.11.013

15. Tassi V, Vannucci J, Ceccarelli S, Gili A, Matricardi A, Avenia N, et al. Stage-related outcome for thymic epithelial tumours. *BMC Surg* (2019) 18(Suppl 1):114. doi: 10.1186/s12893-018-0434-z
16. Detterbeck F, Youssef S, Ruffini E, Okumura M. A review of prognostic factors in thymic malignancies. *J Thorac Oncol* (2011) 6(7 Suppl 3):S1698–704. doi: 10.1097/JTO.0b013e31821e7b12
17. Wright CD, Wain JC, Wong DR, Donahue DM, Gaissert HA, Grillo HC, et al. Predictors of recurrence in thymic tumors: importance of invasion, World Health Organization histology, and size. *J Thorac Cardiovasc Surg* (2005) 130 (5):1413–21. doi: 10.1016/j.jtcvs.2005.07.026
18. Koppitz H, Rockstroh JK, Schüller H, Standop J, Skowasch D, Müller-Hermelink HK, et al. State-of-the-art classification and multimodality treatment of malignant thymoma. *Cancer Treat Rev* (2012) 38(5):540–8. doi: 10.1016/j.ctrv.2011.11.010
19. Dong J, Yuan S, Chang B, Huang J, Geng X, Cai X, et al. An exploratory study of CT-guided percutaneous radiofrequency ablation for stage I thymoma. *Cancer Imaging* (2019) 19(1):80. doi: 10.1186/s40644-019-0267-8
20. Gielda BT, Peng R, Coleman JL, Thomas CR, Cameron RB. Treatment of early stage thymic tumors: surgery and radiation therapy. *Curr Treat Options Oncol* (2008) 9(4-6):259–68. doi: 10.1007/s11864-008-0080-2
21. Conforti F, Pala L, Giaccone G, De Pas T. Thymic epithelial tumors: From biology to treatment. *Cancer Treat Rev* (2020) 86:102014. doi: 10.1016/j.ctrv.2020.102014
22. Zhao Y, Chen H, Shi J, Fan L, Hu D, Zhao H. The correlation of morphological features of chest computed tomographic scans with clinical characteristics of thymoma. *Eur J Cardiothorac Surg* (2015) 48(5):698–704. doi: 10.1093/ejcts/ezu475
23. Ottlakan A, Borda B, Morvay Z, Maraz A, Furak J. The Effect of Diagnostic Imaging on Surgical Treatment Planning in Diseases of the Thymus. *Contrast Media Mol Imaging* (2017) 2017:9307292. doi: 10.1155/2017/9307292
24. Vogel J, Lin L, Litzky LA, Berman AT, Simone CB 2nd. Predicted Rate of Secondary Malignancies Following Adjuvant Proton Versus Photon Radiation Therapy for Thymoma. *Int J Radiat Oncol Biol Phys* (2017) 99(2):427–33. doi: 10.1016/j.ijrobp.2017.04.022
25. Jung KJ, Lee KS, Hun J, Kim J, Kim TS, Kim EA, et al. Malignant thymic epithelial tumors: CT·pathologic correlation. *AJR Am J Roentgenol* (2001) 176:433–9. doi: 10.2214/ajr.176.2.1760433
26. Aran S, Daftari Besheli L, Karcaaltincaba M, Gupta R, Flores EJ, Abujudeh HH. Applications of dual-energy CT in emergency radiology. *AJR Am J Roentgenol* (2014) 202:W314–24. doi: 10.2214/AJR.13.11682

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

REVIEW
published: 29 March 2021
doi: 10.3389/fonc.2021.633176

Understanding Sources of Variation to Improve the Reproducibility of Radiomics

Binsheng Zhao *

Department of Radiology, Columbia University Irving Medical Center, New York, NY, United States

Radiomics is the method of choice for investigating the association between cancer imaging phenotype, cancer genotype and clinical outcome prediction in the era of precision medicine. The fast dispersal of this new methodology has benefited from the existing advances of the core technologies involved in radiomics workflow: image acquisition, tumor segmentation, feature extraction and machine learning. However, despite the rapidly increasing body of publications, there is no real clinical use of a developed radiomics signature so far. Reasons are multifaceted. One of the major challenges is the lack of reproducibility and generalizability of the reported radiomics signatures (features and models). Sources of variation exist in each step of the workflow; some are controllable or can be controlled to certain degrees, while others are uncontrollable or even unknown. Insufficient transparency in reporting radiomics studies further prevents translation of the developed radiomics signatures from the bench to the bedside. This review article first addresses sources of variation, which is illustrated using demonstrative examples. Then, it reviews a number of published studies and progresses made to date in the investigation and improvement of feature reproducibility and model performance. Lastly, it discusses potential strategies and practical considerations to reduce feature variability and improve the quality of radiomics study. This review focuses on CT image acquisition, tumor segmentation, quantitative feature extraction, and the disease of lung cancer.

Keywords: radiomics, lung cancer, reproducibility, variability, CT acquisition, tumor segmentation, feature extraction, quality control

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Sylvain Reuzé, GE Healthcare, France
Chuan Zhou, University of Michigan, United States

***Correspondence:**
Binsheng Zhao
bz2166@cumc.columbia.edu

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *24 November 2020*
Accepted: *01 March 2021*
Published: *29 March 2021*

Citation:
Zhao B (2021) Understanding Sources of Variation to Improve the Reproducibility of Radiomics. Front. Oncol. 11:633176. doi: 10.3389/fonc.2021.633176

INTRODUCTION

Radiomics refers to the determination of tumor imaging phenotypes by extracting and analyzing a large number of quantitative image features, a.k.a. radiomics features (1–3). Unlike molecular- and tissue-based analyses, radiomics strives to characterize differences in tumor phenotypes based on non-invasive radiographic images that can be routinely obtained from clinical practice and clinical trials. Radiomics can capture the heterogeneity of a whole tumor and tumor metastases in multiple body sites and their surrounding tissues, and it can be used to monitor changes in tumor biology (e.g., mutation status) over time. Thus, radiomics is promising to be capable of addressing key issues across the continuum of cancer care.

The hypothesis underpinning radiomics is that disease processes, which produce histopathological and genetic alterations, also manifest in characteristic phenotypes that can be captured by radiographic images. Qualitative visual interpretation of CT features has been used by radiologists in making routine diagnoses for decades, such as lung nodules with spiculated edges indicating malignancy and an enlarged tumor size (diameter) post-therapy indicating a worse prognosis for the treatment. However, the big moment for cancer imaging phenotype was the 2007 article on the reconstruction of global gene expression profiles of hepatocellular carcinoma (HCC) using predefined imaging traits assessed qualitatively by radiologists on contrast-enhanced CT (CECT) (4). A new radiogenomic venous invasion scoring system, derived from three imaging traits (internal arteries, hypodense halos, and tumor-liver difference) on CECT in HCC, was reported to serve as a noninvasive imaging biomarker for histological microvascular invasion, a tissue biomarker associated with early disease recurrence and poor overall survival (5). While human eyes have an incredible ability to recognize both local and global patterns, visual interpretations can be subjective and prone to variation especially when evaluating subtle differences. Radiomics can objectively discern clinically relevant information that human eyes cannot even perceive. Indeed, a fast-growing literature shows the great promise of radiomics signatures (radiomics features and models) as a "virtual biopsy" to assist in cancer diagnosis and prognosis, treatment plan, patient stratification, and assessment of tumor response to therapy. The current status of CT-based radiomics in lung cancer has been well summarized in a recent collection of review articles [e.g., (6–17)].

Radiomics features are well defined, and some are even intuitive (in line with expert radiologists' visual interpretation). Radiomics analysis is a favorable approach for studying tumor imaging phenotypes because performing it requires a relatively small number of patients to train models compared to convolutional neural networks (CNNs), and sometimes it yields explainable analysis results as well. However, multiple sources of variation in every step of the radiomics workflow create an intrinsic methodological weakness that has been recognized since the earliest days of radiomics analysis (18, 19) (**Figure 1**). For instance, radiomics features can be sensitive to heterogeneous image acquisition settings (scanners, scanning techniques, and reconstruction parameters). Unknown ground truth of tumor boundaries can introduce uncertainty into features derived from segmented tumors. Despite an explosive increase in the radiomics literature, this research frequently fails to adequately consider sources of variation and reports isolated results not validated by replication in external data sets (20). The resulting concerns about rigor and reproducibility slow the pace of innovation in radiomics and limit its translational potential.

Recognizing the need to evaluate the scientific merit and clinical utility of radiomics studies, a group of scientists proposed a radiomics quality score (RQS) in 2017 (21) which evaluates a set of essential components in the radiomics workflow, starting with the quality of image protocol and ending with the availability of open science and data. A maximum of 36 possible points is awarded by scoring each component's

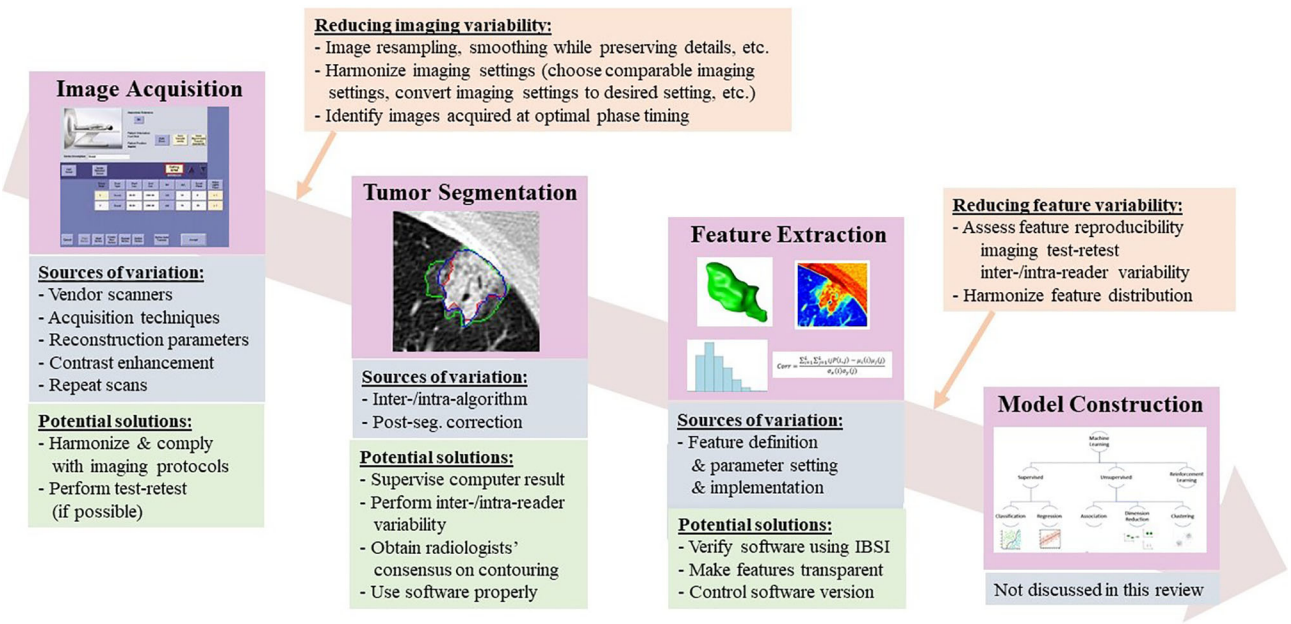


FIGURE 1 | Radiomics workflow, along with sources of variation and potential strategies to reduce feature variability.

accordance to the suggested guidelines, with more important aspects earning more points. While the RQS is not perfect, it does establish a set of practices that can facilitate clinical translation of radiomics research. It also highlights the weakness of the current literature: the mean RQS scores of published radiomics studies are low (<10 points) (22, 23), indicating inadequate scientific quality.

The quality of radiomics studies has recently improved thanks to community-wide efforts to explore and reduce variability in medical imaging and to promote the translation of quantitative imaging biomarkers into clinical practice and clinical trials (24–27) (**Figure 2**). **Figures 2A, B** are drawn based on a research team's recent literature search for the CT-based radiomics studies in lung cancer (6), which we supplemented with studies published as of July, 2020 as well as information about imaging parameters (slice thickness, reconstruction kernel) and segmentation (inter-/intra-variability, software, result supervised or not) (**Table 1** in **Supplementary Material**). Although previous imaging studies have shown the effects of slice thickness and reconstruction kernel on computed features, between ~5% and ~25% of radiomics studies prior to 2020 did not report their study imaging protocols (**Figure 2A**, green color). Most of those who reported their imaging protocols only included the slice thickness information (**Figure 2A**, blue color). It is good to see that the trend of reporting both slice thickness and reconstruction kernel increased from 10% in 2016 to 50% in 2020 (**Figure 2A**, pinkish-orange color). Nevertheless, half of the radiomics studies still do not seem to have considered the effects of reconstruction kernel on radiomics features, especially texture features. The percentages of studies that performed imaging test-retest and inter-/intra-segmentation have been stable over the years, varying between ~20% - 40% (**Figure 2B**, gray and yellow colors). All radiomics studies published in 2019 and 2020 reported human supervision of tumor segmentation (**Figure 2B**, green color), an important step to ensure the accuracy of segmentation, while only ~40% studies did so in 2016.

A valid quantitative imaging biomarker must be informative, or sensitive to underlying biology, as well as reproducible and reliable across various image acquisition settings and quantitative methods. It is essential to understand and regulate the sources of variation to ensure that consistent high quality images can be meaningfully analyzed and biological information can be reliably extracted by advanced quantitative methods. This article starts with image acquisition, then considers tumor segmentation and feature extraction. Readers who are interested in machine learning for radiomics are referred to (10, 16, 17). From the point of view of image analysis, in each of the following sections, it first illustrates how radiomics features can be affected by various factors using demonstrative examples, then reviews a number of published studies exploring sources of variation and offering increased reproducibility of radiomics features and models. Lastly, it discusses potential strategies and practical considerations to reduce feature variability and improve the quality of radiomics studies.

IMAGE ACQUISITION

Radiomics signatures aim to characterize the phenotypes of tumors and surrounding tissue using radiographic images. They can be sensitive to image quality governed by image acquisition settings, or the constellation of factors involved in acquiring the images, which include (but is not limited to) scanner equipment, acquisition techniques, reconstruction parameters, and contrast administration.

Radiomics studies have mostly used retrospective analysis of imaging data from historical studies and clinical trials that were not designed for quantitative feature analysis of tumors. Many of the images studied were acquired in clinical trials to make simple

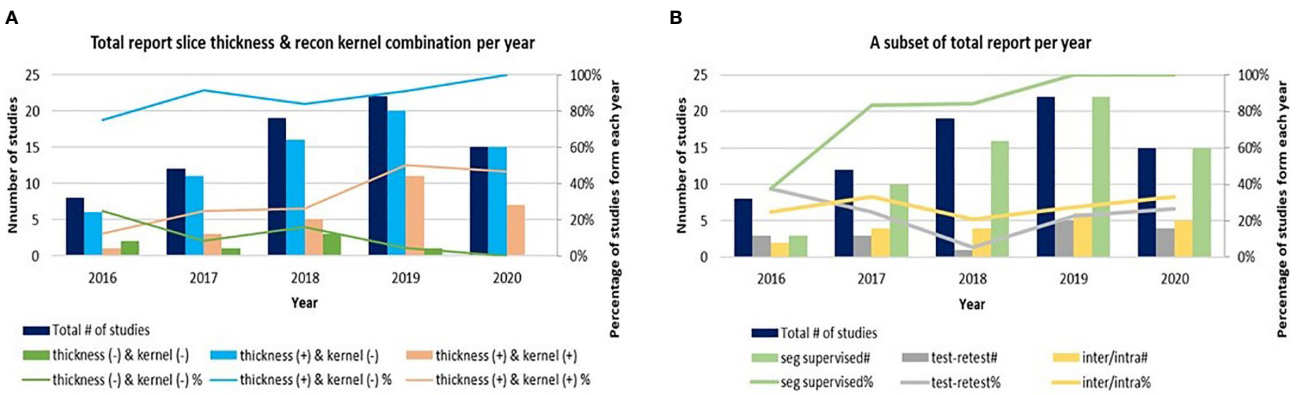


FIGURE 2 | **(A)** A trend of radiomics studies reporting image acquisition parameters of slice thickness, reconstruction kernels and both. **(B)** A trend of radiomics studies reporting re-imaging, re-segmentation and supervised segmentation. Due to the small number of radiomics studies published in 2014 and 2015, those studies are excluded from the graphs.

measurements of tumor diameter on CT images, which did not demand a high degree of standardization in image acquisition parameters. Because these datasets are now being radiomically analyzed retrospectively, and new data sets are being acquired prospectively, the importance of the degree of variation in CT acquisition needs to be determined.

Pioneering efforts revealed that imaging variables, such as repeat CT scans (28), imaging reconstruction slice thicknesses and kernels (29), and scanners (30), could affect the computed values of radiomics features. These studies inspired intensive investigations in feature variability and reproducibility, which have confirmed the initial findings and extended them to broader research areas. Investigations on the sources of variation in CT image acquisition have mainly focused on one or combinations of the following factors: test-retest (28, 31, 32), vendors' scanner (30, 33–36), tube voltage and current (37–41), pitch (36), field of view/pixel spacing (42–44), reconstruction kernel and slice thickness (we do not here distinguish between slice thickness and slice interval, the real physical distance between any two adjacent images) (29, 31, 38, 39, 45–47), contrast administration (48–50), and 4D phases (51, 52). In the following subsections, a number of studies exploring sources of variation in image acquisition is reviewed, followed by a discussion on potential strategies and practical considerations to reduce variability in image acquisition.

In Vivo "Same-Day" Repeat CT Studies

Radiomics features derived from tumor images from two CT scans performed on the same day or during a short time period can be different due to factors such as the patient's relocation, breath holding, and organ movement, even though no biological changes would be expected to be discernable during such a short time period. The reproducibility of radiomics features on repeat CT scans must be demonstrated in order to establish the reliability of radiomics models built using these features.

Repeat CT in Lung Cancer

Early radiomics studies already took into account the effects of repeat CT imaging and re-segmentation on features' reproducibility (3, 28, 53), thanks to the availability of The Reference Image Database to Evaluate Therapy Response's Lung CT Collection (RIDER Lung) (54, 55). RIDER Lung is a unique, publicly available same-day repeat CT image dataset that allows exploration of the reproducibility of quantitative methods, including segmentation and feature extraction, for lung cancer studies. This dataset consists of 31 non-small cell lung cancer (NSCLC) patients' repeat CT scan images reconstructed using 1.25 mm slice thickness and the lung kernel. Unfortunately, RIDER Lung is suboptimal as test-retest for radiomics studies because CT images in the majority of clinical studies were not reconstructed using 1.25 mm slice thickness and the lung kernel.

In order to explore reproducibility and variability in radiomics features due to re-imaging at multiple acquisition settings with same or different imaging parameters, investigators published a pilot study on 89 commonly used radiomics features using same-day repeat CT scan images reconstructed at six imaging settings/series: a combination of three slice thicknesses (1.25 mm, 2.5 mm, 5 mm) and two reconstruction kernels (lung (L): a sharp kernel; standard (S): a smooth kernel) (31). These settings cover the CT acquisition parameters widely used in lung cancer oncology trials and clinical practice. **Figure 3A** shows an example of a lung cancer tumor captured on a CT scan that was reconstructed using six different imaging settings. Given the same slice thickness, tumor heterogeneity can be better seen on sharper images than on smoother ones. The curves beneath the tumor images show the values of two popular GLCM features, Contrast (**Figure 3B**, blue color) and Correlation (**Figure 3B**, orange color), calculated under each imaging setting. The bigger the value of Contrast, the more heterogeneous the tumor. The greater the value of Correlation, the more homogeneous the tumor. In this example, the value differences were caused by different imaging reconstruction parameters, not by the tumor's underlying biological effects. The study found that the radiomics features were generally reproducible when calculated between two repeat scans reconstructed using the same imaging setting. This is indicated by quite uniformly bright red areas (high concordance correlation coefficient (CCC) values) in **Figure 4A(a)**. However, a substantial amount of variability was observed within the same slice thickness when using standard or lung reconstruction kernels, generating smooth and sharp images respectively, as indicated by large dark areas (low CCC values) mostly centered at the texture features [**Figure 4A(b)**]. The authors' conclusion that smooth and sharp reconstructions should not be treated as interchangeable for radiomics studies has been confirmed by other independent studies (29, 36, 56).

Repeat CT in Rectal Cancer

RIDER Lung was a very well controlled clinical study in which the two repeat non-contrast chest CT scans were performed within 15 minutes using the same imaging protocol on the same scanner. Other radiomics studies also reported good reproducibility when testing their quantitative features using RIDER Lung [e.g., (32, 36)]. However, it is possible that repeat CT scan images of other organs may cause different magnitudes of feature reproducibility. A study found much lower feature reproducibility in rectal cancer than in lung cancer (32). The investigators collected repeat CT scan images from 40 rectal cancer patients in a clinical setting; the interval times between two repeat scans ranged from 5 to 19 days. They reported that only 9/542 features had CCC >0.85 in rectal cancer, whereas 446/542 features had higher CCC values for the test-retest analysis of the RIDER Lung dataset. However, this is not surprising because the longer interval times between the two repeat scans in the rectal cancer study, the possible use of different imaging settings for two repeat scans, and presence of more noise in rectal images could all contribute to the decreased reproducibility.

Four-Dimensional CT (4D CT)

The same-day repeat CT images in the RIDER Lung collection were acquired with each patient holding their breath. Radiotherapy scan images, however, are often acquired under free breathing of the patients. Respiratory motion can cause changes in tumor location, volume, shape and intensity (57)

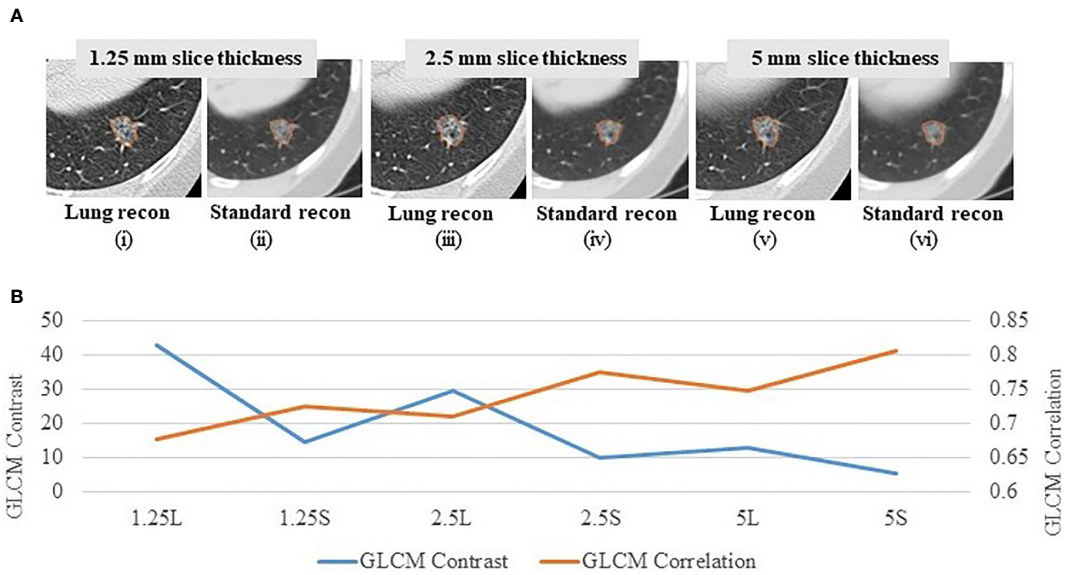


FIGURE 3 | Effects of imaging parameters on radiomics features **(A)** A lung tumor captured on one CT scan reconstructed at 6 different imaging settings: 1.25 mm slice thickness with the lung reconstruction algorithm (sharp image) (1.25L) (i) and the standard reconstruction algorithm (smooth image) (1.25S) (ii); 2.5 mm slice thickness with lung reconstruction (2.5L) (iii) and standard reconstruction (2.5S) (iv); 5 mm slice thickness with lung reconstruction (5L) (v) and standard reconstruction (5S) (vi) (31). **(B)** GLCM Contrast (blue color) and Correlation (orange color) features computed at the 6 corresponding imaging reconstruction settings. The bigger the value of Contrast, the more heterogeneous the tumor. The greater the value of Correlation, the more homogeneous the tumor.

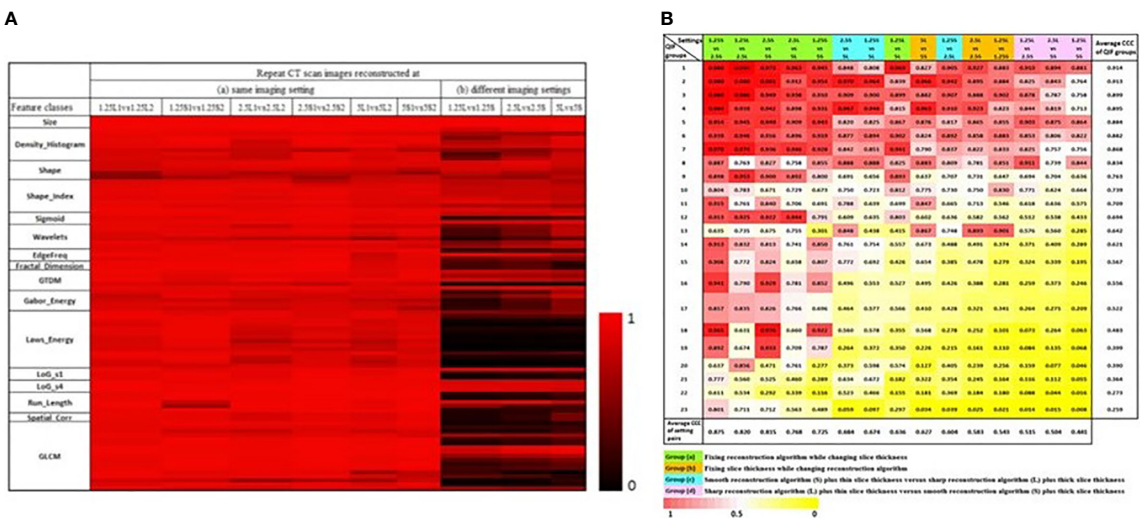


FIGURE 4 | **(A)** CCC heat map of radiomics features. The CCC (0 to 1) of the studied 89 radiomics features were computed from same-day repeat CT images reconstructed at (a) six identical imaging settings or (b) three different imaging settings. The brighter the red color, the higher the CCC values (i.e., the more reproducible) of a feature computed for the repeat scans (31). **(B)** CCC heat map of 23 non-redundant radiomics feature groups (rows) under 15 inter-setting comparisons (columns). Columns are arranged in descending order according to the average CCC of the inter-setting comparisons. Rows are arranged in descending order according to average CCCs of non-redundant feature groups (45).

leading to more uncertainty of target tumors and enlarged margins in the delineation of the treatment volumes. To decrease the amount of radiation exposures to healthy tissues, an emerging modality of gated or 4D CT imaging has been developed and used in radiation treatment planning (58). During a thoracic 4D CT study, multiple CT images are acquired over a period of at least one full respiratory cycle (8 or 10 phases) at each table position. Moving the table and synchronizing the

scans according to the patient's air-flow-volume curve, a spatial-temporal 4D CT dataset can be acquired. After being sorted, the motion-reduced 3D CT image series acquired at each respiratory cycle can be generated.

A recent study investigated respiration-related 4D stability of radiomics features across 8 individual respiratory phases for NSCLC (59). Eight hundred forty-one features were extracted from all individual phases of each patient. The relationship between individual coefficients of variation (COVs) and tumor motion magnitude was also inspected. The study found that some features (e.g., skewness, many GLDM features) were sensitive to respiration, whereas others (e.g., shape related features, many GLCM features) were not. The study did not observe a clear trend between the feature stability and the motion magnitude due to respiration. In the second part of the study, the value of utilizing 4D stability to preselect radiomics features to build prognostic prediction models for overall survival in early-stage NSCLC radiotherapy patients was explored. By comparing the performance of the models built with and without 4D stability feature preselection, the study showed an improved prediction performance with the preselection. Other studies in radiation oncology also suggested using phase images of already acquired 4D CT data as an alternative way to determine and remove unstable radiomics features prior to radiomics model construction when test-retest images were not available (51, 52, 60).

Radiomics Phantoms

Due to concerns such as radiation dose to patients, comprehensive investigations of image acquisition's effects on radiomics features have to rely on phantom studies. However, there is a significant disparity between tumor phenotypes that are seen in patient clinical CT images and traditional physical phantoms (e.g., simple shape, homogeneous density) (61).

Credence Cartridge Radiomics (CCR) Phantom

A group of medical physicists designed the CCR phantom to assist in exploring intra- and inter-scanner robustness and reproducibility of radiomics features (30). The CCR phantom embraces ten cartridges of an equal size of 10.1×10.1×3.2 cm^3, each filled with different materials in different patterns. The phantoms were scanned on 17 scanners from the four major CT vendors at multiple medical centers using their local thoracic imaging protocols. Both histogram-based and texture features were extracted using the open source radiomics software package of IBEX (62). The study results showed that the phantom's dynamic density range covered that observed in the tumors seen in 20 NSCLC patients. The authors noticed that inter-scan variability of the features varied depending on the feature itself and the cartridge material. One of the drawbacks of the CCR phantom is its uniform cartridge shape, which cannot study radiomics features that describe tumor shape and the interrelation between tumors and surrounding tissue.

Other studies also used the CCR phantom to explore the reproducibility and robustness of radiomics features across CT scanners, scanning techniques, and reconstruction parameters. An example was to study the effect of CT tube current on radiomics features. Using the ten cartridges in the CCR phantom, one study showed no clear effect of tube current on radiomics features (33). Another study, however, showed that tube current affected features extracted from homogeneous materials more than from tumor-like textured phantoms when splitting 6 cartridges contained in the CCR phantom into two groups, one filled with homogeneous materials and the other filled with more tissue-like texture materials (40).

3D Printed Phantoms

Although the CCR phantom has been widely used to investigate variability in radiomics features across scanners and scanning parameters, it does not contain lesion shape information, and its density textures/patterns are not anatomically informed. Recently, advances in 3D printing technology have made it possible to design and fabricate synthetic phantoms with realistic lesion sizes, shapes, intensities and internal textures while knowing the ground truth of their characteristics.

Using a subset of lung nodules taken from the database of Lung Image Database Consortium (LIDC), a series of corresponding virtual nodule models were created using the investigators' software and its built-in fitting and texture modeling routines (63). A multi-material 3D printer then distributed 2 base materials in the desired proportions according to the dithered nodule model to achieve lesion sizes, shapes, and internal density textures similar to those of the real nodules. The heterogeneous nodule phantoms were imbedded in an anthropomorphic thoracic phantom and scanned using different acquisition parameters of dose level, slice thickness, and reconstruction kernel. The study demonstrated that the printed textured phantoms can be used to determine the variability and accuracy of texture features extracted from CT images acquired at varying imaging settings.

In order to determine robust shape features, researchers used spherical harmonic functions to create mathematical tumor models with increasing degrees of complexity/spiculatedness and printed the models using a single material 3-D printer (64). They studied the relationship of a set of commonly used shape features (e.g., Volume, Surface area, Compactness, Sphericity) with varying degrees of spiculatedness under different conditions (slice thickness, resampling, and surface and volume computing algorithms). As expected, they found that surface-specific features, such as Surface area, were positively correlated with tumor spiculatedness, whereas global shape features, such as Compactness, were negatively correlated with tumor spiculatedness. They also found that the shape features are less affected by the aforementioned variables and less dependent to tumor volume.

Efforts Made in Imaging Harmonization

Image acquisition settings can vary considerably in datasets collected from retrospective or ongoing multi-center studies. Radiomics signatures that are influenced by variations in the source imaging settings may assign significance to differences such as an imaging parameter used to reconstruct images, rather

than the biologically significant differences in tumor images. Establishing the consistency of image data is vitally important for the discovery of robust imaging biomarkers which can be validated and applied to multi-center clinical trials and clinical practice. Harmonization of imaging protocols is an effective approach to reduce imaging-induced variability in radiomics.

Identifying Comparable Imaging Parameters

Different imaging settings can be said to be comparable when similar feature values can be computed from the images they produce. An early effort to identify comparable imaging settings was reported by the team who had contributed the RIDER Lung dataset. As a subsequent analysis of the same-day repeat CT study, the investigators used the six-setting CT image data to assess the feature agreements across the 3 slice thicknesses and 2 reconstruction kernels (45). Three inter-setting comparisons, i.e., 1.25S vs 2.5S, 1.25L vs 2.5L and 2.5S vs 5S, show high average CCC values (> 0.8 for all feature groups; bottom row in **Figure 4B**), indicating that these imaging parameters can be used interchangeably in radiomics studies. The study also found that changing slice thickness alone can generate better agreements, especially when the range of slice thickness is limited to 1.25mm and 2.5mm. Furthermore, combining thicker slices with sharper reconstruction algorithms can have the same effects as combining thinner slices with smoother reconstruction algorithms for the computation of radiomics features.

Controlling Imaging Protocols

The team who developed the CCR phantom studied whether a controlled imaging protocol could reduce variability in radiomics features (35) by scanning an updated version of the CCR phantom on 100 scanners using both local and study-specified CT protocols for chest and head & neck (H&N). The local imaging protocols were heterogeneous, e.g., the slice thickness ranged from 1 mm to 5 mm, while the study-specified protocols were controlled by using comparable imaging parameters across scanners, e.g., the reconstruction used slice thicknesses of 2.5 mm or 3 mm and smooth kernels. The size of cylindrical ROIs was 8.2 cm in diameter. The IBEX radiomics package was used to calculate 49 features including Neighborhood Grey Tone Difference Matrix (NGTDM) and Grey Level Co-Occurrence Matrix (GLCM). A linear mixed effects model was used to determine the overall variability contributed by the manufacturer, scanner of a given manufacturer, cartridge material, and residual to the variability in the measurements. The authors found that, compared to the local chest and H&N imaging protocols, the controlled protocols could reduce the overall variability by 57% and 52%, respectively.

Optimal standardization of chest imaging protocol parameters did not ensure the reproducibility of 27 texture features from the NGTDM and GLCM families, which were also computed using the IBEX radiomics package, across three CT vendor scanners (34) in a study using an anthropomorphic lung phantom with inserted lesions of different materials that simulated the attenuation properties of a human tissue. The imaging parameters were optimally chosen for lung cancer studies except the reconstruction slice thickness of 5 mm, which was rather thick for the small phantom lesions that ranged from 1 cm - 1.5 cm. One limitation of this study, as discussed by the authors, was the small size of the lesion inserts. The authors planned to conduct a follow up study to investigate the impact of ROI size on feature reproducibility, as calculating texture features such as GLCM from relatively small lesions on thick slice thickness can be problematic.

It is hard to make a direct comparison between the findings of these two phantom studies exploring benefits of the imaging parameter harmonization across CT scanners due to the differences of the phantoms, image preprocessing, etc.

Converting Imaging Settings to Desired Setting

Artificial intelligence (AI) offers the potential to automatically harmonize images which were acquired and reconstructed at different imaging settings. A recent study reported the use of a CNN to improve the reproducibility of radiomics features between different reconstruction kernels (soft and sharp) (65). The investigators developed a CNN architecture using residual learning and an end-to-end approach. To demonstrate the effectiveness of this CNN model, a total of 702 radiomics features were extracted from 104 pulmonary nodules or masses (all >= 6 mm; 51 non-enhanced and 53 enhanced CTs) using Pyradiomics (66), an open-source feature extraction package. The CCCs of the total features extracted from images reconstructed at the different kernels and the different kernels after image conversion were 0.38 and 0.84, respectively. Among the features, the CCCs of the wavelet features increased the most after the image conversion of the reconstruction kernels. The authors concluded that CNN-based CT image conversion can reduce the effect of reconstruction kernels on radiomics features. Another study showed that CNN-based super-resolution methods can improve the reproducibility of radiomics features extracted from CT images reconstructed at different slice thicknesses (67).

Matching Image Appearance/Quality

Differences in image quality between special vendors' CT systems are unavoidable. In addition to the scanner equipment, tube voltage and current, FOV, slice thickness, and reconstruction kernels, there are many other acquisition-related "hidden" factors that may affect image quality. It is impossible to study all affecting factors, known or unknown, one by one.

An alternative way to reduce feature variability caused by imaging is to identify the similarity of images acquired at different settings. Phantom studies can help match image appearance and thus identify comparable imaging settings across different vendors' scanners, scanning techniques and parameters, etc. (68, 69). For example, by analyzing noise power spectrum (NPS), a group of medical physicists conducted a study using the ACR CT phantom to quantitatively compare noise texture between two CT systems, GE and Siemens (68). Under a consistent acquisition protocol (120 kVp, 0.625/0.6 mm slice thickness, 250 mAs, and 22 cm field of view), using filtered back projection and a wide selection of available reconstruction kernels, a systematic kernel-by-kernel

comparison was performed. The study found that the GE's "Soft," "Standard," "Chest," and "Lung" kernels closely matched the Siemens' "B35f," "B43f," "B41f," and "B80f" kernels, respectively. More research in matching image quality can be found in (69).

Identifying Images Acquired at Optimal Phase Timing

Multi-phase CT scans after contrast administration are most widely used for liver cancer diagnosis, prognosis, and response assessment. Bolus tracking is used clinically during image acquisition to control (arterial and portal venous) phase timing to increase the likelihood of the phase timing being optimal. However, bolus tracking does not consider individual patient's biological variation and thus cannot ensure that the optimal timing was successfully reached in a given patient. In a pilot study, investigators explored the effect of portal venous phase (PVP) timing on the density measurement of liver metastases (LM) from colorectal cancer (CRC) and found that LM-CRC density was significantly decreased at non-optimal PVP timing by 14.8%: 16.7% at early PVP and 12.6% at late PVP (49). The same group then developed both semi-automated and AI-based fully-automated programs to identify optimal from non-optimal PVP timing as well as to differentiate five contrast-enhancement phases (49, 70, 71). They applied the developed PVP optimal-timing quality assurance (QA) method to their study developing an on-treatment signature to detect metastatic CRC patients sensitive to FOLFIRI+cetuximab using radiomics analysis of tumor changes between baseline and 8-week CT images. The radiomics signature showed higher performance on optimal imaging (AUC=0.80; 95%CI:0.69, 0.94) than on non-optimal imaging (AUC=0.72; 95%CI:0.59, 0.83) (72).

The effect of optimal timing on radiomics features is an understudied area. Automated AI-based QA algorithms to identify optimally acquired CT scan images for radiomics analyses can help ensure image quality and consistency and thus increase the chances to develop reproducible and reliable radiomics signatures.

Influence of Imaging Harmonization and Optimization on Radiomics Models

Imaging harmonization has shown potential for improving the reproducibility of radiomics features. The following subsections review and discuss how the performance of predictive models built using radiomics features is affected by the harmonization and optimization of image acquisition parameters.

Diagnosis of Solitary Pulmonary Nodule (SPN)

In a study using radiomics signatures to help the diagnosis of SPN, investigators assessed the effects of contrast enhancement, slice thickness, and reconstruction kernel on the diagnostic performance of the model they developed (73). In total, 240 SPN patients (malignant:benign = 180:60) had both non-contrast CT (NECT) and contrast-enhanced CT (CECT) scans, each reconstructed using two different slice thicknesses of 1.25 mm and 5 mm and two reconstruction kernels of lung (sharp kernel) and standard (smooth kernel). At each CT imaging setting, 150 radiomics features were extracted from each SPN and the diagnostic performance of the resulting signature was assessed based on its AUC. The validation results showed better discrimination capability of the radiomics signature derived from NECT than CECT (AUC: 0.750 vs. 0.735, p=0.014), from thin-slice than thick-slice CT (AUC: 0.750 vs. 0.725, p = 0.025), and from smooth kernel than sharp kernel (AUC: 0.725 vs.0.686, p = 0.039). The authors thus concluded that the non-contrast, thin-slice (1.25mm) and smooth reconstruction kernel-based CT was more informative for SPN diagnosis compared to the other imaging parameters studied.

Prediction of EGFR Mutation Status in Lung Adenocarcinoma (LAC)

Investigators evaluated whether the optimal selection of CT reconstruction settings improved the construction of a radiomics model to predict EGFR mutation status in LAC using standard of care CT images (74). In this study, CT scans of 51 patients (EGFR : WT = 23:28) with LACs of clinical stage I/II/IIIA were reconstructed at the following four image setting groups: 1) Thin-Sharp, 2) Thin-Smooth, 3) Thick-Sharp, and 4) Thick-Smooth (Thin: 1 mm; Thick: 5 mm; Sharp: B70f/B70s; Smooth: B30f/B31f/B31s). In total, 1,160 radiomics features were extracted and used to build machine learning prediction models at each of the four settings and a mixture setting (cases randomly selected from the groups 1-4). The study showed the best AUC (95%CI) of 0.83 (0.68, 0.92) when using the Thin-Smooth setting and the worst AUC (95%CI) of 0.75 (0.59, 0.86) when using the mixture setting ($P<10^{-3}$).

Prediction of Overall Survival (OS) in Head and Neck Cancer

A recent radiomics study in head and neck cancer found that models built with patients on a controlled imaging protocol did not predict OS better than models built using varying imaging protocols (75). In this study, investigators retrospectively collected 726 patients' CT images from one U.S. and two European institutions, among which the largest subset of 511 patients' CT images was acquired using a GE scanner with the reconstruction parameters of a standard kernel and 1.25 mm image thickness. The radiomics features were computed using IBEX (62). Radiomics models to predict OS were built using the full patient dataset (heterogeneous imaging protocols) and the largest subset (controlled imaging protocol). This study did not find increased performance of the outcome prediction model when the imaging protocol was controlled (AUCs: full set vs. subset = 0.72 vs. 0.55). Moreover, volume and HPV status were selected as covariates in the OS prediction model built on the full patient dataset. The authors further reported that volume alone or volume and HPV status provided an AUC of 0.73, indicating that adding radiomics features did not improve the model performance. This again suggests that radiomics texture features can be a surrogate for/correlated with tumor volume and points to the need to remove redundant features prior to model building (3, 76).

Discussion: Potential Strategies and Practical Considerations to Reduce Variability in Image Acquisition

Image acquisition is the first essential step in the radiomics workflow, directly determining the quality of images upon which all subsequent analyses rely. Some strategies and considerations to improve image consistency and reduce feature variability are highlighted below.

Controlling Imaging Protocols To Increase Image Consistency. CT scanners, scanning techniques, and reconstruction parameters can affect radiomics features and models. The degree of variation caused by these factors depends on the tumor's characteristics and the radiomics feature itself. Studies should report imaging acquisition settings in detail so that they can be reproduced by others. Ensuring high quality and consistent images across scanner devices and imaging protocols is the key for the successful development and application of radiomics signatures.

Potential Optimal Imaging Parameters For Studying Lung Cancer Phenotypes. Controlling CT imaging protocols and complying with these protocols are essential to the acquisition of high quality and consistent image data for radiomics studies. Preliminary data suggest that the most suitable imaging parameter setting for phenotype studies in lung cancer is thin slice thicknesses (e.g., 1 mm, 1.25 mm) and smooth reconstruction kernels (e.g., standard, B31f/B31s). Moreover, same-day repeat CT studies found that the settings of 1.25S and 2.5S generated the most reproducible features (**Figures 4A, B**). These independent findings support the use of thinner slice thickness and smoother kernel for prospective lung cancer phenotype studies. However, this approach warrants further investigation, especially because of the conflicting findings in the H&N study.

Test-Retest With Proper Imaging Parameters. The purpose of test-retest is to identify radiomics features that are sensitive to re-imaging and remove them from subsequent analyses. Because image acquisition parameters can affect computed feature values, the imaging parameters should be matched, or adequately similar, between the test-retest data and the individual studies' data so that the testing results are reliable. In addition, different disease sites should have their own test-retest image data, which can be acquired from either patients or (texture) phantoms. Before re-shooting a phantom, make sure to relocate/re-orient the phantom. When test-retest imaging is not available for the phenotype of interest, image perturbation such as noise addition, image translation and rotation, and volume growth or shrinkage can be considered (77).

4D CT - An Alternative For Test-Retest. Scanning patients twice during a short time period is impractical. However, 4D CT imaging has been used in radiotherapy to reduce respiratory motion-induced changes in tumor location and morphology. Such image datasets are often available in radiation oncology departments. Due to its ability to generate 3D CT image series at multiple respiratory phases, the 4D CT scan images can serve as a candidate of test-retest dataset to investigate feature variability. Studies show that certain radiomics features are sensitive to respiration and the preselection of 4D stability features can improve the performance of radiomics prediction models. Moreover, the end of the exhale phase, which is less affected by respiratory motion compared to the other phase images, is recommended to reduce feature variability for radiomics studies.

Radiomics Phantoms. Phantom studies play an essential role in exploring different sources of variation and their magnitudes across vendor scanners, scanning techniques and reconstruction parameters. However, traditional physical phantoms are usually constructed of materials that are radiologically equivalent to tissues and contain simple geometric features such as cartridges, cylinders, line-pair patterns, and ramps. Anthropomorphic phantoms typically mimic the overall shape of a human being but don't include detailed intra-organ/lesion features and are mostly used for dosimetry measurements. Thus, there is a significant gap between the intricate anatomical details that are seen in clinical CT images and the mostly uniform and simple nature of traditional physical phantoms. Characterizing such synthetic lesions using the cutting-edge 3D printing technology would be instrumental toward assessing the variability of features across different CT platforms and protocols.

Quantitative Metrics To Determine Image Quality and/or Similarity. The wide range of vendors' scanners, scanning techniques, and reconstruction parameters used in clinical practice and clinical trials makes it impossible to study the effects of all possible variables on radiomics features and models. Developing quantitative methods/metrics to determine image quality and/or similarity can be an alternative way to identify comparable images that can be used interchangeably or to decide whether an image's quality is adequate for computing radiomics features. This should be done based on the acquired images such as identifying optimal phase-timing, with no need to know the exact acquisition parameters of the images.

Imaging Harmonization Through AI/CNN. Image processing methods can reduce variability in images acquired with heterogeneous image acquisition settings. Voxel size resampling followed by Butterworth smoothing (an image processing method) has been found to improve feature reproducibility (42). Traditional image processing methods cannot be automatically adapted to harmonize a multitude of imaging settings that could exist in an image dataset. AI/CNN, however, shows great promise in converting CT imaging settings to a desired setting and in identifying whether images are acquired at the optimal phase timing. There is no doubt that AI, especially generative adversarial network (GAN)-based networks, will play a significant role in image-to-image translation including CT imaging conversion/harmonization (78).

Reproducible Features vs. Clinically Informative Features. When investigators report the improved reproducibility of radiomics features, a common method is to count the increased number of the studied features that have increased CCC values or CCC values greater than a predefined threshold (e.g., CCC >0.85). It is true that when more features are reproducible, there is a greater likelihood to identify robust radiomics models which are built using these reproducible features. However, the reproducibility of a feature does not

necessarily mean that it is clinically informative. On the other hand, it is likely that heterogeneous imaging settings may have little effects on some coarse clinically informative radiomics signatures such as ground glass opacity (GGO) portion and necrosis component of a tumor. Nevertheless, successful radiomics models must be built upon reproducible and robust features.

LESION SEGMENTATION

Lesion segmentation is a prerequisite for feature extraction, a critical step in radiomics workflow. Segmentation is an essential part of computer vision and image processing and is still an active research area today. Artificial intelligence (AI) promises fully-automated detection and segmentation of lesions (79).

Segmentation Methods and Variability

Segmentation can be performed manually, semi-automatically, or fully-automatically. Variability of the lesion segmentation may come from diverse segmentation algorithms and human supervised post-segmentation correction.

Manual Segmentation

Manual segmentation, a hand-drawing method using a computer mouse, is used only when there is no access to reliable semi-automated segmentation software because it is time consuming, subjective, and prone to variability due to radiologists' different opinions on identifying lesion boundaries (inter-reader variability) or a radiologist's inconsistency in delineating lesion boundaries at different time points (intra-reader variability). Manual segmentation was still used in about 40% of the lung cancer radiomics studies in our literature search (**Table 1** in **Supplemental Materials**).

Semi-Automated Segmentation

Semi-automated segmentation requires an operator to use a computer mouse to manually initiate a segmentation algorithm that can be developed using different strategies such as clustering, region-growing, active contours, and watershed transform.

Inter-Algorithm Variability

Different strategies employed in different segmentation algorithms can yield different results (inter-algorithm variability). In a "moist run" dataset (40 lung lesions and 12 lung phantom lesions) collected by the Quantitative Imaging Network (QIN) for a lung segmentation challenge, large variations were seen when three different segmentation algorithms were applied to the same GGO lung lesion (**Figure 5A**) (80). Briefly, the algorithm Alg01 was based on the marker-controlled watershed transform and required a region-of-interest (ROI) manually drawn outside the lesion as the algorithm's initial input (**Figure 5A**, top-left). Alg02 and Alg03 used the region-growing approach, with either one or multiple clicks to determine seed points (Alg02) (**Figure 5A**, top-middle) or a seed circle (Alg03) (**Figure 5A**, top-right) inside the lesion as the initial input. For heterogeneous lesions, the region-growing based algorithms can easily be trapped by a local homogeneous region, creating a high risk of under-segmentation. In this example, Alg02 segmented only the solid part of the lesion (under-segmentation) when the seed point was placed in a high-density area of the lesion.

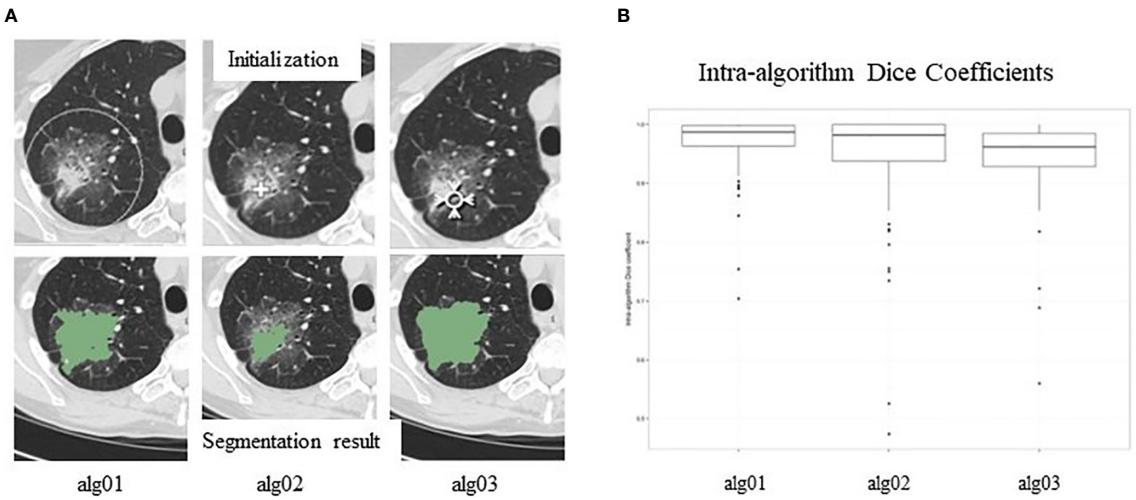


FIGURE 5 | **(A)**. Inter-reader variability in segmentation. Top panel: manual initializations (seed point/ROIs) of three segmentation algorithms; bottom panel: corresponding segmentation results. **(B)** Intra-reader variability in segmentation. Segmentation results are affected by seed points/ROIs. Reproduced with permission from (80).

Intra-Algorithm Variability

After manual initialization, each algorithm analyzes the density distribution of the pixels provided by the initial ROI and then automatically separates the lesion from its background using the input information and its own segmentation strategy/objective criteria. Therefore, initial ROIs can affect segmentation results. **Figure 5B** shows variations (Dice Coefficient distribution) of the segmentation results (lesion volumes) of each of the three algorithms when initial ROIs are placed differently. This kind of variation is called intra-algorithm variability. A good segmentation algorithm should be insensitive to initial ROIs. Studies reported that radiomics features extracted from segmented lesions had higher reproducibility when using the same algorithm with different initial inputs than when using different segmentation algorithms (81, 82).

Fully-Automated Segmentation

Fully-automated segmentation is performed without any human-machine interaction. The input of such algorithms is the entire image series and the output is the image series containing automatically segmented lesions. A fully-automated segmentation method needs to perform lesion detection and segmentation simultaneously. The challenge for automated lesion detection is to avoid false negative and false positive lesions. Unlike manual and semi-automated segmentations, repeatedly running a fully-automated algorithm on one image series won't change the output result. However, the impact of image acquisition settings on fully-automated segmentation algorithms needs to be explored (83).

Human Supervised Post-Segmentation Correction

Ideally, a lesion segmentation algorithm should be fully automated, reproducible, and accurate. However, both lesions and relationships between lesions and their surrounding tissues can manifest in complex patterns on CT, making a satisfactory segmentation for all lesions unrealistic. To avoid segmentation errors, a radiologist needs to review and correct computer-generated lesion contours. Over the past few years, awareness of the need for human supervised post-segmentation correction has increased (**Figure 2B**; green color). Supervised segmentation is influenced by the radiologist's subjective judgement. However, only the modified parts of the lesion contours are affected by the manual correction and the unmodified contour parts are still determined by objective criteria. This explains why radiomics features extracted from lesions segmented manually were less reproducible than those extracted from lesions segmented algorithmically with supervision by a radiologist (84, 85).

Segmentation of Multiple Disease Sites

Solid tumors, including primary and metastatic lesions, exist in various organs. They can present various intra-tumoral patterns and contrast levels to the surrounding tissues on CT images, which challenges lesion segmentation to different degrees. For instance, lung lesions are usually easy to be segmented due to their high contrast to the surrounding lung parenchyma. However, when lung lesions attach to blood vessels or chest walls possessing similar densities to those of the lesions, segmentations can become difficult. Lymph nodes are well-known for their low contrast to their surrounding backgrounds. Segmentation of liver lesions can suffer from their heterogeneity, low contrast against liver parenchyma (contrast-enhancement dependent), and noisy abdominal images. Various strategies have been developed to better delineate tumors of different types.

In general, texture features are affected more than volume feature by image acquisition parameters. Over-segmentation, i.e., inclusion of surrounding non-lesion tissues in the lesion segmentation, can have a large effect on texture features when there is a large density difference between the lesion and its surrounding tissues (e.g., lung lesion and lung parenchyma). A tight segmentation is thus more desired than a loose segmentation in radiomics studies. Lesion segmentation can hit lesion boundary-related features harder than others.

A study preliminarily analyzed the effect of inter-observer variability between three manual contours on the stability of 1,404 radiomics features in head and neck squamous cell carcinoma (HNSCC), malignant pleural mesothelioma (MPM), and NSCLC (86). There were 11 lesions for each type. The authors found that the inter-observer delineation variability was the highest in MPM and the lowest in NSCLC, and the stability rate of radiomics features negatively correlated with delineation variability. Shape-related features showed the weakest stability among the 3 tumor types.

Effect of Inter-Reader Variability on Radiomics Prediction Model

The last example in this section shows a pilot study exploring the effects of inter-reader variability on radiomics prediction models. In the study, the investigators predicted EGFR mutational status in early stage NSCLC patients treated with a targeted therapy (Gefitinib) using the change in 89 radiomics features over 3 weeks (delta features) extracted from 1.25 mm and lung kernel images (87). Lung lesions in 46 patients (EGFR:wildtype = 20:26) were independently segmented by three radiologists using in-house software that allowed manual post-segmentation correction. Univariate analysis identified the most significant delta features computed from each of the three radiologists' segmentation results. The best EGFR prediction performance expressed by AUC values differed for each radiologist's segmentation: 0.79 (top feature: compact factor – a shape feature), 0.85 (top feature: mean density) and 0.91 (top feature: volume), respectively. Delta volume was the only feature that was among the top 5 most significant features in all three radiologists' results. The prediction performances using the delta volumes obtained by the three radiologists were (AUC=) 0.77, 0.80 and 0.91, respectively. All outperformed the corresponding unidimensional performances of 0.63, 0.53 and 0.66. Unidimensional measurement (i.e., tumor in-plane diameter) is used to assess tumor change by conventional Response Evaluation Criteria in Solid Tumors (RECIST) (88). None of the three radiologists' results included the delta diameter in its

top 5 most significant feature list. The results of this study warrant validation on larger data.

Open Source Software for Lesion Segmentation

3D Slicer and ITK-Snap are the most popular open source platforms for interactive segmentation, registration, and volume rendering/visualization of medical images. Built over two decades through support from the National Institutes of Health (NIH) and software engineers worldwide, 3D Slicer has provided researchers with a set of free image processing tools (89). ITK-SNAP is another open source tool that offers free semi-automatic segmentation software (90). Both platforms provide manual delineation functions. So far, about 25% of lung cancer radiomics studies were conducted with the help of open source segmentation tools (**Table 1** in **Supplemental Materials**).

Discussion: Potential Strategies and Practical Considerations to Reduce Variability in Lesion Segmentation

Accurate, reproducible, and efficient segmentation tools that can be widely distributed are essential to accelerating and advancing cancer imaging research. Semi-automated segmentation tools have commonly been used in radiology-oncology imaging studies. An imaging platform providing lesion segmentation software should also provide a manual editing/correction function. Certainly, computer segmentation methods are more efficient when they require fewer human-machine interactions.

Inter- and/or Intra-Reader Test. The Purpose of Inter-Reader (or intra-reader) testing is to recognize radiomics features that are sensitive to lesion segmentation so that they can be removed from subsequent analyses. Features that are sensitive to segmentation can be identified by asking multiple radiologists to delineate the same lesions or an individual radiologist to delineate a set of lesions at two or more sessions, with a sufficient time interval between any two annotation sessions to avoid the effects of the radiologist's reading memory.

Radiologists' Consensus on Lesion Contouring. Radiologists are not specifically trained in identifying tumor boundaries; big variations can happen especially when segmenting partial solid tumors. The Tumor Segmentation step shown in **Figure 1** (Radiomics workflow) offers an example of three radiologists' manually delineated tumor contours; some tend to delineate contours tightly surrounding a solid tumor component, while others tend to delineate contours loosely including more GGO areas. Although there may not be "gold standard" lesion boundaries, obtaining radiologists' consensus about lesion boundaries can help reduce variability in segmentation and thus in computed radiomics features.

Proper Use of Segmentation Software. Different semi-automated algorithms use different strategies to obtain information about lesions and/or their surrounding tissue from initial ROIs, which can help the algorithms identify lesion pixels/voxels. For instance, to properly start a region-growing based algorithm, seed ROI points/circle should be placed in both hypo and hyper density areas inside a heterogeneous lesion so that the range of lesion densities can be fully captured and used to guide the region growing algorithm. Proper use of a segmentation algorithm can improve the segmentation's accuracy and consistency. Again, to avoid unpredictable surrounding tissues of possible high (or low) contrasts, tight segmentation results are more preferred than loose segmentation results in radiomics studies.

In radiation oncology, the standard treatment planning process has generated a large amount of annotated tumors that can be readily used in radiomics studies. However, it should be noted that the quality of the segmentations might not be sufficiently precise for radiomics. For instance, there is no need to accurately delineate the speculated edges of a tumor for the purposes of treatment planning while radiomics requires a very precise delineation of the tumor. Therefore, segmentation results taken from radiotherapy data may need to be refined prior to feature extraction.

Effects of Imaging on Segmentation. Acquisition settings determine image quality and can thus affect segmentation algorithms [e.g., (91, 92)]. The ultimate goal of image pre-processing is to reduce noise while maintaining image details. Generally, pre-processing methods using smoothing filters (e.g., Gaussian filter) are applied for the region-growing based algorithms, whereas sharpening filters (e.g., Laplacian filter) are used by the edge-based segmentation algorithms. When investigating volumetric imaging biomarkers, variables affecting volumetry/tumor segmentation have been intensively studied, particularly by the RSNA-organized Quantitative Imaging Biomarkers Alliance (QIBA) (93–95), which is not further discussed in this review.

FEATURE EXTRACTION

Radiomics features are also known as quantitative image features. In the past decades, pattern recognition using quantitative image features has been widely used for tasks such as image segmentation, classification, and computer-aided detection and diagnosis (96).

Radiomics Features

Radiomics features can be grouped into two categories: agnostic and quantified semantic features (18). Agnostic features are derived to quantify lesion morphology and density heterogeneity through mathematical equations/descriptors, while quantified semantic features are developed to characterize visual patterns of lesions (ROIs) based on radiology lexicons. Agnostic features are usually further divided into the following four categories based on: 1) morphology (e.g., size, shape), 2) histogram-statistics (e.g., mean, standard deviation, skewness, kurtosis), 3) texture (e.g., Run-Length, GLCM), and 4) transformation (e.g., Wavelet transform). Histogram-based features, a.k.a. first-order statistics, describe tumor density distribution without considering spatial information, whereas texture features, a.k.a. second-order statistics, characterize tumor heterogeneity by considering the spatial interrelations of image pixel/voxel densities. Transformation-based features are

computed from transformed images rather than original images. Of note, there is another type of quantitative features that can provide additional information, i.e., features that characterize density transition between a lesion and its surrounding tissues/parenchyma. An example is the feature class of Sigmoid Function; the feature, Sigmoid-slope, can be used to quantify lesion edge (density) sharpness.

Quantified semantic features are perceptive because they are created based on a radiologist's visual observations. For instance, GGO volume percentage, a quantified visual feature, was found to be significantly higher in tumors with exon 21 missense mutation than that in tumors with other EGFR mutation status (97). Some agnostic features can also be intuitively interpreted. Skewness, an agnostic feature that measures the asymmetry of a density distribution about its mean (e.g., density distribution of a solid tumor is left-skewed with a negative skewness value), was found to be predictive for disease-free-survival (DFS) associated with certain histologic subgroups of lung adenocarcinoma; the lower the skewness value is, the poorer the DFS will be (98). Another example is the Laws' Energy features. This feature class emphasizes texture patterns of edge, spot, ripple and wave through the Laws filters. Whether such tumor image patterns are clinically informative needs to be investigated. However, meanings of many agnostic features can be hard to be intuitively interpreted. Nevertheless, it is believed that models built upon one or multiple radiomics features can distinguish imaging phenotypes that can or cannot be visually observed by human.

Sources of Variation in Feature Computation

Traditional radiomics features are computed from predefined mathematical equations/descriptors that can be found in textbooks and/or published literature (99). Theoretically, these radiomics features are clearly defined and thus fully controllable. However, sometimes there are multiple choices to define a feature with an identical name, select specific values for feature parameters, and implement a feature calculation. In reality, values of radiomics features computed using different feature extraction software can vary considerably, which makes it hard to compare radiomics studies especially if details of the feature definitions, parameter settings, and implementations are not disclosed adequately.

Feature Definition

Variations in feature definition can happen when multiple equations/descriptors are used to define a same feature. A simple example is Compactness, a shape feature that is defined to quantify how spherical a 3D object's shape is. Although Compactness is a function of an object's surface area (S) and volume (V), there are different equations to define it, e.g., $V/(\pi^{1/2} * S^{3/2})$ and $36*\pi* V^2/S^3$. Even if these two equations are related, the computed values from the two equations are different. This type of variance can be controlled by making feature definitions transparent.

Feature Parameter Setting

Many features, especially texture features, have parameters in their definitions so that they can be used to quantify image patterns at multiple scales and different orientations. The feature parameter used most often is the number of gray-level (density) bins, a.k.a. the bin level. Density discretization groups the entire density range of images into bins of equal width. Reducing the bin level or increasing the bin width can improve the computational efficiency for certain features such as GLCM features. Moreover, density discretization can lessen noise interference. In general, bin width should not be lower than random noise level. However, a large bin width may not be able to capture the subtle differences in density (texture).

The GLCM feature class is an excellent example to explain feature specific parameters (**Figure 6**) (100). A GLCM matrix is created by counting how often pairs of pixels with specific gray-level values occur in a specified distance and direction over the ROI. The GLCM features are the computed statistics from the matrix (101). There are 3 key parameters: the bin level of the original images (i.e., the dimension of GLCM), the distance of pixel pairs, and the direction of the line spanned by the pixel pair. In **Figure 6A**, starting with an original image, the figures show the feature computation process. Two example GLCM matrices are generated with the bin levels of 4 [**Figures 6A(b)–(d)**] and 8 [**Figures 6A(e)–(g)**], both with the distance of 1 pixel and direction of 0° (**Figure 6B**). For each GLCM matrix, two common GLCM features, Contrast and Homogeneity, are calculated and their values are different due to the different bin levels, **Figures 6A(d)** and **(g)**.

The influence of density discretization (bin levels) on radiomics features was investigated using the CCR phantom (33). The effect of the bin width (5 to 50 HU) on the stability of 114 studied texture features was found to be marginal compared to the effect of scanners, slice thicknesses, and tube currents. Although the study concludes that feature stability may not be compromised during the optimization of gray-level discretization when attempting to improve model performance, evidence from clinical studies is needed.

Feature Implementation

Often, there are multiple choices to implement certain radiomics features. For instance, a lesion surface area can be evaluated by a mesh-based representation of the outer surface or by areas of voxel faces toward the outside of the lesion. For features that are derived from pre-processed images using a filtering technique such as Gabor filter, filter length is a feature parameter, and the method for handling the ROI edge when moving the filter over the ROI is a hidden variable in the implementation of the Gabor filter. Moreover, features can be computed in 2D, 2.5D (a combination of 2D features), or 3D and extracted from the original images as well as from pre-processed images using different filtering techniques. In 2D image processing, for instance, 4 or 8 connected pixels are usually considered as neighboring pixels and 4 or 8 directions are chosen. All these and more unspecified variances during feature computation/implementation can add unknown variation to the computed feature values. To date, no radiomics studies have provided sufficient details about their feature definitions, parameter settings and implementations so that others can reproduce this aspect of their studies.

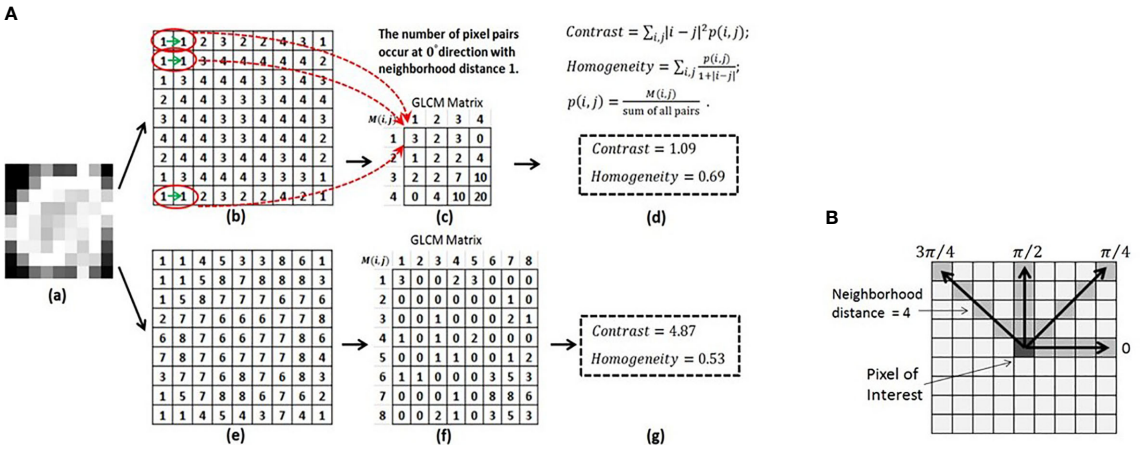


FIGURE 6 | **(A)** Computing the GLCM features of Contrast and Homogeneity using different bin levels. (a) Original image. (b) Normalized image using the bin level of 4. (c) GLCM matrix derived from (b). (d) Contrast and Homogeneity computed from GLCM in (c). (e) Normalized image using the bin level of 8. (f) GLCM matrix derived from (e). (g) Contrast and Homogeneity computed from GLCM in (f). **(B)** Calculation of the GLCM features for a 9X9 2D image at four directions and a neighborhood distance of 4 pixels. Reproduced with permission from (100).

Studies Exploring and Reducing Variability in Feature Computation

Research has investigated sources of variation in feature computation (81, 102–104). Two collaborative studies on this topic are reviewed and discussed in the following subsections.

Preliminary Effort by the Quantitative Imaging Network (QIN)

This study, conducted by ten teams from the PET/CT working group of the QIN funded by the National Cancer Institute (NCI), explored the agreement of 13 software packages on nine basic radiomics features including volume, 2D and 3D diameters, mean density, standard deviation, kurtosis, surface area, sphericity, and GLCM entropy (103). The investigators applied the feature extraction software used by the teams (about half open source and half in-house) to both Digital Reference Objects (DROs) and patient image data. The DROs consisted of three objects with both texture and uniform densities and spherical and spiculated shapes (105). The patient data contained images from 10 patients taken from the LIDC database, a publicly accessible database (106). One pre-annotated contour for each DRO/lesion was used to extract radiomics features. Percentage coefficient of variation (CV) was used to evaluate agreement of the computed features. The results showed that for the DROs, six out of the nine features, i.e., volume, 2D and 3D diameters, mean density, standard deviation and kurtosis (after Fisher correction), demonstrated excellent agreement (CV < 1%). The features of surface and sphericity showed moderate agreement (CV: ~13%). GLCM entropy had big variations (texture DRO: ~50%; uniform DRO: CV > 600%). For the patient data, CV values of 2D and 3D diameters, surface, and sphericity increased but were still moderate. CV of the GLCM entropy decreased to ~36%. All other features remained in excellent agreement.

From the DROs to real lesions, ROI shapes became more irregular and densities became less uniform. This was why the software packages turned out to agree less with each other when computing features that relied more on ROI boundaries/surfaces, such as 2D and 3D diameters and surface and sphericity features. It was not surprising that the GLCM entropy feature showed such big variations between feature extraction software packages. Harmonization of some key parameters (e.g., bin level, pixel pair's distance and direction) was found to reduce the average CV value of GLCM entropy from ~36% to ~20%.

Comprehensive Study by the Image Biomarker Standardization Initiative (IBSI)

Since 2016, the IBSI, an independent international collaboration, has focused on standardizing definition and implementation of quantitative image features and providing benchmark data sets and consensus-based reference values (26). The IBSI reference manual is written to provide consensus-based recommendations and guidelines to improve reproducibility and transparency of radiomics features and studies.

Recently, the IBSI published a large scale study that standardized 169 commonly used radiomics features (104). This multi-year, multi-phase study involved 25 research teams using their own feature extraction and image processing software and showed the investigators' first-hand experience in the calibration and certification of various feature extraction software packages. The study utilized a consensus-based and iterative approach. Phase I (25 participating teams) obtained the reference values of radiomics features based on a 3D digital phantom. Phase II (15 teams) defined a general image processing scheme, implemented it at different configurations, and obtained corresponding reference values of radiomics features using a lung cancer CT image series. Initially, only weak consensus (<3

teams matched) existed for 76.8% features at phase I and 65.4% features at phase II. At the final iteration, strong or better consensus (6-9 team matches) was achieved to 95.1% and 90.6% at phase I and II, respectively. Phase III (9 teams) prospectively assessed reproducibility of the 169 standardized features against a public dataset of CT, PET, and MR images from 51 sarcoma patients. More than 97% of the features studied reached an excellent reproducibility (ICC > 0.9), showing the value of feature standardization in reducing variability between different feature extraction software.

The study identified several causes of deviation. For instance, lesion volumes can be represented by simple voxel cubes or polygonal models (or meshes). This affects the computation of surface area and thus morphological features. Sometimes, there are "holes", which are dark regions inside segmented lesions. The decision whether to fill such small holes prior to feature computation can influence the computed value. Differences of this kind are controllable and can be reduced or eliminated through feature standardization.

Feature Distribution Harmonization – Combat

Image acquisition-induced variations in radiomics features are intensively discussed in the early section of Image acquisition, where the suggested solutions to reduce such variability are mainly focused on obtaining consistent and/or comparable images through controlling image acquisition protocols and/or post-processing of acquired images using both conventional and AI-based methods.

Recently, a new data-driven method based on the empirical Bayes frameworks, called ComBat harmonization, was introduced into radiomics to reduce feature variability caused by scanners and scanning parameters (107–109). This method was initially developed for large-scale genomic data analysis (110). When combining different datasets collected from microarray experiments, a big challenge is to remove non-biological variations caused by the systematic technical differences while handling samples, i.e., to remove the so-called "batch effects", where the batches denote operators, array types, etc. In radiomics, batches refer to scanners, imaging protocols, individual imaging parameters, etc. Unlike the imaging harmonization, the ComBat method operates directly on the computed feature values to remove batch-induced bias. This eliminates/reduces, for example, the demands for sharing and transferring medical images between institutions that can be limited by specific regulations and standardizing image acquisition settings that can be hard to be implemented in routine clinical practice.

Figure 7 shows two examples of harmonization/realignment of a feature, GLCM Homogeneity (108). The example shows two distributions of the feature, computed from the images reconstructed at two different reconstruction kernels, Lung vs. Standard (example 1, **Figure 7A**), and at two different slice thicknesses, 1.25 mm vs. 5 mm (example 2, **Figure 7B**). In each example, feature distributions were better overlapped after applying the ComBat harmonization function (https://github.com/Jfortin1/ComBatHarmonization). A follow-up study independently verified the published results by applying the ComBat method to harmonize a larger set of radiomics features computed from a broader range of imaging protocols in a larger cohort of patients. The investigators noticed that the harmonization also increased the repeatability of texture features (109). This promising technology warrants validation for its clinical usefulness in radiomics.

Experience With Open Source Software and Open Source Databases in Building Radiomics Prediction Models

There are a number of free open source software packages to compute radiomics features. Based on the literature searching results, open source and in-house feature software were used almost equally frequently in the lung cancer radiomics studies published from 2014 to July 2020 (**Supplemental Materials**; excluding ~9% "not specified" software). Pyradiomics (~15%) (66) and Imaging Biomarker Explorer (IBEX) (~8%) (62) are the two most popular open source software to study radiomics.

Recently, a new radiomic feature calculator, called RaCaT, became available (111). It calculates a large number of features that are in compliance with the IBSI standard. Although the calculator can be downloaded and used without requiring any programming skills, it does not provide any Graphical User Interface. Users need to call the calculator either from their own programming environments or from the command line.

A research group recently reported its first-hand experience in building a radiomics model to predict EGFR mutation status in NSCLC patients using two open source databases, TCIA (The Cancer Imaging Archive) (112) and TCGA (The Cancer Genome Atlas) (113), and three feature extraction software packages, the open source Pyradiomics (1319 features) and IBEX (1563 features), and an in-house package (1160 features) (114). Although they encountered some obstacles, they reported a smooth experience overall with the public datasets and open source feature extraction software. They were able to collect both image data and clinical data for the majority of patients satisfying the inclusion criteria of their study. However, the TCGA-LUAD and the TCGA-LUSC datasets contained image data and genomic data that were stored separately on the TCIA and the TCGA, respectively, for the majority of cases. In addition, the genomic data was often incomplete, which reduced the number of useable cases. The two open-source software packages had clear instructions that made them amenable to beginners. Radiomics feature definitions were well documented and were able to be extracted from the majority of lesions. Some errors did occur during the extraction in both open-source software packages that could not solved. The study found that although the three software packages selected different features to build their prediction models, the models' performances were similar. The correlations found between those selected features by the different software indicate that these features may describe

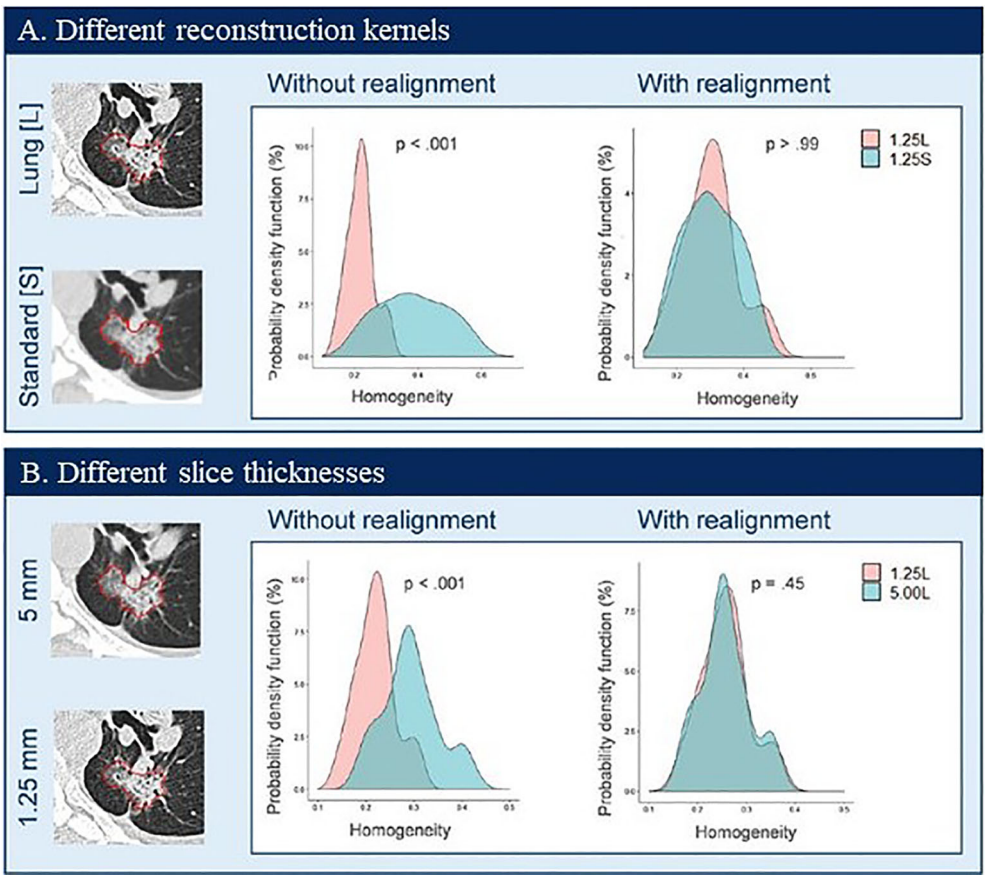


FIGURE 7 | Probability density distributions of Homogeneity before (without realignment) and after (with realignment) ComBat in patient data by using two CT reconstruction kernels **(A)** and two slice thicknesses **(B)**. Reproduced with permission from (108).

similar tumor imaging phenotypes that are associated with underlying biological characteristics.

Discussion: Potential Strategies and Practical Considerations to Improve Feature Extraction

Variations in feature computation are caused by possible differences in feature definition, parameter setting, and implementation. Variations also come from the previous steps of image acquisition, lesion segmentation, and image preprocessing, which exaggerate variability in radiomics features and models built using these features (12, 115, 116).

Feature Definition Standardization. One way to reduce feature variability, enhance collaboration, and accelerate the development and validation of radiomics signatures is to standardize feature definition, parameter setting, and implementation. The IBSI's effort in standardizing the feature extraction process is a significant step toward increasing feature transparency, reducing feature variation, and providing reference images and reference feature values to help verify/calibrate feature extraction software developed by researchers globally (104). Customizable 3D DROs can be created to help standardize radiomics features and uncover coding errors (105). It should be noted that promoting feature standardization does not mean that investigators shouldn't develop and use their own feature definitions, parameter settings, and implementation methods that are different than those suggested by the IBSI.

Feature Parameter Setting. Normally, we only use about 100 or less fundamental radiomics features. However, with multiple settings of feature-specific parameters, different implementation methods, and various image pre-processing methods, the total number of features that can be provided by a feature extraction software package can easily reach multiple thousands. Currently, the settings of many feature parameters are "randomly" chosen or simply adopted from the literature where the image types and contents can be very different than those of the investigators' own clinical studies. As a result, the same features, same feature parameter settings, and/or same image pre-processing methods are often used to study different clinical questions for different disease sites using different imaging modalities. This so called one-size-fits-all scenario may delay or prevent the discovery of radiomics signatures. In order to increase the opportunity to identify biologically relevant features while studying lung cancer,

for instance, understanding the density range of lung tumors and image noise characteristics may help choose proper values of feature parameters.

Feature Redundancy. On one hand, multiple parameter values allow quantification of lesion textures at different scales, contrasts, and directions, which can increase the chance to identify biologically relevant features. On the other hand, multiple parameters can drastically increase the total number of features, many of which are correlated. The high dimensionality of features can also lead to model overfitting. Reducing feature dimension is necessary prior to building prediction models using machine learning methods. Feature reduction and identification of potential confounding variables such as image acquisition parameters (e.g., slice thickness) and clinically used prognosticators (e.g., tumor size) are beyond the discussions of this review paper.

Feature Transparency. Inadequate descriptions of feature extraction in the current literature is a big burden for widespread adoption of the features and replication/validation of the developed radiomics signatures. For researchers who are capable of writing their own feature extraction algorithms, it is important for them to track their changes of the codes using version control software and describe the feature extraction details as much as possible in publications. For the groups offering open source feature extraction software, the software version numbers along with the release dates and upgrades should be clearly documented and provided for the purposes of record tracking.

Image Pre-Processing. Image pre-processing includes, but is not limited to, smoothing, sharpening, and/or resampling of images prior to feature extraction. Generally speaking, image smoothing can improve density-based feature reproducibility. For instance, the LoG (Laplacian of Gaussian) texture features computed from the same-day repeat CT scan images reconstructed at different imaging settings is an example (**Figure 4A**). LoG_s1 denotes no pre-processing and LoG_s4 indicates that a large Gaussian kernel is applied to strongly smooth the original images before the feature calculation. The reproducibility of LoG features calculated on the smoothed images is drastically improved (CCC heat map colors changed from dark to bright) even when the features are calculated from images reconstructed using different kernels. However, over smoothing can suppress image texture details, which may lose clinically useful information related to low contrast textures. There is a trade-off between reproducible features and informative features.

Another image pre-processing operation is to resample CT images to isotropic resolutions in x-, y-, and z-directions. Studies show that isotropic resolutions can improve feature reproducibility (42). It is worth mentioning that, in 3D image segmentation, the isotropic resampling of images is often a precondition for direct use of 3D image processing operators that are employed by many 3D segmentation algorithms.

Reproducible and Reliable Features. Both re-imaging and re-segmentation can introduce variation into radiomics features. To assess the reproducibility due to re-imaging, features are extracted from a set of lesions imaged and segmented from two repeat scans acquired within a short time interval. To assess reproducibility due to re-segmentation, features need to be extracted from a set of lesions segmented by the same radiologist in at least two different sessions (intra-reader variability) and/or by at least two independent radiologists (inter-reader variability). If repeat scan image data are available, re-segmentation of lesions on repeat images can take into the account the variability caused by both re-imaging and re-segmentation simultaneously. The concordance correlation coefficient (CCCs) is a widely accepted statistical method to assess the reproducibility of radiomics features (117). Only reproducible features will be retained for the subsequent machine learning analysis. Once features are extracted, checking outliers for each feature is a practical way to help identify imaging artifacts, segmentation errors, etc.

ComBat Feature Harmonization. The ComBat is an easy-to-use and fast feature harmonization method recently introduced to remove batch effects in radiomics. Based on calculated feature values, the ComBat method has the ability to adjust for the batch effects at multiple layers, e.g., at institution, scanner, imaging protocol and individual imaging parameter levels. With the ComBat method, more features can become robust and be analyzed, historical image data can be better reanalyzed and multi-center data can be properly combined and/or compared. Future research includes, for instance, incorporating clinical and biological variables into the ComBat method to preserve biological variation while maximally removing batch effects. The ComBat feature harmonization opens a new and efficient avenue to accelerate the development, validation and dissemination of robust and generalized radiomics signatures and their transfers to clinical practice.

CNN Features. Given sufficient data, features derived from a CNN can be expected to overcome the limitations of pre-defined traditional radiomics features because a CNN's backward propagation of errors for training purposes enables the network to self-learn novel features which are most useful for a specific application. The automated learning and iterative image filtering performed by a CNN may also make the CNN models less likely to be confounded by heterogeneous image acquisition settings. The CNN also eliminates the step of lesion segmentation, a major source of variation in radiomics. Nevertheless, radiomics can build tumor imaging phenotype models using small datasets, a necessity for many medical studies. Radiomics signatures can often be intuitively interpreted, which also makes radiomics favorable over the "black box" approach of using a CNN. In the foreseeable future, there is no doubt that both radiomics and AI/CNN will be mainstream approaches to study quantitative imaging biomarkers in precision medicine.

SUMMARY

Radiomics has shown promise for a variety of clinical applications in lung and other cancers, and in particular for

diagnosis, prognosis, and response assessment. Radiomics derives strength from hypothesis neutral techniques that can identify subtle details or changes in patterns/features of medical images that are associated with biological activities and clinical outcomes. This, however, also creates a potential weakness: the values of computed radiomics features and the performance of radiomics models incorporating them can be sensitive to many variables intrinsic to the radiomics workflow. Given heterogeneous image acquisition settings, varied quantification software packages, different diseases' characteristics, and small and mixed patient populations, the development of reproducible and generalizable radiomics signatures is not as straightforward as it initially appeared. Indeed, radiomics is a multidisciplinary research field. Its success relies on close collaborations among physicians, medical imaging physicists, biomedical engineers, statisticians, and computer scientists. Over the past years, great community efforts have been made to better understand sources of variation, improving reproducibility and reliability of radiomics features and models through imaging and feature harmonization and increasing transparency and quality of radiomics studies. Ever-growing open source imaging and genomic databases as well as open source software packages help accelerate the development and external validation of radiomics signatures.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The author would like to acknowledge the support from National Institute of Health (Award Number U01 CA225431).

ACKNOWLEDGMENTS

The author would like to thank Tavis Allison for assistance in editing of the manuscript and Jin Yoon for assistance in expanding and presenting of the data in **Supplementary Table**.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.633176/full#supplementary-material

REFERENCES

1. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
2. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* (2012) 30 (9):1234–48. doi: 10.1016/j.mri.2012.06.010
3. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* (2014) 5(1):1–9. doi: 10.1038/ncomms5644
4. Segal E, Sirlin CB, Ooi C, Adler AS, Gollub J, Chen X, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat Biotechnol* (2007) 25(6):675–80. doi: 10.1038/nbt1306
5. Banerjee S, Wang DS, Kim HJ, Sirlin CB, Chan MG, Korn RL, et al. A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. *Hepatology* (2015) 62(3):792–800. doi: 10.1002/hep.27877
6. Fornacon-Wood I, Faivre-Finn C, O'Connor JP, Price GJ. Radiomics as a personalized medicine tool in lung cancer: Separating the hope from the hype. *Lung Cancer* (2020) 146:197–208. doi: 10.1016/j.lungcan.2020.05.028
7. Song J, Yin Y, Wang H, Chang Z, Liu Z, Cui L. A review of original articles published in the emerging field of radiomics. *Eur J Radiol* (2020) 127:1–8. doi: 10.1016/j.ejrad.2020.108991
8. Rogers W, Thulasi Seetha S, Refaee TA, Lieverse RI, Granzier RW, Ibrahim A, et al. Radiomics: from qualitative to quantitative imaging. *Br J Radiol* (2020) 93(1108):20190948. doi: 10.1259/bjr.20190948
9. Dercle L, Henry T, Carré A, Paragios N, Deutsch E, Robert C. Reinventing Radiation Therapy with Machine Learning and Imaging Bio-markers (Radiomics): state-of-the-art, challenges and perspectives. *Methods* (2020) 188:44–60. doi: 10.1016/j.ymeth.2020.07.003
10. Papanikolaou N, Matos C, Koh DM. How to develop a meaningful radiomic signature for clinical use in oncologic patients. *Cancer Imaging* (2020) 20:1–10. doi: 10.1186/s40644-020-00311-4
11. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging—"how-to" guide and critical reflection. *Insights Imaging* (2020) 11(1):1–16. doi: 10.1186/s13244-020-00887-2
12. Espinasse M, Pitre-Champagnat S, Charmettant B, Bidault F, Volk A, Balleyguier C, et al. CT Texture Analysis Challenges: Influence of Acquisition and Reconstruction Parameters: A Comprehensive Review. *Diagnostics* (2020) 10(5):258. doi: 10.3390/diagnostics10050258
13. Bodalal Z, Trebeschi S, Nguyen-Kim TDL, Schats W, Beets-Tan R. Radiogenomics: bridging imaging and genomics. *Abdominal Radiol* (2019) 44(6):1960–84. doi: 10.1007/s00261-019-02028-w
14. Liu Z, Wang S, Di Dong JW, Fang C, Zhou X, Sun K, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. *Theranostics* (2019) 9(5):1303. doi: 10.7150/thno.30309
15. Bogowicz M, Vuong D, Huellner MW, Pavic M, Andratschke N, Gabrys HS, et al. CT radiomics and PET radiomics: ready for clinical implementation? *Q J Nucl Med Mol Imaging: Off Publ Ital Assoc Nucl Med (AIMN) Int Assoc Radiopharmacol (IAR) Section Soc* (2019) 63(4):355–70. doi: 10.23736/S1824-4785.19.03192-3
16. Park JE PS, Kim HJ. Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives. *Korean J Radiol* (2019) 20:1124–37. doi: 10.3348/kjr.2018.0070
17. Avanzo M, Wei L, Stancanello J, Vallieres M, Rao A, Morin O, et al. Machine and deep learning methods for radiomics. *Med Phys* (2020) 47(5):e185–202. doi: 10.1002/mp.13678
18. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
19. Yip SS, Aerts HJ. Applications and limitations of radiomics. *Phys Med Biol* (2016) 61(13):R150. doi: 10.1088/0031-9155/61/13/R150
20. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. *Korean J Radiol* (2019) 20(3):405–10. doi: 10.3348/kjr.2019.0025
21. Lambin P, Leijenaar RT, Deist TM, Peerlings J, De Jong EE, Van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141

22. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, et al. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. *Eur Radiol* (2020) 30(1):523–36. doi: 10.1007/s00330-019-06360-z
23. Chetan MR, Gleeson FV. Radiomics in predicting treatment response in non-small-cell lung cancer: current status, challenges and future perspectives. *Eur Radiol* (2021) 31(2):1049–58. doi: 10.1007/s00330-020-07141-9
24. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC. Group. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. *Radiology* (2011) 258(3):906–14. doi: 10.1148/radiol.10100799
25. Clarke LP, Nordstrom RJ, Zhang H, Tandon P, Zhang Y, Redmond G, et al. The quantitative imaging network: NCI's historical perspective and planned goals. *Trans Oncol* (2014) 7(1):1. doi: 10.1593/tlo.13832
26. Hatt M, Vallieres M, Visvikis D, Zwanenburg A. IBSI: an international community radiomics standardization initiative. *J Nucl Med* (2018) 59 (supplement 1):287–. doi: 10.2967/jnumed.117.200501
27. deSouza NM, Achten E, Alberich-Bayarri A, Bamberg F, Boellaard R, Clément O, et al. Validated imaging biomarkers as decision-making tools in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European Society of Radiology (ESR). *Insights Imaging* (2019) 10(1):87. doi: 10.1186/s13244-019-0764-0
28. Balagurunathan Y, Kumar V, Gu Y, Kim J, Wang H, Liu Y, et al. Test–retest reproducibility analysis of lung CT image features. *J Digit Imaging* (2014) 27 (6):805–23. doi: 10.1007/s10278-014-9716-x
29. Zhao B, Tan Y, Tsai WY, Schwartz LH, Lu L. Exploring variability in CT characterization of tumors: a preliminary phantom study. *Trans Oncol* (2014) 7(1):88–93. doi: 10.1593/tlo.13865
30. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring CT scanner variability of radiomics features. *Invest Radiol* (2015) 50(11):757. doi: 10.1097/RLI.0000000000000180
31. Zhao B, Tan Y, Tsai W-Y, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* (2016) 6 (1):1–7. doi: 10.1038/srep23428
32. van Timmeren JE, Leijenaar RT, van Elmpt W, Wang J, Zhang Z, Dekker A, et al. Test–retest data for radiomics feature stability analysis: generalizable or study-specific? *Tomography* (2016) 2(4):361. doi: 10.18383/j.tom.2016.00208
33. Larue RT, van Timmeren JE, de Jong EE, Feliciani G, Leijenaar RT, Schreurs WM, et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. *Acta Oncol* (2017) 56(11):1544–53. doi: 10.1080/0284186X.2017.1351624
34. Mahmood U, Apte A, Deasy JO, Schmidtlein CR, Shukla-Dave A. Investigating the robustness Neighborhood Grey Tone Difference Matrix (NGTDM) and Grey Level Co-Occurrence Matrix (GLCM) radiomic features on clinical Computed Tomography systems using Anthropomorphic Phantoms: Evidence from a multivendor study. *J Comput Assist Tomography* (2017) 41(6):995. doi: 10.1097/RCT.0000000000000632
35. Ger RB, Zhou S, Chi P-CM, Lee HJ, Layman RR, Jones AK, et al. Comprehensive investigation on controlling for CT imaging variabilities in radiomics studies. *Sci Rep* (2018) 8(1):1–14. doi: 10.1038/s41598-018-31509-z
36. Berenguer R, Pastor-Juan M, Canales-Vázquez J, Castro-García M, Villas MV, Mansilla Legorburo F, et al. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. *Radiology* (2018) 288(2):407–15. doi: 10.1148/radiol.2018172361
37. Fave X, Cook M, Frederick A, Zhang L, Yang J, Fried D, et al. Preliminary investigation into sources of uncertainty in quantitative imaging features. *Comput Med Imaging Graph* (2015) 44:54–61. doi: 10.1016/j.compmedimag.2015.04.006
38. Lo P, Young S, Kim H, Brown M, McNitt-Gray M. Variability in CT lung-nodule quantification: effects of dose reduction and reconstruction methods on density and texture based features. *Med Phys* (2016) 43(8Part1):4854–65. doi: 10.1118/1.4954845
39. Midya A, Chakraborty J, Gönen M, Do RK, Simpson AL. Influence of CT acquisition and reconstruction parameters on radiomic feature reproducibility. *J Med Imaging* (2018) 5(1):011020. doi: 10.1117/1.JMI.5.1.011020
40. Mackin D, Ger R, Dodge C, Fave X, Chi P-C, Zhang L, et al. Effect of tube current on computed tomography radiomic features. *Sci Rep* (2018) 8(1):1–10. doi: 10.1038/s41598-018-20713-6
41. Hepp T, Othman A, Liebgott A, Kim JH, Pfannenberg C, Gatidis S. Effects of simulated dose variation on contrast-enhanced CT-based radiomic analysis for Non-Small Cell Lung Cancer. *Eur J Radiol* (2020) 124:108804. doi: 10.1016/j.ejrad.2019.108804
42. Mackin D, Fave X, Zhang L, Yang J, Jones AK, Ng CS, et al. Harmonizing the pixel size in retrospective computed tomography radiomics studies. *PloS One* (2017) 12(9):e0178524. doi: 10.1371/journal.pone.0178524
43. Shafiq-ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* (2017) 44(3):1050–62. doi: 10.1002/mp.12123
44. Rastegar S, Beigi J, Saeedi E, Shiri I, Qasempour Y, Rezaei M, et al. Radiographic Image Radiomics Feature Reproducibility: A Preliminary Study on the Impact of Field Size. *J Med Imaging Radiat Sci* (2020) 51 (1):128–36. doi: 10.1016/j.jmir.2019.11.006
45. Lu L, Ehmke RC, Schwartz LH, Zhao B. Assessing agreement between radiomic features computed for multiple CT imaging settings. *PloS One* (2016) 11(12):e0166550. doi: 10.1371/journal.pone.0166550
46. Kim H, Park CM, Lee M, Park SJ, Song YS, Lee JH, et al. Impact of reconstruction algorithms on CT radiomic features of pulmonary tumors: analysis of intra-and inter-reader variability and inter-reconstruction algorithm variability. *PloS One* (2016) 11(10):e0164924. doi: 10.1371/journal.pone.0164924
47. Shafiq-ul-Hassan M, Zhang GG, Hunt DC, Latifi K, Ullah G, Gillies RJ, et al. Accounting for reconstruction kernel-induced variability in CT radiomic features using noise power spectra. *J Med Imaging* (2017) 5(1):011013. doi: 10.1117/1.JMI.5.1.011013
48. Yang J, Zhang L, Fave XJ, Fried DV, Stingo FC, Ng CS, et al. Uncertainty analysis of quantitative imaging features extracted from contrast-enhanced CT in lung tumors. *Comput Med Imaging Graph* (2016) 48:1–8. doi: 10.1016/j.compmedimag.2015.12.001
49. Dercle L, Lu L, Lichtenstein P, Yang H, Wang D, Zhu J, et al. Impact of variability in portal venous phase acquisition timing in tumor density measurement and treatment response assessment: metastatic colorectal cancer as a paradigm. *JCO Clin Cancer Inf* (2017) 1(1):1–8. doi: 10.1200/CCI.17.00108
50. Kakino R, Nakamura M, Mitsuyoshi T, Shintani T, Hirashima H, Matsuo Y, et al. Comparison of radiomic features in diagnostic CT images with and without contrast enhancement in the delayed phase for NSCLC patients. *Phys Med* (2020) 69:176–82. doi: 10.1016/j.ejmp.2019.12.019
51. Larue RT, Van De Voorde L, van Timmeren JE, Leijenaar RT, Berbée M, Sosef MN, et al. 4DCT imaging to assess radiomics feature stability: An investigation for thoracic cancers. *Radiother Oncol* (2017) 125(1):147–53. doi: 10.1016/j.radonc.2017.07.023
52. Tanaka S, Kadoya N, Kajikawa T, Matsuda S, Dobashi S, Takeda K, et al. Investigation of thoracic four-dimensional CT-based dimension reduction technique for extracting the robust radiomic features. *Phys Med* (2019) 58:141–8. doi: 10.1016/j.ejmp.2019.02.009
53. Hunter LA, Krafft S, Stingo F, Choi H, Martel MK, Kry SF, et al. High quality machine-robust image features: Identification in nonsmall cell lung cancer computed tomography images. *Med Phys* (2013) 40(12):121916. doi: 10.1118/1.4829514
54. Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, et al. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non–small cell lung cancer. *Radiology* (2009) 252 (1):263–72. doi: 10.1148/radiol.2522081593
55. Prior F, Smith K, Sharma A, Kirby J, Tarbox L, Clark K, et al. The public cancer radiology imaging collections of The Cancer Imaging Archive. *Sci Data* (2017) 4:170124. doi: 10.1038/sdata.2017.124
56. Lu L, Liang Y, Schwartz LH, Zhao B. Reliability of Radiomic Features Across Multiple Abdominal CT Image Acquisition Settings: A Pilot Study Using ACR CT Phantom. *Tomography* (2019) 5(1):226. doi: 10.18383/j.tom.2019.00005

57. Lewis JH, Jiang SB. A theoretical model for respiratory motion artifacts in free-breathing CT scans. *Phys Med Biol* (2009) 54(3):745. doi: 10.1088/0031-9155/54/3/018
58. Pan T, Lee TY, Rietzel E, Chen GT. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med Phys* (2004) 31(2):333–40. doi: 10.1118/1.1639993
59. Du Q, Baine M, Bavitz K, McAllister J, Liang X, Yu H, et al. Radiomic feature stability across 4D respiratory phases and its impact on lung tumor prognosis prediction. *PloS One* (2019) 14(5):e0216480. doi: 10.1371/journal.pone.0216480
60. Lafata K, Cai J, Wang C, Hong J, Kelsey CR, Yin F-F. Spatial-temporal variability of radiomic features and its effect on the classification of lung cancer histology. *Phys Med Biol* (2018) 63(22):225003. doi: 10.1088/1361-6560/aae56a
61. Valladares A, Beyer T, Rausch I. Physical imaging phantoms for simulation of tumor heterogeneity in PET, CT, and MRI: an overview of existing designs. *Med Phys* (2020) 47(4):2023. doi: 10.1002/mp.14045
62. Zhang L, Fried DV, Fave XJ, Hunter LA, Yang J, Court LE. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys* (2015) 42(3):1341–53. doi: 10.1118/1.4908210
63. Samei E, Hoye J, Zheng Y, Solomon JB, Marin D. Design and fabrication of heterogeneous lung nodule phantoms for assessing the accuracy and variability of measured texture radiomics features in CT. *J Med Imaging* (2019) 6(2):021606. doi: 10.1117/1.JMI.6.2.021606
64. Limkin EJ, Reuzé S, Carré A, Sun R, Schernberg A, Alexis A, et al. The complexity of tumor shape, spiculatedness, correlates with tumor radiomic shape features. *Sci Rep* (2019) 9(1):1–12. doi: 10.1038/s41598-019-40437-5
65. Choe J, Lee SM, Do K-H, Lee G, Lee J-G, Lee SM, et al. Deep learning–based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses. *Radiology* (2019) 292 (2):365–73. doi: 10.1148/radiol.2019181960
66. Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77(21):e104–e7. doi: 10.1158/0008-5472.CAN-17-0339
67. Park S, Lee SM, Do KH, Lee JG, Bae W, Park H, et al. Deep Learning Algorithm for Reducing CT Slice Thickness: Effect on Reproducibility of Radiomic Features in Lung Cancer. *Korean J Radiol* (2019) 20:1431–40. doi: 10.3348/kjr.2019.0212
68. Solomon JB, Christianson O, Samei E. Quantitative comparison of noise texture across CT scanners from different manufacturers. *Med Phys* (2012) 39(10):6048–55. doi: 10.1118/1.4752209
69. Winslow J, Zhang Y, Samei E. A method for characterizing and matching CT image quality across CT scanners from different manufacturers. *Med Phys* (2017) 44(11):5705–17. doi: 10.1002/mp.12554
70. Ma J, Dercle L, Lichtenstein P, Wang D, Chen A, Zhu J, et al. Automated identification of optimal portal venous phase timing with convolutional neural networks. *Acad Radiol* (2020) 27(2):e10–e8. doi: 10.1016/j.acra.2019.02.024
71. Dercle L, Ma J, Xie C, A-p C, Wang D, Luk L, et al. Using a single abdominal computed tomography image to differentiate five contrast-enhancement phases: A machine-learning algorithm for radiomics-based precision medicine. *Eur J Radiol* (2020) 125:1–9. doi: 10.1016/j.ejrad.2020.108850
72. Dercle L, Lu L, Schwartz LH, Qian M, Tejpar S, Eggleton P, et al. Radiomics response signature for identification of metastatic colorectal cancer sensitive to therapies targeting EGFR pathway. *JNCI: J Natl Cancer Inst* (2020) 112 (9):902–12. doi: 10.1093/jnci/djaa017
73. He L, Huang Y, Ma Z, Liang C, Liang C, Liu Z. Effects of contrast-enhancement, reconstruction slice thickness and convolution kernel on the diagnostic performance of radiomics signature in solitary pulmonary nodule. *Sci Rep* (2016) 6:34921. doi: 10.1038/srep34921
74. Li Y, Lu L, Xiao M, Dercle L, Huang Y, Zhang Z, et al. CT slice thickness and convolution kernel affect performance of a radiomic model for predicting EGFR status in non-small cell lung cancer: a preliminary study. *Sci Rep* (2018) 8(1):1–10. doi: 10.1038/s41598-018-36421-0
75. Ger RB, Zhou S, Elgohari B, Elhalawani H, Mackin DM, Meier JG, et al. Radiomics features of the primary tumor fail to improve prediction of overall survival in large cohorts of CT-and PET-imaged head and neck cancer patients. *PloS One* (2019) 14(9):e0222509. doi: 10.1371/journal.pone.0222509
76. Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: the need for safeguards. *Radiother Oncol* (2019) 130:2–9. doi: 10.1016/j.radonc.2018.10.027
77. Zwanenburg A, Leger S, Agolli L, Pilz K, Troost EG, Richter C, et al. Assessing robustness of radiomic features by image perturbation. *Sci Rep* (2019) 9(1):1–10. doi: 10.1038/s41598-018-36938-4
78. Zhu J-Y, Park T, Isola P, Efros AA, eds.. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proc IEEE Int Conf Comput Vis*. (2017) 2242–51. doi: 10.1109/ICCV.2017.244
79. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell* (2017) 39(4):640–51. doi: 10.1109/TPAMI.2016.2572683
80. Kalpathy-Cramer J, Zhao B, Goldgof D, Gu Y, Wang X, Yang H, et al. A comparison of lung nodule segmentation algorithms: methods and results from a multi-institutional study. *J Digit Imaging* (2016) 29(4):476–87. doi: 10.1007/s10278-016-9859-z
81. Kalpathy-Cramer J, Mamomov A, Zhao B, Lu L, Cherezov D, Napel S, et al. Radiomics of lung nodules: a multi-institutional study of robustness and agreement of quantitative imaging features. *Tomography* (2016) 2(4):430–7. doi: 10.18383/j.tom.2016.00235
82. Tunali I, Hall LO, Napel S, Cherezov D, Guvenis A, Gillies RJ, et al. Stability and reproducibility of computed tomography radiomic features extracted from peritumoral regions of lung cancer lesions. *Med Phys* (2019) 46 (11):5075–85. doi: 10.1002/mp.13808
83. Paul R, Hassan M-u, Moros EG, Gillies RJ, Hall LO, Goldgof DB. Deep feature stability analysis using CT images of a physical phantom across scanner manufacturers, cartridges, pixel sizes, and slice thickness. *Tomography* (2020) 6(2):250. doi: 10.18383/j.tom.2020.00003
84. Parmar C, Velazquez ER, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PloS One* (2014) 9(7):e102107. doi: 10.1371/journal.pone.0102107
85. Qiu Q, Duan J, Duan Z, Meng X, Ma C, Zhu J, et al. Reproducibility and non-redundancy of radiomic features extracted from arterial phase CT scans in hepatocellular carcinoma patients: impact of tumor segmentation variability. *Quant Imaging Med Surg* (2019) 9(3):453. doi: 10.21037/qims.2019.03.02
86. Pavic M, Bogowicz M, Würms X, Glatz S, Finazzi T, Riesterer O, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. *Acta Oncol* (2018) 57(8):1070–4. doi: 10.1080/0284186X.2018.1445283
87. Huang Q, Lu L, Dercle L, Lichtenstein P, Li Y, Yin Q, et al. Interobserver variability in tumor contouring affects the use of radiomics to predict mutational status. *J Med Imaging* (2017) 5(1):011005. doi: 10.1117/1.JMI.5.1.011005
88. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026
89. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* (2012) 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001
90. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* (2006) 31 (3):1116–28. doi: 10.1016/j.neuroimage.2006.01.015
91. Tan Y, Guo P, Mann H, Marley SE, Scott MLJ, Schwartz LH, et al. Assessing the effect of CT slice interval on unidimensional, bidimensional and volumetric measurements of solid tumours. *Cancer Imaging* (2012) 12 (3):497. doi: 10.1102/1470-7330.2012.0046
92. Zhao B, Tan Y, Bell DJ, Marley SE, Guo P, Mann H, et al. Exploring intra- and inter-reader variability in uni-dimensional, bi-dimensional, and volumetric measurements of solid tumors on CT scans reconstructed at different slice intervals. *Eur J Radiol* (2013) 82(6):959–68. doi: 10.1016/j.ejrad.2013.02.018

93. Petrick N, Kim HJG, Clunie D, Borradaile K, Ford R, Zeng R, et al. Comparison of 1D, 2D, and 3D nodule sizing methods by radiologists for spherical and complex nodules on thoracic CT phantom images. *Acad Radiol* (2014) 21(1):30–40. doi: 10.1016/j.acra.2013.09.020
94. McNitt-Gray MF, Kim GH, Zhao B, Schwartz LH, Clunie D, Cohen K, et al. Determining the variability of lesion size measurements from CT patient data sets acquired under "no change" conditions. *Trans Oncol* (2015) 8 (1):55–64. doi: 10.1016/j.tranon.2015.01.001
95. Athelogou M, Kim HJ, Dima A, Obuchowski N, Peskin A, Gavrielides MA, et al. Algorithm variability in the estimation of lung nodule volume from phantom CT scans: results of the QIBA 3A public challenge. *Acad Radiol* (2016) 23(8):940–52. doi: 10.1016/j.acra.2016.02.018
96. Summers RM. Texture analysis in radiology: Does the emperor have no clothes? *Abdominal Radiol* (2017) 42(2):342–5. doi: 10.1007/s00261-016-0950-1
97. Lee H-J, Kim YT, Kang CH, Zhao B, Tan Y, Schwartz LH, et al. Epidermal growth factor receptor mutation in lung adenocarcinomas: relationship with CT characteristics and histologic subtypes. *Radiology* (2013) 268(1):254–64. doi: 10.1148/radiol.13112553
98. Lu L, Wang D, Wang L, Linning E, Guo P, Li Z, et al. A quantitative imaging biomarker for predicting disease-free-survival-associated histologic subgroups in lung adenocarcinoma. *Eur Radiol* (2020) 30(7):3614–23. doi: 10.1007/s00330-020-06663-6
99. Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. *Med Image Anal* (2014) 18(1):176–96. doi: 10.1016/j.media.2013.10.005
100. Lu L, Schwartz L, Zhao B. Chapter 3. Quantitative imaging using CT. In: Li R, Lei Z, Napel S, Rubin DL, editors. *RADIOMICS AND RADIOGENOMICS: Technical Basis and Clinical Applications*. London: CRC Press (2019).
101. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. *IEEE Trans Syst Man Cybern* (1973) 6):610–21. doi: 10.1109/TSMC.1973.4309314
102. Foy JJ, Robinson KR, Li H, Giger ML, Al-Hallaq H, Armato SG. Variation in algorithm implementation across radiomics software. *J Med Imaging* (2018) 5(4):044505. doi: 10.1117/1.JMI.5.4.044505
103. McNitt-Gray M, Napel S, Jaggi A, Mattonen S, Hadjiiski L, Muzi M, et al. Standardization in quantitative imaging: a multicenter comparison of radiomic features from different software packages on digital reference objects and patient data sets. *Tomography* (2020) 6(2):118–28. doi: 10.18383/j.tom.2019.00031
104. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJ, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
105. Jaggi A, Mattonen MS, McNitt-Gray M, Napel S. Stanford DRO Toolkit: Digital Reference Objects for Standardization of Radiomic Features. *Tomography* (2020) 6:111–7. doi: 10.18383/j.tom.2019.00030
106. Armato SG II, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Med Phys* (2011) 38(2):915–31. doi: 10.1118/1.3528204
107. Orlhac F, Boughdad S, Philippe C, Stalla-Bourdillon H, Nioche C, Champion L, et al. A postreconstruction harmonization method for multicenter radiomic studies in PET. *J Nucl Med* (2018) 59(8):1321–8. doi: 10.2967/jnumed.117.199935
108. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of a method to compensate multicenter effects affecting CT radiomics. *Radiology* (2019) 291 (1):53–9. doi: 10.1148/radiol.2019182023
109. Mahon R, Ghita M, Hugo G, Weiss E. ComBat harmonization for radiomic features in independent phantom and lung cancer patient computed tomography datasets. *Phys Med Biol* (2020) 65(1):015010. doi: 10.1088/1361-6560/ab6177
110. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* (2007) 8 (1):118–27. doi: 10.1093/biostatistics/kxj037
111. Pfaehler E, Zwanenburg A, de Jong JR, Boellaard R. RaCaT: An open source and easy to use radiomics calculator tool. *PloS One* (2019) 14(2):e0212223. doi: 10.1371/journal.pone.0212223
112. Freymann JB, Kirby JS, Perry JH, Clunie DA, Jaffe CC. Image data sharing for biomedical research—meeting HIPAA requirements for de-identification. *J Digit Imaging* (2012) 25(1):14–24. doi: 10.1007/s10278-011-9422-x
113. Zanfardino M, Pane K, Mirabelli P, Salvatore M, Franzese M. TCGA-TCIA impact on radiogenomics cancer research: a systematic review. *Int J Mol Sci* (2019) 20(23):6033. doi: 10.3390/ijms20236033
114. Lu L, Sun SH, Yang H. Radiomics prediction of EGFR status in lung cancer—our experience in using multiple feature extractors and the Cancer Imaging Archive data. *Tomography* (2020) 6(2):223–30. doi: 10.18383/j.tom.2020.00017
115. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics* (2017) 37(5):1483–503. doi: 10.1148/rg.2017170056
116. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys* (2018) 102(4):1143–58. doi: 10.1016/j.ijrobp.2018.05.053
117. Lawrence I, Lin K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* (1989) 45(1):255–68. doi: 10.2307/2532051

Conflict of Interest: The author BZ receives royalties from Varian Medical Systems and Keosys Medical Imaging and funding from NIH.

frontiers in Oncology

ORIGINAL RESEARCH
published: 15 April 2021
doi: 10.3389/fonc.2021.658669

The Establishment of New Thresholds for PLND-Validated Clinical Nomograms to Predict Non-Regional Lymph Node Metastases: Using ^{68}Ga-PSMA PET/CT as References

Jianhua Jiao[1†], Zhiyong Quan[2†], Jingliang Zhang[1†], Weihong Wen[3], Jun Qin[1], Lijun Yang[1], Ping Meng[1], Yuming Jing[1], Shuaijun Ma[1], Peng Wu[1], Donghui Han[1], Andrew A. Davis[4], Jing Ren[5], Xiaojian Yang[1], Fei Kang[2], Qiang Zhang[4], Jing Wang[2*] and Weijun Qin[1*]*

[1] Department of Urology, Xijing Hospital, Fourth Military Medical University, Xi'an, China, [2] Department of Nuclear Medicine, Xijing Hospital, Fourth Military Medical University, Xi'an, China, [3] Institute of Medical Research, Northwestern Polytechnical University, Xi'an, China, [4] Department of Medicine, Division of Hematology/Oncology, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, [5] Department of Radiology, Xijing Hospital, Fourth Military Medical University, Xi'an, China

Purpose: PLND (pelvic lymph node dissection)-validated nomograms are widely accepted clinical tools to determine the necessity of PLND by predicting the metastasis of lymph nodes (LNMs) in pelvic region. However, these nomograms are in lacking of a threshold to predict the metastasis of extrareolar lymph nodes beyond pelvic region, which is not suitable for PLND. The aim of this study is to evaluate a threshold can be set for current clinical PLND-validated nomograms to predict extrareolar LN metastases beyond pelvic region in high-risk prostate cancer patients, by using ^{68}Ga-PSMA PET/CT as a reference to determine LN metastases (LNMs).

Experimental Design: We performed a retrospective analysis of 57 high-risk treatment-naïve PC patients in a large tertiary care hospital in China who underwent ^{68}Ga-PSMA-617 PET/CT imaging. LNMs was detected by ^{68}Ga-PSMA-617 PET/CT and further determined by imaging follow-up after anti-androgen therapy. The pattern of LN metastatic spread of PC patients were evaluated and analyzed. The impact of ^{68}Ga-PSMA PET/CT on clinical decisions based on three clinical PLND-validated nomograms (Briganti, Memorial Sloan Kettering Cancer Center, Winter) were evaluated by a multidisciplinary prostate cancer therapy team. The diagnostic performance and the threshold of these nomograms in predicting extrareolar LNMs metastasis were evaluated *via* receiver operating characteristic (ROC) curve analysis.

OPEN ACCESS

Edited by:
Fatima-Zohra Mokrane, Université Toulouse III Paul Sabatier, France

Reviewed by:
Kaijie Wu, Xi'an Jiaotong University, China
Sikandar Shaikh, Shadan Hospital and Institute of Medical Sciences, India

***Correspondence:**
Weijun Qin
qinwj@fmmu.edu.cn
Jing Wang
wangjing@fmmu.edu.cn
Qiang Zhang
q-zhang2@northwestern.edu

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *26 January 2021*
Accepted: *09 March 2021*
Published: *15 April 2021*

Citation:
Jiao J, Quan Z, Zhang J, Wen W, Qin J, Yang L, Meng P, Jing Y, Ma S, Wu P, Han D, Davis AA, Ren J, Yang X, Kang F, Zhang Q, Wang J and Qin W (2021) The Establishment of New Thresholds for PLND-Validated Clinical Nomograms to Predict Non-Regional Lymph Node Metastases: Using ^{68}Ga-PSMA PET/CT as References. Front. Oncol. 11:658669. doi: 10.3389/fonc.2021.658669

Results: LNMs were observed in 49.1% of the patients by ^{68}Ga-PSMA PET/CT, among which 65.5% of LNMs were pelvic-regional and 34.5% of LNMs were observed in extrareolar sites (52.1% of these were located above the diaphragm). The Briganti, MSKCC and Winter nomograms showed that 70.2%-71.9% of the patients in this study need to receive ePLND according to the EAU and NCCN guidelines. The LN staging information obtained from ^{68}Ga-PSMA PET/CT would have led to changes of planned management in 70.2% of these patients, including therapy modality changes in 21.1% of the patients, which were mainly due to newly detected non-regional LNMs. The thresholds of nomograms to predict non-regional LNMs were between 64% and 75%. The PC patients with a score >64% in Briganti nomogram, a score >75% in MSKCC nomogram and a score >67% in Winter nomogram were more likely to have non-regional LNMs. The AUCs (Area under curves) of the clinical nomograms (Briganti, MSKCC and Winter) in predicting non-regional LNMs were 0.816, 0.830 and 0.793, respectively.

Conclusions: By using ^{68}Ga-PSMA PET/CT as reference of LNM, the PLND-validated clinical nomograms can not only predict regional LNMs, but also predict non-regional LNMs. The additional information from ^{68}Ga-PSMA PET/CT may provide added benefit to nomograms-based clinical decision-making in more than two-thirds of patients for reducing unnecessary PLND. We focused on that a threshold can be set for current clinical PLND-validated nomograms to predict extrareolar LN metastases with an AUC accuracy of about 80% after optimizing the simple nomograms which may help to improve the efficiency for PC therapy significantly in clinical practice.

Keywords: prostate cancer, PET/CT, distant, lymph node metastases, PSMA, impact, nomogram

HIGHLIGHTS

^{68}Ga-PSMA PET/CT demonstrated that nearly one-third of LNMs in high-risk treatment-naïve PC patients were non-regional and the new information obtained from whole-body ^{68}Ga-PSMA PET/CT has the potential to benefit the nomograms-based clinical decision-making in more than two-thirds of the patients, leading to more personalized treatment. ^{68}Ga-PSMA PET/CT can be used to exclude patients with non-regional lymph node metastases (LNMs) before pelvic lymph node dissection (PLND) in order to improve the efficiency of prostate cancer therapy in clinical practice. With a higher cutoff value, the clinical nomograms have the potential to predict non-regional LNMs as well.

Abbreviations: LNMs, lymph node metastases; PC, prostate cancer; MDT, multidisciplinary cancer therapy; PSMA, prostate specific membrane antigen; PET/CT, positron emission tomography computed tomography; PLND, pelvic lymph node dissection; ePLND, extend PLND; sPLND, sentinel PLND; CT, computed tomography; mpMRI, multiparametric magnetic resonance imaging; DWI, diffusion-weighted imaging; USPIO, ultra-small paramagnetic iron oxide particles; tPSA, total prostate specific antigen; GS, Gleason score; ADT, androgen deprivation treatment; RP, radical prostatectomy; RT, radiotherapy; IHC, immunohistochemistry; IF, immunofluorescence; l.n., lymph nodes; BMs, bone metastases; IGRT, image-guided radiotherapy; SUV_{max}, Maximum Standardized Uptake Values; HE, hematoxylin-eosin; RP, radical prostatectomy; TRUS, transrectal ultrasound; AUC, Area under curve.

INTRODUCTION

Prostate cancer (PC) is the most commonly diagnosed malignant tumor and the second leading cause of cancer associated death in men worldwide (1). The most common route of metastasis is lymphogenous spread (2), and regional lymph node metastases (LNMs) can be observed in high-risk PC patients at 19.4% (3). The situation of LNMs can greatly influence clinical decision making in PC patients. Currently, the conventional imaging techniques such as computed tomography (CT) and multiparametric magnetic resonance imaging (mpMRI) had limited utility because of their relatively low sensitivities of approximately 40.0% for detection of LNMs (4). To better predict the risk of regional LNMs in PC patients, multiple pelvic lymph node dissection (PLND) validated nomograms, such as the Briganti, Memorial Sloan Kettering Cancer Center (MSKCC) and Winter nomograms were created to identify optimal candidates for PLND in treatment-naïve PC patients, according to EAU (European Association of Urology) and NCCN (National Comprehensive Cancer Network) guidelines (5–7). PLND should be performed in the patients with LNMs risk higher than 2% in MSKCC nomogram, 5% in Briganti nomogram or 7% in Winter nomogram (5–7).

The incidence rate of non-regional LNMs leading to unnecessary PLND in treatment-naïve PC patients were as high as 36.0% (8). However, according to the EAU–ESTRO–SIOG and NCCN guidelines, there is no specific examination for

detecting non-regional LNMs in guidelines and only a cross-sectional abdominopelvic imaging (CT/MRI) and a bone scan are routinely recommended for staging purposes (7, 9). Therefore, many PC patients may receive unnecessary PLND because they already have non-regional LNMs. There were no thresholds in PLND-validated clinical nomograms (MSKCC, Briganti and Winter) to exclude the PC patients with non-regional LNMs. If there were thresholds to identify the PC patients with non-regional LNMs in PLND-validated clinical nomograms, the nomograms can be better used in selection of PLND candidates.

However, the detection of non-regional LNMs and development of nomograms to predict non-regional LNMs were limited by the imaging techniques to detect non-regional LNMs. In clinical practice, there was no routine imaging examination for non-regional LNMs and the nodal staging largely depended on preoperative pelvic CT/MRI or histopathological results from PLND. The whole-body MRI and biopsies for non-regional LNMs were hard to performed routinely. With high accuracy, ^{68}Ga-PSMA PET/CT can be an effective imaging modality to detect non-regional LNMs. Prostate-specific membrane antigen (PSMA), a type II transmembrane protein, was overexpressed in 98.0% of PC associated LNMs (10). ^{68}Ga-labeled prostate-specific membrane antigen PET/CT (^{68}Ga-PSMA PET/CT) is a novel molecular imaging technique with promise compared to conventional imaging techniques for detection of LNMs for primary staging of PC patients (4, 11, 12). The use of ^{68}Ga-PSMA PET in probing regional LNMs of PC was validated in a series of prior studies (13–16). The specificity of ^{68}Ga-PSMA PET for detecting regional LNMs ranged from 80.0% to 100.0% (13) and confirmed high specificity of over 95.0% in large cohorts (17, 18). In hence, ^{68}Ga-PSMA PET is now an established imaging technique to improve the detection of non-regional LNMs in prostate cancer (19).

The disease with regional LNMs can be considered as locoregional progressive disease between localized disease and the oligometastatic disease or systemic metastatic disease with distant LNMs (20, 21). Therefore, the aim of the present study was to investigate whether clinical nomograms (Briganti, MSKCC and Winter) can predict non-regional LNMs and to generate cutoff values to predict distant LNMs. Further, we analyzed the potential added benefit of visualizing newly detected distant LNMs by ^{68}Ga-PSMA PET/CT to existing nomograms-based clinical decision-making. The non-regional LNMs observed by ^{68}Ga-PSMA PET/CT were detected and monitored by multiple imaging techniques in follow-ups of the patients. In addition, we depicted metastatic pattern of LNMs in high-risk treatment-naïve PC patients.

PATIENTS AND METHODS

Patients and Study Design

We retrospectively reviewed a database from a large tertiary care hospital in China for patients with pathologically confirmed PC who underwent ^{68}Ga-PSMA PET/CT from April 2017 to October 2019. Patients were included in the study if they met the following inclusion criteria: (1) transrectal ultrasound (TRUS)-guided 12-core biopsy to pathologically confirm PC; (2) Gleason score (GS); (3) clinical tumor stage; (4) pretreatment total prostate specific antigen (tPSA). Patients were excluded if they received any treatment before ^{68}Ga-PSMA PET/CT, such as androgen deprivation treatment (ADT), radical prostatectomy (RP), radiotherapy (RT), or chemotherapy. Other exclusion criteria included patients with negative PSMA expression on primary PC tumor validated by immunohistochemistry (IHC), and patients who had an interval between tPSA data and ^{68}Ga-PSMA PET/CT that was more than 30 days. All patients had a bone scan and an mpMRI as well as ^{68}Ga-PSMA PET/CT. Ultimately, a total of 57 patients with sufficient clinical data were eligible for analysis. Mean patient age was 69.4 ± 8.2 years (Median 68.5, range 40-84) and the mean serum PSA at imaging was 283.9 ng/ml (median 28.91, range 0.09-8447). All patients had high risk PC, according to the D'Amico standard (22).

This study was performed in the Urology Department and Nuclear Medicine Department of the Fourth Military Medical University Affiliated Hospital (Xijing Hospital, Xi'an, Shaanxi, China). The study was approved by the Ethics Committee of Fourth Military Medical University, and all participating patients provided written informed consent. The research was conducted in accordance with the Declaration of Helsinki and national regulations. Deidentified data were collected in a central database at Fourth Military Medical University.

Histological Examination

A TRUS-guided 12-core prostate biopsy with necessary additional target biopsy was performed for each patient's biopsy. All resected tissue of primary PC tumors and lymph nodes from surgeries were formalin-fixed and routinely processed for hematoxylin-eosin (HE) staining and immunohistochemistry (IHC) analysis. The Gleason score (ISUP grade) was considered as the highest score on the biopsy specimen. The histopathological results served as a reference and were stratified in accordance with the 7th edition of the American Joint Committee on Cancer (AJCC) staging system for PC (23). The pathological results were confirmed by the consensus of two board-certified specialists in genitourinary pathology, as previously reported (24). The pathologists were blinded to both the ^{68}Ga-PSMA PET/CT results and the clinical evaluation of the tissues from the surgeons.

Immunohistochemistry Staining

The tissue samples were formalin-fixed and routinely processed for IHC staining to evaluate PSMA expression with anti-PSMA antibody (1:100, MAB-0575, MXB Biotechnologies), as previously reported (25). Further methods related to IHC are included in the **Supplementary Materials** and *Methods*.

Immunofluorescence Staining

The slides were processed for IF staining of PSMA (1:50, MAB-0575, MXB Biotechnologies) and p504s (a biomarker of PC, 1:50, RMA-0546, MXB Biotechnologies), a technique that was

previously reported (26). Details regarding IF procedures are included in the **Supplementary Materials** and *Methods*.

Imaging Evaluation of mpMRI and Bone Scan

All mpMRI evaluations were performed on a 3.0-T MR scanner (Achieva 3.0 T TX, Philips Medical Systems, The Netherlands) by using a 16-channel phased-array coil, as we previously described (27). LNs were rated as malignant if they have a short-axis diameter > 10 mm and if they showed restricted diffusion on the DWI and ADC map or increased contrast enhancement (28). Bone scan was performed on Symbia T2 (Siemens Medical Solutions, Erlangen, Germany) with at 3-5 hours after injection of 20-25mCi $^{99}Tc^{m}$-MDP. The evaluation of mpMRI or bone scan were reviewed by two radiologists or two board-certified nuclear medicine specialists

Imaging Protocol and Evaluation of ^{68}Ga-PSMA PET/CT

All ^{68}Ga-PSMA PET/CT evaluations were performed at a single center (Fourth Military Medical University, Xijing Hospital, Xi'an, Shaanxi, China). Patients underwent ^{68}Ga-PSMA PET imaging on a Biograph 40 system (Siemens Medical Solutions, Erlangen, Germany). The ^{68}Ga/^{68}Ge generator system was produced by ITG GmbH (Munich, Germany), and the DOTA ligand was acquired from ABX GmbH (Radeberg, Germany). The ^{68}Ga-PSMA-617 was synthesized as we previously reported (24), and the patients were intravenously injected with 1.8-2.2 MBq/kg body weight ^{68}Ga-PSMA-617. Mean injection activity of ^{68}Ga-PSMA PET was 141.7 ± 21.9 MBq. Low-dose CT (pitch 0.8, 50 mA, 120KV[peak]) scans for PET attenuation were obtained (automatic mA, 120keV, 512x512 matrix, 5-mm slice thickness, 1.0-s rotation time, and 0.8 pitch), followed by a PET scan with 5 bed positions (3 min/bed, from head to the proximal thighs) performed about 60 minutes after tracer injection. The PET/CT images were then transferred to a multimodal workstation for data analysis (Syngo Truepoint Siemens Medical Solutions).

The scans of ^{68}Ga-PSMA PET/CT were reviewed by two board-certified nuclear medicine specialists (Z.Q. and F.K.) with more than ten years' experience in reading PET imaging and one board-certified radiation oncologist (J.W.). According to prior studies, lymph nodes with a SUV_{max} of 2.0 or more and a diameter of 5 mm or more were considered PSMA-positive on ^{68}Ga-PSMA PET (29, 30). Scans were evaluated using a Siemens MIWP workstation (Syngo MIWP; Siemens Medical Solutions, Erlangen, Germany), according to the Joint EANM and SNMMI procedure guidelines (version 1.0) (31, 32).

Statistical Analysis

Descriptive statistics were calculated and presented as the frequency (percentage) for categorical variables, the mean (standard deviation) for continuous variables of normal distribution and the median (quartile) for continuous variables of skewness distribution. All data were analyzed by IBM SPSS statistics software, version 23.0 (IBM, Inc., Chicago, IL, USA).

RESULTS

Patient Characteristics and Pattern of Metastatic Spread

The characteristics of the 57 patients included in the study are summarized in **Table 1**. To better understand the pattern of metastatic spread of LNMs, we depicted non-regional LNMs from all PC patients (**Figure 1**). ^{68}Ga-PSMA PET/CT visual analysis found 206 PSMA-positive LNMs from 49.1% (28/57) of the patients (mean 7.4 nodes per patient; range: 1-30 positive nodes per patient). A representative PC patient with non-regional oligometastatic lymph nodes is shown in **Figure 2** and a PC patient without LNMs is shown in **Figure S1**. **Figure S2** shows a representative PC patient with LNMs had negative results in mpMRI but positive results in ^{68}Ga-PSMA PET/CT. If staining confirmed co-expression of PSMA and P504s, a biomarker of PC cells, on resected LNMs (**Figure S4**).

The Detection and Distribution of Non-Regional LNMs by ^{68}Ga-PSMA PET/CT

To clarify the non-regional LNMs detected by ^{68}Ga-PSMA PET/CT and its distribution, we depicted non-regional LNMs from all PC patients (**Figure 1**). For 57 patients, 34.5% (71/206) of the LNMs were non-regional and 65.5% (135/206) of the LNMs were regional.

As shown in **Table 2**, non-regional LNMs were observed in 57.1% (16/28) of the PC patients with LNMs. Nearly half of all

TABLE 1 | Characteristics of patients and tumors at diagnosis.

Characteristic		value
Age (ages)		
	Mean ± SD	68.5 ± 8.2
	Median (range)	69.0 (40-84)
tPSA at PSMA PET/CT (ng/mL)		
	Median (P_{25}-P_{75})	30.7 (8.8-149.5)
SUV_{max}		
	Mean ± SD	21.3 ± 19.2
	Median (P_{25}-P_{75})	16.0 (7.9-25.1)
Injection dose (MBq)		
	Mean ± SD	141.9 ± 21.5
	Median (range)	142.1(67.0-181.3)
Uptake time (minutes)		
	Mean ± SD	66.617 ± 14.5
	Median (range)	62(40-98)
T-stage, n (%)		
	T2a	5 (8.8%)
	T2b	5 (8.8%)
	T2c	38 (66.7%)
	T3a	2 (3.5%)
	T3b	4 (7.0%)
	T4	3 (5.3%)
Gleason score, n (%)		
	6	3 (5.3%)
	3+4 = 7	2 (3.5%)
	4+3 = 7	8 (14.0%)
	8	22 (38.6%)
	9	17(29.8%)
	10	5(8.8%)
Risk-group according to D'Amico, n (%)		
	High	57(100.0%)

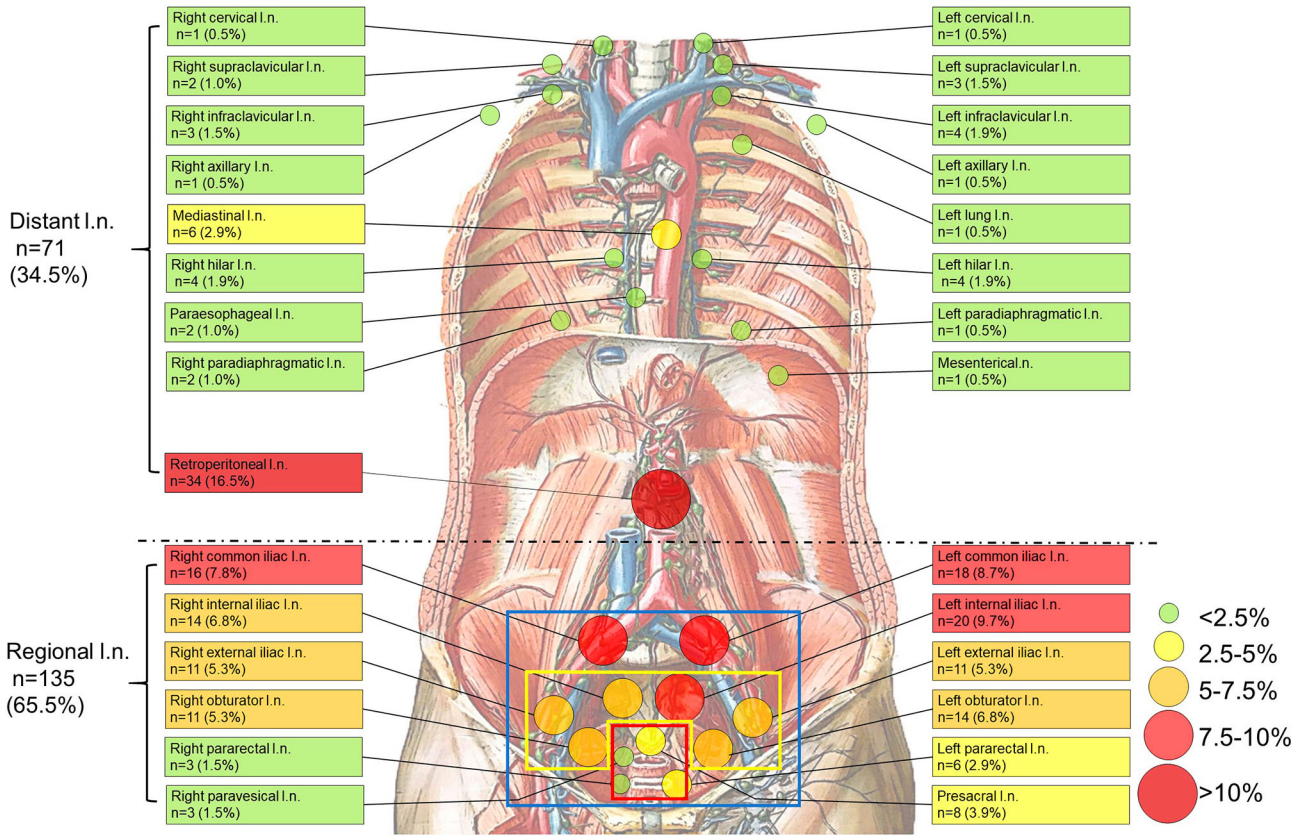


FIGURE 1 | Anatomical distribution of PSMA-avid lymph node metastases in the whole body. (blue square, the scope of ePLND; yellow square, the scope of limited PLND; red square, the selective resected lymph nodes if radioactive in sentinel PLND). Anatomical distribution of non-regional PSMA-avid lymph node metastases. All non-regional lymph node metastases (n=71, 34.5%) are beyond pelvic mpMRI, the currently recommended imaging by the guidelines (7, 9). Background picture (Lymph Vessels and Nodes of Posterior Abdominal Wall) was used under the permission of © Elsevier Inc.

LNMs occurred in retroperitoneal lymph nodes (n=34, 16.5%), while the remaining LNMs (n=37, 18.0%) were observed in other areas above the diaphragm. The most commonly observed LNMs above the diaphragm were seen in the mediastinal area (n=6, 2.9%). PSMA-avid LNMs in the left infraclavicular lymph nodes, the left hilar lymph nodes, and the right hilar lymph nodes were equally observed (n=4, 1.9%). PSMA-positive LNMs were also observed in the left supraclavicular lymph nodes (Virchow nodes, n=3, 4.2%) and the right infraclavicular lymph nodes (n=3, 4.2%). A patient with Virchow nodes is shown in **Figures 3A, B**, and a patient with bilateral supraclavicular LNMs is shown in **Figures 3C, D**. In addition, LNMs were observed in the left/right cervical lymph nodes, the right supraclavicular lymph nodes, the left/right axillary lymph nodes, the left/right hilar lymph nodes, the paraesophageal lymph nodes, and the left/right paradiaphragmatic lymph nodes. One mesenteric lymph node and one metastatic node of the left lung were also observed. In total, more than one-third of all LNMs were non-regional LNMs, which were beyond the range of PLND (**Figure 1**, **Table 2**).

Thus, more than one-third of all LNMs shown in the ^{68}Ga-PSMA PET/CT were non-regional LNMs, which were beyond the range of PLND. Totally, 28.1% (16/57) of all patients have both regional and non-regional LNMs, and no skip metastases was observed in each patient.

The Evaluation of Regional LNMs Risks According to Three Clinical Nomograms

To compare the scores of having regional LNMs, we calculated the risks of having LNMs for the patients in the study according to the three PLND-based nomograms (5, 6, 33).

As shown in **Figure 4**, the PC patients with nonregional LNMs have higher scores than those without LNMs. The risks of LNMs according to the Briganti, MSKCC and Winter nomograms were 48.0% (median; range 1.0-95.0%), 63.0% (median; range 9.0-99.0%), and 70.0% (median; range 10.0-89.0%), respectively (**Table 3**). For the Briganti nomogram, the risks varied from low (<10.0%) for 13 men (22.8%) to very high (>50.0%) for 25 men (43.9%) in our cohort. For the MSKCC nomogram, the risks varied from low (<10.0%) for 10 men (17.5%) to very high (>50.0%) for 35 men (61.4%) in our cohort. For the Winter nomogram, the risks varied from low (<10.0%) for 2 men (3.5%) to very high (>50.0%) for 35 men (61.4%). Due to high tPSA values or GS, the risk of LNMs in 18 patients was more than 90.0% from the MSKCC nomogram.

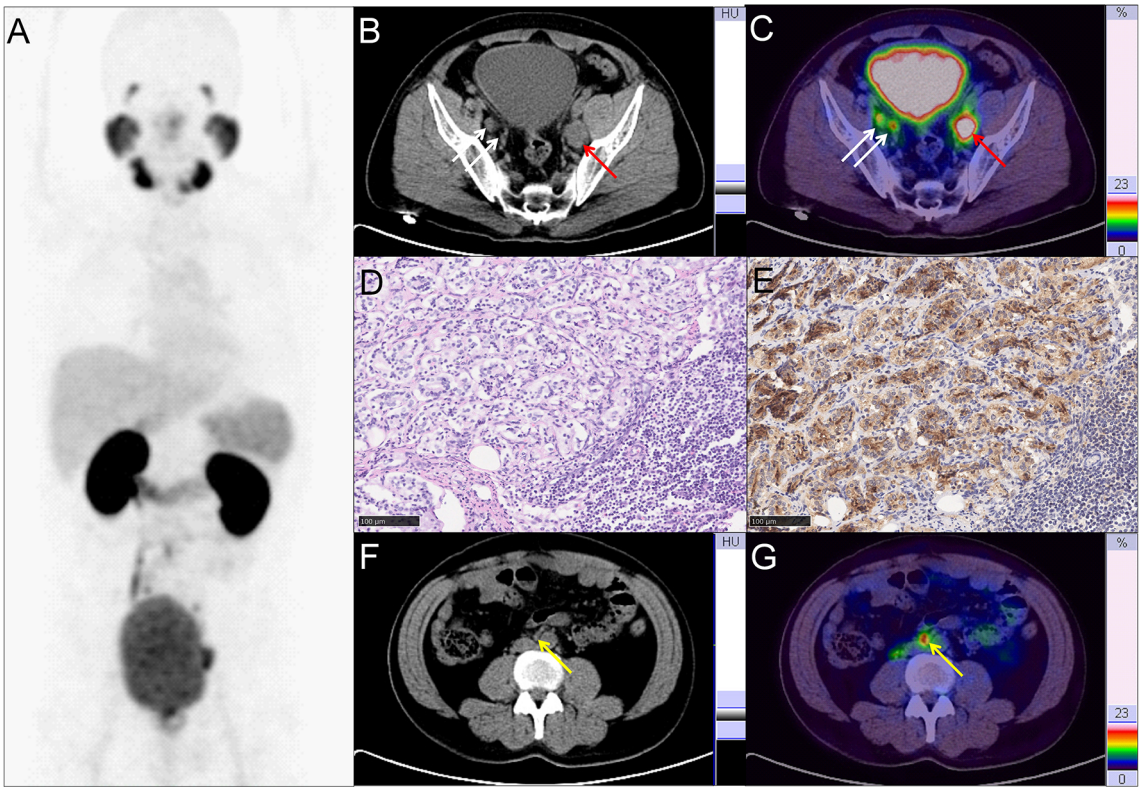


FIGURE 2 | Representative PC patient with non-regional oligometastatic lymph nodes. **(A–C, F, G)** ^{68}Ga-PSMA PET/CT results of a 40 y/o PC patient with LNMs (cT3a, GS 4 + 4 = 8, tPSA 47.99 ng/ml). **(B, C)** Typical para-iliac LNMs (white arrow, right para-iliac LNMs; red arrow, typical left para-iliac LNMs, SUV_{max} 13.2). **(D)** HE staining and **(E)** PSMA staining of the resected left para-iliac LNMs. **(F, G)** non-regional oligometastatic LNMs in paraaortic area (yellow arrow, SUV_{max} 10.9). The LNMs risks were 94.0%, 93.0%, and 89.0% according to the Briganti, MSKCC and Winter nomograms, respectively. The LNMs risk (>65%) indicated that the PC patients may have non-regional LNMs. After ^{68}Ga-PSMA PET/CT, radiotherapy was also performed on the oligometastatic paraaortic LNM after RP with ePLND.

TABLE 2 | Overview of distant lymph nodes metastases.

Location	Σ
Total	71 (100.0%)
left supraclavicular lymph nodes	3(4.2%)
right supraclavicular lymph nodes	2(2.8%)
left infraclavicular lymph nodes	4(5.6%)
right infraclavicular lymph nodes	3(4.2%)
left cervical lymph nodes	1(1.4%)
right cervical lymph nodes	1(1.4%)
paraesophageal lymph nodes	2(2.8%)
left axillary lymph nodes	1(1.4%)
right axillary lymph nodes	1(1.4%)
left hilar lymph nodes	4(5.6%)
right hilar lymph nodes	4(5.6%)
left lung	1(1.4%)
mediastinal lymph nodes	6(8.5%)
left paradiaphragmatic lymph nodes	1(1.4%)
right paradiaphragmatic lymph nodes	2(2.8%)
mesenteric lymph nodes	1(1.4%)
Retroperitoneal lymph nodes	34(47.9%)

Thus, clinical nomograms indicated that most of the patients in this study need to receive ePLND to deal with regional LNMs or potential regional LNMs. The PC patients with distant LNMs has higher scores than those with regional LNMs. Then, we tried to clarify the non-regional LNMs by ^{68}Ga-PSMA PET/CT.

The Tangible Benefit of ^{68}Ga-PSMA PET/CT in Nomogram-Based Therapy Choices of High-Risk Prostate Cancer Patients With Non-Regional Lymph Node Metastases

Next, we analyzed whether an additional ^{68}Ga-PSMA PET/CT have potential benefit on nomogram-based therapy choices of high-risk PC patients, including sparing unnecessary ePLND in PC patients with non-regional LNMs. Four examples were used to show the potential benefit on clinical decision-making, especially in those with non-regional LNMs.

Therapy modality changes were made in the PC patients with non-regional oligometastatic lymph nodes. As shown in **Figure 2**, a 40-year-old patient (T3a stage) with tPSA 48.0 ng/ml and GS 4 + 4 = 8 with pelvic LNMs on pelvic mpMRI and no BMs on bone scan. The LNMs risks were 94.0%, 93.0%, and 89.0% according to the Briganti, MSKCC and Winter

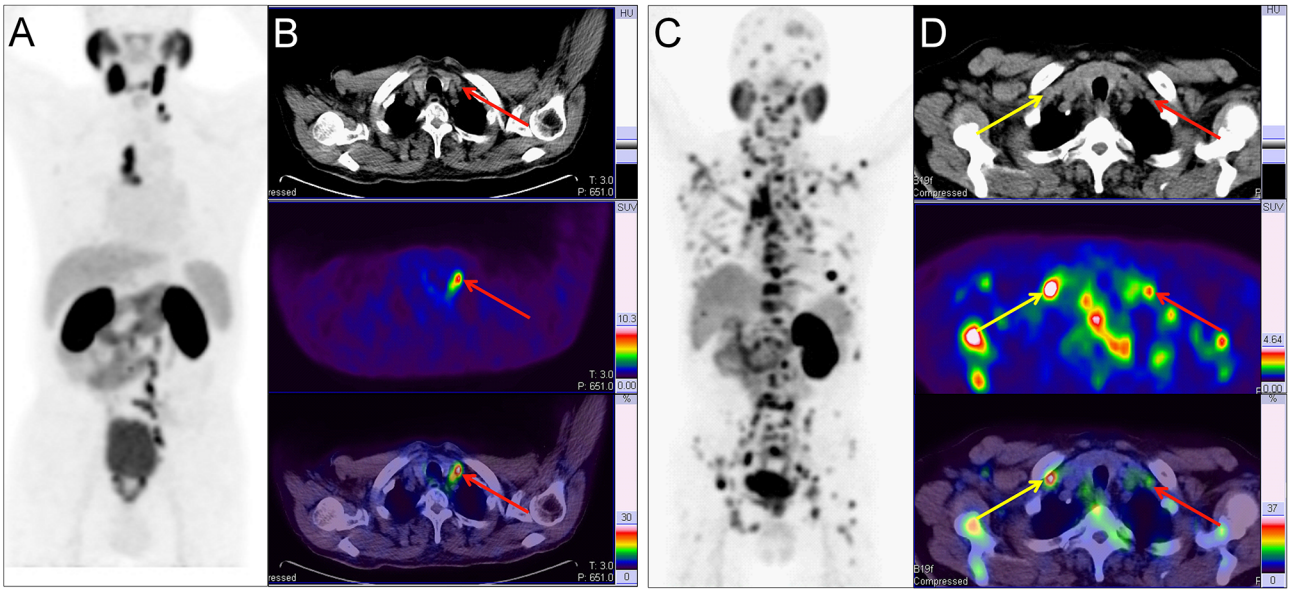


FIGURE 3 | Two example PC patients with Virchow nodes. **(A, B)** ^{68}Ga-PSMA PET/CT results of a 76 y/o patient (T2cN1M1a, GS 4 + 4 = 8, tPSA 150.9 ng/ml). Virchow nodes (red arrow, SUV_{max} 8.42). After ^{68}Ga-PSMA PET/CT, LNMs were shown in the Virchow nodes, mediastinal lymph nodes, and retroperitoneal lymph nodes so RP with ePLND could be replaced by RT. **(C, D)** ^{68}Ga-PSMA PET/CT of a 73 y/o patient (pT4N1M1b; GS 5 + 4 = 9; tPSA 747.9 ng/ml). right supraclavicular LNMs (yellow arrow, SUV_{max} 7.89) and Virchow nodes (red arrow, SUV_{max} 2.98). In the first patient, the LNMs risks were 81.0%, 99.0%, and 70.0% according to the Briganti, MSKCC and Winter nomograms, respectively. The LNMs risk (>65%) indicated that the PC patients may have non-regional LNMs. After ^{68}Ga-PSMA PET/CT, the radiation scope of radiotherapy could be modified for more accurate location.

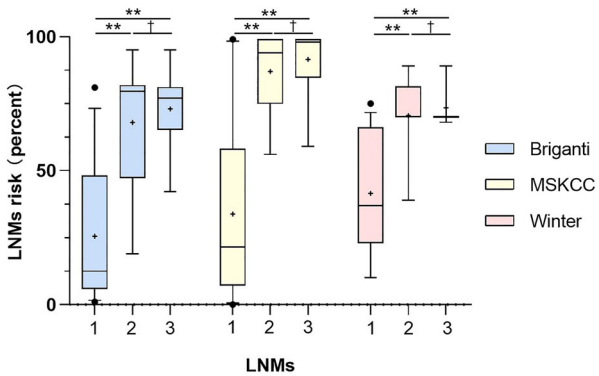


FIGURE 4 | Comparison of risks from the three nomograms between the patients with no LNMs (1), the patients with regional LNMs (2) and the patients with distant LNMs (3). Vertical borders of box represent 25th and 75th percentiles and middle bar represents median while "+" represents mean. Mann-Whitney U tests showed significant difference between the patients with LNMs and without LNMs (**P<0.001) and the PC patients with non-regional LNMs have higher scores than those with regional LNMs. The LNMs risks were calculated from Briganti, MSKCC and Winter nomograms.

nomograms, respectively. As a result, the patient underwent RP and extended PLND (ePLND) according to standard guidelines (7, 9). However, ^{68}Ga-PSMA PET/CT revealed that non-regional oligometastatic lymph nodes existed in the paraaortic area. Therefore, after surgery, radiotherapy (RT) should be performed on the paraaortic LNMs.

As another example, changes in therapeutic modality may also occur in PC patients based on the detection of non-regional LNMs. As shown in **Figure 3A**, a 76-year-old patient (T2c stage) with tPSA 150.9 ng/ml and GS 4 + 4 = 8 with regional LNMs on pelvic mpMRI and no BMs on bone scan. The LNMs risks were 81.0%, 99.0%, and 70.0% according to the Briganti, MSKCC and Winter nomograms, respectively. As a result, the patient received RP and ePLND according to standard guidelines (7, 9). However, ^{68}Ga-PSMA PET/CT revealed non-regional LNMs in the Virchow nodes, mediastinal lymph nodes, and retroperitoneal lymph nodes. Hence, RP and ePLND should be replaced by image-guided radiotherapy (IGRT). ^{68}Ga-PSMA PET/CT has the potential to reduce unnecessary ePLND in this kind of PC patients.

In addition, changes in RT scope may be recommended for PC patients with many non-regional LNMs. As shown in **Figure 3B**, a 73-year-old patient (T4 stage) with tPSA 747.9 ng/ml and GS 4 + 5 = 9 with regional LNMs on pelvic mpMRI and BMs on bone scan. However, all non-regional LNMs, such as Virchow nodes, were missed by pelvic mpMRI. With ^{68}Ga-PSMA PET/CT, non-regional LNMs were clearly shown for accurate modification of the radiation area of IGRT.

Furthermore, ^{68}Ga-PSMA PET/CT has the potential to reduce unnecessary ePLND in PC patients without LNMs. As shown in **Figure S1**, a 71-year-old patient (T2a stage) with tPSA

TABLE 3 | Risk of lymph nodes metastases according to three clinical nomograms.

	Briganti nomogram	MSKCC nomogram	Winter nomogram
Patient number	57	57	57
Median (range) [%]	48.0 (1.0-95.0)	63.0 (0.0-99.0)	70.0 (10.0-89.0)
Subgroups			
≤10.0%	13(22.8%)	10(17.5%)	2(3.5%)
10.1-20.0%	7(12.3%)	4(7.0%)	1(1.8%)
20.1-30.0%	3(5.3%)	5(8.8%)	10(17.5%)
30.1-40.0%	1(1.8%)	1(1.8%)	8(14.0%)
40.1-50.0%	8(14.0%)	2(3.5%)	1(1.8%)
50.1-60.0%	4(7.0%)	6(10.5%)	1(1.8%)
60.1-70.0%	6(10.5%)	4(7.0%)	30(52.6%)
70.1-80.0%	3(5.3%)	2(3.5%)	1(1.8%)
80.1-90.0%	10(17.5%)	5(8.8%)	3(5.3%)
>90.0%	2(3.5%)	18(31.6%)	0 (0.0%)
Unknown	0 (0.0%)	0 (0.0%)	0 (0.0%)

of 9.3ng/ml and GS 4 + 4 = 8 with no pelvic LNMs on pelvic mpMRI and no BMs on bone scan. The LNMs risks were 11.0%, 21.0%, and 39.0% according to the Briganti, MSKCC and Winter nomograms, respectively. Therefore, the patient underwent RP and ePLND to treat the potential LNMs according to the guidelines (7, 9). However, no LNMs were observed on ^{68}Ga-PSMA PET/CT. After ePLND, the pathological results from resected lymph nodes confirmed that no LNMs existed, which were in agreement with the PSMA PET/CT findings.

In general, the new information of ^{68}Ga-PSMA PET/CT can lead to changes of therapy choices in 70.2% (40/57) of the PC patients when using the Briganti nomogram in clinical decision-making (**Figure 5A**). Therapy modality changes could be made in 21.1% of the patients and other changes include details of RT scope (17.5%) or type and extent of surgery (31.6%). Changes in therapeutic treatment modality were due to newly detected non-regional LNMs in 12.3% of the patients and newly detected BMs in 8.7% of the patients. ^{68}Ga-PSMA PET/CT has the potential to reduce unnecessary ePLND in 12.3% of the patients because of newly detected non-regional LNMs. Newly detected non-regional LNMs can also lead to modification of RT in 17.5% of the patients. Similarly, 75.4% (43/57) of the patients would have had different clinical decision-making as compared to the MSKCC nomogram (**Figure 5B**) and 73.7% (42/57) of the patients would have had different clinical decision-making as compared to the Winter nomogram (**Figure 5C**) after ^{68}Ga-PSMA PET/CT. Therefore, ^{68}Ga-PSMA PET/CT can provide added benefit to the three standard nomograms and affect clinical decision-making, especially for patients with non-regional LNMs.

The Establishment of New Thresholds for PLND-Validated Clinical Nomograms to Predict Non-Regional Lymph Node Metastases

According to the guidelines, the PC patients with LNMs risk higher than 5% in Briganti nomogram, 2% in MSKCC nomogram or 7% in Winter nomogram need receive ePLND to deal with potential regional LNMs (5–7). However, whether the PLND-validated clinical nomograms have the potential to evaluate the risk of non-regional LNMs remains unclear. The PC patients with LNMs risk higher than the above mentioned cutoff values, such as 5%, may also have non-regional LNMs and these patients with non-regional LNMs should not receive ePLND. Next, we analyzed whether the clinical nomograms have the potential to predict non-regional LNMs.

To determine whether the clinical scores of nomograms can predict PSMA PET positive non-regional LNMs by setting a higher cutoff value, we plotted the ROC curves of the three nomograms to compare the accuracy of nomograms in predicting PSMA PET positive non-regional LNMs; for each nomogram, a cutoff value corresponding to highest level accuracy was utilized. The AUC of the clinical scores to predict PSMA PET positive non-regional LNMs is shown in **Figure 6**. We found that the PC patients with a score >64% in Briganti nomogram, a score >75% in MSKCC nomogram and a score >67% in Winter nomogram were more likely to have non-regional LNMs. We found the AUC of MSKCC and Briganti nomograms was slightly higher than those of Winter nomogram. The AUC of clinical nomograms was shown in **Table 4**. Using the above higher cutoff values, 10.5% (6/57) high-risk PC patients with non-regional LNMs can be excluded by Briganti and Winter nomograms, and 8.7% (5/57) high-risk PC patients with non-regional LNMs can be excluded by MSKCC nomograms (**Figure 7**).

Conclusively, the ePLND-based clinical nomograms have the potential to predict non-regional LNMs as well. The surgeons should pay close attention to the PC patients with LNMs risk higher than approximately 65% for potential non-regional LNMs.

DISCUSSION

This study depicted the pattern of metastatic spread by ^{68}Ga-PSMA PET/CT and demonstrated the potential benefit of this imaging modality on clinical decision-making for high-risk PC patients, especially in patients with non-regional LNMs. ^{68}Ga-PSMA PET/CT can provide pattern of metastatic spread of LNMs with a higher sensitivity and specificity compared with conventional morphological imaging (34). One previous study indicated that, compared with mpMRI, more LNMs can be observed by ^{68}Ga-PSMA PET/CT in 27.8% of PC patients (35). In this study, 96.7% of the patients with LNMs who received ^{68}Ga-PSMA PET/CT had detectable LNMs, while pelvic mpMRI could only detect regional LNMs in 60.7% of patients. Therefore, pelvic mpMRI alone has limited sensitivity for detection of regional LNMs and cannot include non-regional LNMs.

In our study, 34.5% of the LNMs were detected in non-regional sites. This finding was comparable to prior work, including a study of 280 treatment-naïve PC patients that reported the use of ^{68}Ga-PSMA PET/CT detected 36.0% of LNMs as non-regional and 15.5% of LNMs above the diaphragm (8). The most commonly observed non-regional

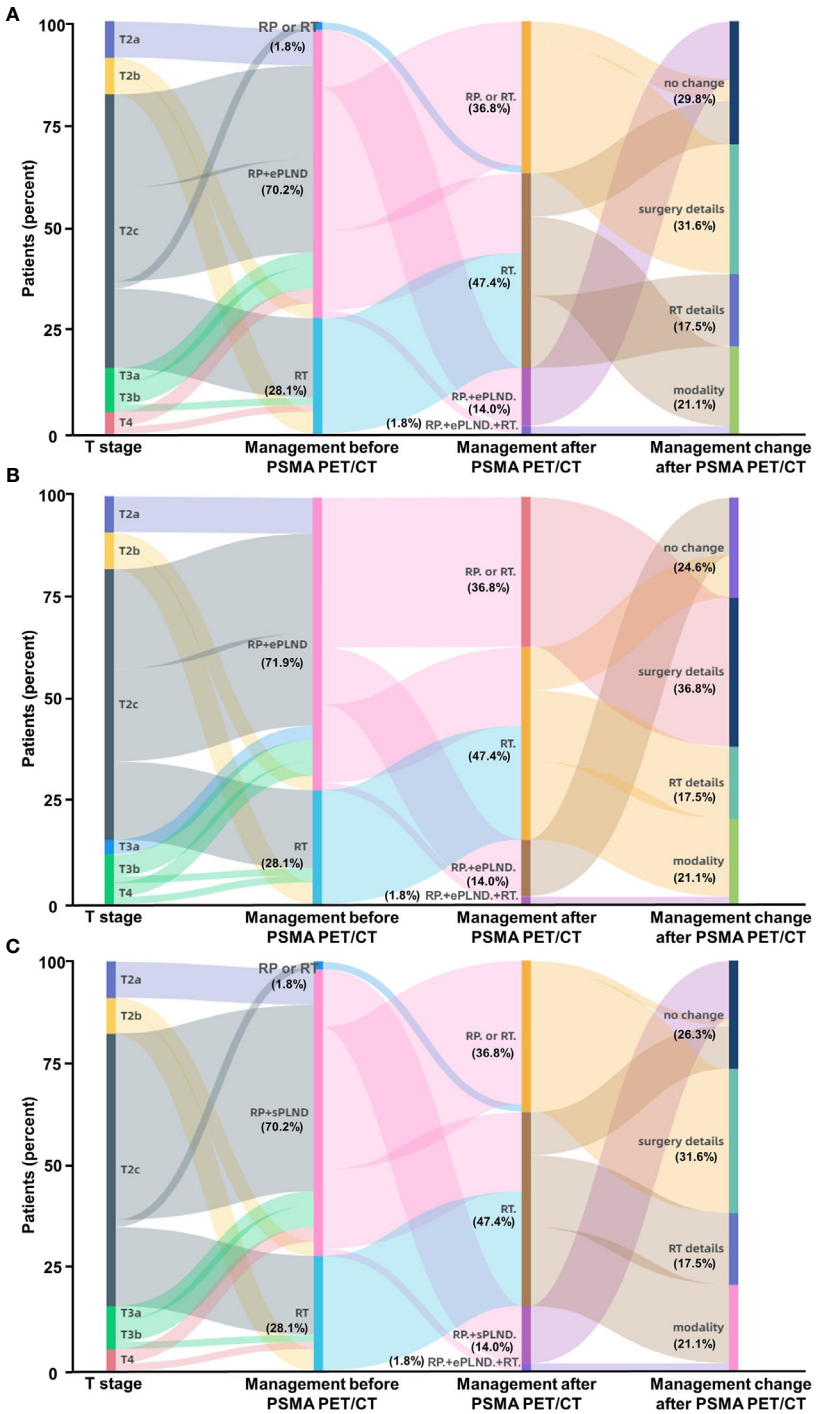


FIGURE 5 | Impact of ^{68}Ga-PSMA PET/CT on clinical decision-making in high-risk treatment-naïve prostate cancer patients. **(A)** According to the Briganti nomogram, ^{68}Ga-PSMA PET/CT would have led changes of planned therapy in 70.2% of these patients and avoided unnecessary ePLND in 12.3% of the patients who have previously undetected non-regional LNMs. **(B)** According to the MSKCC nomogram, ^{68}Ga-PSMA PET/CT would have led changes of planned therapy in 75.4% of the patients; **(C)** According to the Winter nomogram, ^{68}Ga-PSMA PET/CT would have led changes of planned therapy in 73.7% of the patients. As the basic treatment for nearly all high-risk PC patients, ADT or systemic therapy was not included. (Modality: changes of therapy modality; surgery details: changes in surgery details; RT details: change in RT field; RP: radical prostatectomy; RT: radiotherapy; ePLND: extend PLND; sPLND: sentinel PLND).

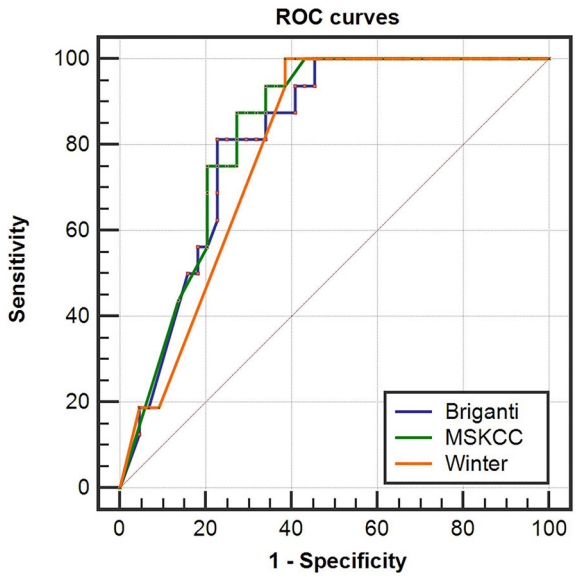


FIGURE 6 | Receiver operating characteristic (ROC) curves of the Briganti, MSKCC and Winter nomograms for predicting non-regional PSMA PET positive LNMs in high-risk PC patients. The AUCs of the clinical nomograms (Briganti, MSKCC and Winter) in predicting non-regional LNMs were 0.816, 0.830 and 0.793, respectively. The PC patients with a score >64% in Briganti nomogram, a score >75% in MSKCC nomogram and a score >67% in Winter nomogram were more likely to have non-regional LNMs. The above cutoff values can be used to predict non-regional LNMs in high-risk PC patients.

TABLE 4 | Distant PSMA PET positive LNMs of ROC analyses by three clinical nomograms.

Stage (number)	Characteristics	AUC	SE	95% CI
T1-T3 (n=57)	MSKCC nomogram	0.830	0.051	0.710-0.914
	Briganti nomogram	0.816	0.054	0.695-0.904
	Winter nomogram	0.793	0.05	0.668-0.886

LNMs were the paraaortic lymph nodes (12.8%), while the most commonly encountered non-regional LNMs above the diaphragm were the mediastinal lymph nodes (6.2%) (8). Our results were consistent with the prior study because 97.8% of the patients in this work were intermediate- to high-risk PC patients (8). In another study, ^{68}Ga-PSMA PET/CT revealed that 16.0% of the treatment-naïve PC patients had LNMs and 6.0% of these patients had non-regional LNMs (36). Meanwhile, Nurhan et al. reported that 9.0% of the treatment-naïve PC patients had non-regional LNMs (37). For the intermediate- to high-risk treatment-naïve PC patients, one study indicated that 23.3% of the patients had regional LNMs and 9.5% of them had non-regional LNMs (38). One previous study showed that retroperitoneal LNMs were encountered in 12.8% of the high-risk PC patients by ^{68}Ga-PSMA I&T PET/CT (39). In our study, the retroperitoneal LNMs were the most commonly visualized non-regional LNMs. We identified 57.1% compared to 47.7% (in BJU study) non-regional LNMs of the PC patients with LNMs (40). Compared with previous studies, more non-regional LNMs were observed in our study because more high-risk PC patients with higher grade of PC were included in our cohort. Furthermore, we focused on that a threshold can be set for current clinical PLND-validated nomograms to predict extrareolar LN metastases with an AUC accuracy of about 80% after optimizing the simple nomograms which may help to improve the efficiency for PC therapy significantly in clinical practice. PSMA expression, which could be reflected by Maximum Standardized Uptake Values (SUV_{max}) of the primary tumor in ^{68}Ga-PSMA PET/CT, was closely correlated with higher grades of PC (41). As shown in **Figure S3**, a significantly higher SUV_{max} was found in the patients with LNMs versus those without LNMs (P<0.01, Mann-Whitney-U-test). The Virchow nodes were observed in our study (**Figure 3**). One previous study also reported that 3.2% of the treatment-naïve patients have LNMs in the Virchow nodes (8). The study showed that the Virchow nodes were significantly more frequent with a GS ≥8 as compared to a GS ≤7b primary PC tumor (8, 42). In our study, all PC patients with Virchow nodes had a GS ≥8.

The new information obtained *via* ^{68}Ga-PSMA PET/CT can benefit clinical decision-making and provide additive benefit to existing PLND-validated nomograms, particularly for individuals with non-regional LNMs. For treatment-naïve PC patients, the most frequent new findings were LNMs (17.2%) and ^{68}Ga-PSMA PET/CT can impact therapeutic decision making in 27.6% of patients (38). In our study, based on the ^{68}Ga-PSMA PET/CT findings, the new information could have resulted in a change in clinical management in more than two-thirds of patients. In our study, the newly detected non-regional LNMs by ^{68}Ga-PSMA PET/CT can lead to therapy modality change in 12.3% of the patients and modification of RT in 17.5% of the patients. In previous work, treatment-naïve PC patients and recurrent PC patients were not studied separately and most prior studies focused on RT management exclusively. Florian et al. reported 27.5% of the PC patients had non-regional LNMs, and the new information from ^{68}Ga-PSMA PET/CT led to radiotherapeutic management in 50.8% of the cases (43). In this study including 26.0% treatment-naïve patients, the new information of ^{68}Ga-PSMA PET/CT led to changes of planned RT in 26.4% of the patients (43). In another study including 48.1% treatment-naïve patients, ^{68}Ga-PSMA PET/CT changed RT in 46.3% of the cases, and changed hormone therapy in 33.3% of the patients, with an overall change in decision-making in 53.7% of the patients (44). In an Australian prospective multi-center study, ^{68}Ga-PSMA PET/CT led to changes of management intent in 21.3% of the treatment-naïve patients while 61.5% of the patients with biochemical recurrence had changes in planned therapy (45). More metastatic lesions were discovered in high-risk PC patients to facilitate RT planning by mapping PSMA-avid lesions (46). One study indicated that ^{68}Ga-PSMA PET/CT changed TNM stage and RT in 26.0% and 44.0% of the 50 treatment-naïve PC patients, respectively (36). In our study, ^{68}Ga-PSMA PET/CT also revealed the potential to reduce ePLND in 31.6% of the patients. Similarly, one previous study demonstrated that ^{68}Ga-PSMA PET imaging had the potential to facilitate patient

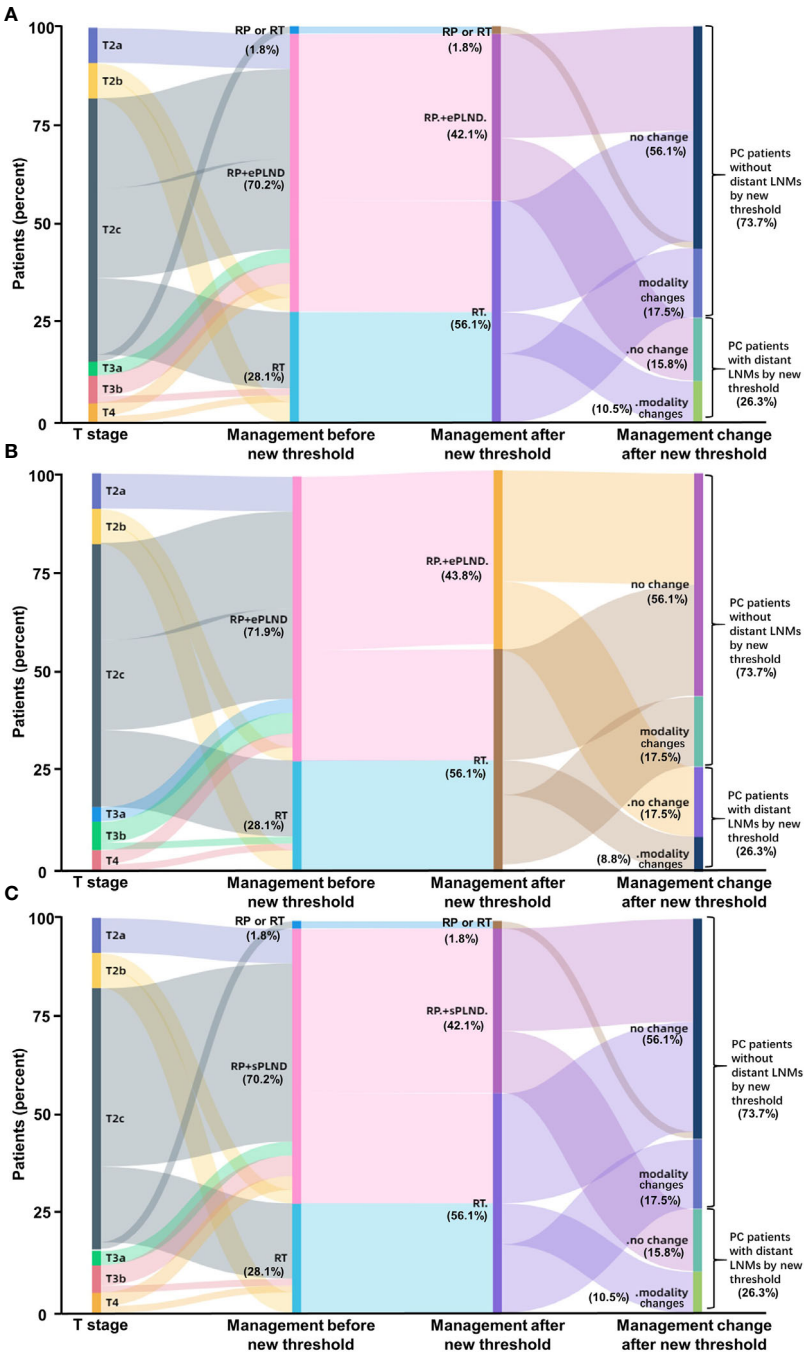


FIGURE 7 | Impact of new thresholds for PLND-validated clinical nomograms on clinical decision-making in high-risk treatment-naïve prostate cancer patients (**A**: Briganti nomogram; **B**: MSKCC nomogram; **C**: Winter nomogram). The new thresholds in Briganti **(A)** and Winter **(C)** nomograms excluded 93.75% (15/16) of the PC patients with non-regional LNMs and 87.5% (14/16) of the patients for MSKCC **(B)** nomogram. (Modality, changes of therapy modality; RP, radical prostatectomy; RT, radiotherapy; ePLND, extend PLND; sPLND, sentinel PLND).

selection for ePLND in intermediate- to high-risk PC patients, and these authors created a novel PLND-validated nomogram including tPSA, GS, and PSMA positive volume ($PSMA_{total}$) of ^{68}Ga-PSMA PET/CT (47). The therapy modality changes and therapy detail changes were modified from one previous study (38).

In the current study, our study demonstrated that the PLND-validated clinical nomograms (Briganti, MSKCC, Winter) have the potential to predict non-regional LNMs with a higher cutoff value of scores, although the three clinical nomograms are widely used to predict the risks of regional LNMs in previous studies

(5–7). The PC patients with LNMs risk higher than 65% need to receive PSMA PET/CT to exclude non-regional LNMs to reduce unnecessary ePLND.

Our study had some limitations. The first limitation in this study is that ^{68}Ga-PSMA PET/CT may have limitation on detecting LNMs in the prostate cancer patients with negative PSMA expression (like neuroendocrine prostate cancer), although expression of PSMA has been recognized in approximately 95% of prostate cancer, both primary and metastatic (10, 41). Another limitation included small sample size of patients because we focused on the high-risk prostate cancer patients who had non-regional LNMs in this project, which will be helpful for clinical decision making.

In conclusion, our study demonstrated that the clinical nomograms have the potential to predict non-regional LNMs. Our findings also demonstrate the proportion of non-regional metastases in the initial staging by ^{68}Ga-PSMA PET/CT and demonstrate how this information may impact clinical decision-making in high-risk treatment-naïve PC patients, especially in those with non-regional LNMs. In comparison to the current standards for PLND, our study revealed that non-regional LNMs can be observed in more than one-third of patients, and ^{68}Ga-PSMA PET/CT has the potential to add additional information to existing nomograms-based clinical decision-making in more than two-thirds of the high-risk PC patients. We focused on that a threshold can be set for current clinical PLND-validated nomograms to predict extrareolar LN metastases with an AUC accuracy of about 80% after optimizing the simple nomograms which may help to improve the efficiency for PC therapy significantly in clinical practice. If validated in a larger prospective study, ^{68}Ga-PSMA PET/CT, as well as clinical nomograms, can be used to exclude patients with non-regional LNMs before PLND with the potential to more accurately identify the appropriate treatment modality for patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because N/A. Requests to access the datasets should be directed to qinwj@fmmu.edu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Fourth Military Medical University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This study is funded by the National Natural Science Foundation of China (grant nos. 81772734, 81971646, 91959208, 81871379 and 81372748), Lynn Sage cancer research OncoSET program, Innovation Capability Support Program of Shaanxi (grant nos. 2020PT-021 and 2021TD-39) and Key project of Shaanxi Natural Science Basic Research Program (grant no. 2021JZ-25).

ACKNOWLEDGMENTS

The authors thank Mrs. Honglei Li (HL) and Yingmei Wang (YW) (Department of Pathology, Xijing Hospital, Fourth Military Medical University, China) for their detailed pathological diagnosis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.658669/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2020. *CA Cancer J Clin* (2020) 70(1):7–30. doi: 10.3322/caac.21590
2. Harisinghani MG, Barentsz J, Hahn PF, Deserno WM, Tabatabaei S, van de Kaa CH, et al. Noninvasive detection of clinically occult lymph-node metastases in prostate cancer. *N Engl J Med* (2003) 348(25):2491–9. doi: 10.1056/NEJMoa022749
3. Eden CG, Zacharakis E, Bott S. The learning curve for laparoscopic extended pelvic lymphadenectomy for intermediate- and high-risk prostate cancer: implications for compliance with existing guidelines. *BJU Int* (2013) 112(3):346–54. doi: 10.1111/j.1464-410X.2012.11671.x
4. Hovels AM, Heesakkers RA, Adang EM, Jager GJ, Strum S, Hoogeveen YL, et al. The diagnostic accuracy of CT and MRI in the staging of pelvic lymph nodes in patients with prostate cancer: a meta-analysis. *Clin Radiol* (2008) 63(4):387–95. doi: 10.1016/j.crad.2007.05.022
5. Winter A, Kneib T, Rohde M, Henke RP, Wawroschek F. First Nomogram Predicting the Probability of Lymph Node Involvement in Prostate Cancer Patients Undergoing Radioisotope Guided Sentinel Lymph Node Dissection. *Urol Int* (2015) 95(4):422–8. doi: 10.1159/000431182
6. Briganti A, Larcher A, Abdollah F, Capitanio U, Gallina A, Suardi N, et al. Updated nomogram predicting lymph node invasion in patients with prostate cancer undergoing extended pelvic lymph node dissection: the essential importance of percentage of positive cores. *Eur Urol* (2012) 61(3):480–7. doi: 10.1016/j.eururo.2011.10.044
7. Mohler JL, Antonarakis ES, Armstrong AJ, D'Amico AV, Davis BJ, Dorff T, et al. Prostate Cancer, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* (2019) 17(5):479–505. doi: 10.6004/jnccn.2019.0023

8. Koerber SA, Stach G, Kratochwil C, Haefner MF, Rathke H, Herfarth K, et al. Lymph node involvement in treatment-naive prostate cancer patients - correlation of PSMA-PET/CT imaging and Roach formula in 280 men in the Radiotherapeutic management. *J Nucl Med: official publication, Society of Nuclear Medicine* (2020) 61(1): 46–50. doi: 10.2967/jnumed.119.227637
9. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* (2017) 71 (4):618–29. doi: 10.1016/j.eururo.2016.08.003
10. Sweat SD, Pacelli A, Murphy GP, Bostwick DG. Prostate-specific membrane antigen expression is greatest in prostate adenocarcinoma and lymph node metastases. *Urology* (1998) 52(4):637–40. doi: 10.1016/s0090-4295(98)00278-7
11. Wu H, Xu T, Wang X, Yu YB, Fan ZY, Li DX, et al. Diagnostic Performance of (6)(8)Gallium Labelled Prostate-Specific Membrane Antigen Positron Emission Tomography/Computed Tomography and Magnetic Resonance Imaging for Staging the Prostate Cancer with Intermediate or High Risk Prior to Radical Prostatectomy: A Systematic Review and Meta-analysis. *World J Mens Health* (2019) 22(1): 6–9. doi: 10.5534/wjmh.180124
12. Koschel S, Murphy DG, Hofman MS, Wong LM. The role of prostate-specific membrane antigen PET/computed tomography in primary staging of prostate cancer. *Curr Opin Urol* (2019) 29(6):569–77. doi: 10.1097/MOU.0000000000000677
13. Luiting HB, van Leeuwen PJ, Busstra MB, Brabander T, van der Poel HG, Donswijk ML, et al. Use of (68) Ga-PSMA PET for detecting lymph node metastases in primary and recurrent prostate cancer and location of recurrence after radical prostatectomy: an overview of the current literature. *BJU Int* (2019) 125(2): 206–14. doi: 10.1111/bju.14944
14. Perera M, Papa N, Christidis D, Wetherell D, Hofman MS, Murphy DG, et al. Sensitivity, Specificity, and Predictors of Positive (68)Ga-Prostate-specific Membrane Antigen Positron Emission Tomography in Advanced Prostate Cancer: A Systematic Review and Meta-analysis. *Eur Urol* (2016) 70(6):926–37. doi: 10.1016/j.eururo.2016.06.021
15. Hijazi S, Meller B, Leitsmann C, Strauss A, Meller J, Ritter CO, et al. Pelvic lymph node dissection for nodal oligometastatic prostate cancer detected by 68Ga-PSMA-positron emission tomography/computerized tomography. *Prostate* (2015) 75(16):1934–40. doi: 10.1002/pros.23091
16. Obek C, Doganca T, Demirci E, Ocak M, Kural AR, Yildirim A, et al. The accuracy of (68)Ga-PSMA PET/CT in primary lymph node staging in high-risk prostate cancer. *Eur J Nucl Med Mol Imaging* (2017) 44(11):1806–12. doi: 10.1007/s00259-017-3752-y
17. Maurer T, Gschwend JE, Rauscher I, Souvatzoglou M, Haller B, Weirich G, et al. Diagnostic Efficacy of (68)Gallium-PSMA Positron Emission Tomography Compared to Conventional Imaging for Lymph Node Staging of 130 Consecutive Patients with Intermediate to High Risk Prostate Cancer. *J Urol* (2016) 195(5):1436–43. doi: 10.1016/j.juro.2015.12.025
18. Maurer T, Eiber M, Schwaiger M, Gschwend JE. Current use of PSMA-PET in prostate cancer management. *Nat Rev Urol* (2016) 13(4):226–35. doi: 10.1038/nrurol.2016.26
19. Perera M, Papa N, Roberts M, Williams M, Udovicich C, Vela I, et al. Gallium-68 Prostate-specific Membrane Antigen Positron Emission Tomography in Advanced Prostate Cancer-Updated Diagnostic Utility, Sensitivity, Specificity, and Distribution of Prostate-specific Membrane Antigen-avid Lesions: A Systematic Review and Meta-analysis. *Eur Urol* (2020) 77(4):403–17. doi: 10.1016/j.eururo.2019.01.049
20. Maurer T, Graefen M, van der Poel H, Hamdy F, Briganti A, Eiber M, et al. Prostate-Specific Membrane Antigen-Guided Surgery. *J Nucl Med Off publication Soc Nucl Med* (2020) 61(1):6–12. doi: 10.2967/jnumed.119.232330
21. van Leeuwen FWB, Winter A, van Der Poel HG, Eiber M, Suardi N, Graefen M, et al. Technologies for image-guided surgery for managing lymphatic metastases in prostate cancer. *Nat Rev Urol* (2019) 16(3):159–71. doi: 10.1038/s41585-018-0140-8
22. D'Amico AV, Desjardin A, Chung A, Chen MH. Assessment of outcome prediction models for localized prostate cancer in patients managed with external beam radiation therapy. *Semin Urol Oncol* (1998) 16(3):153–9.
23. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* (2010) 17(6):1471–4. doi: 10.1245/s10434-010-0985-4
24. Zhang J, Shao S, Wu P, Liu D, Yang B, Han D, et al. Diagnostic performance of (68)Ga-PSMA PET/CT in the detection of prostate cancer prior to initial biopsy: comparison with cancer-predicting nomograms. *Eur J Nucl Med Mol Imaging* (2019) 46(4):908–20. doi: 10.1007/s00259-018-4255-1
25. Jiao D, Li Y, Yang F, Han D, Wu J, Shi S, et al. Expression of Prostate-Specific Membrane Antigen in Tumor-Associated Vasculature Predicts Poor Prognosis in Hepatocellular Carcinoma. *Clin Transl Gastroenterol* (2019) 10 (5):1–7. doi: 10.14309/ctg.0000000000000041
26. Shi SJ, Wang LJ, Han DH, Wu JH, Jiao D, Zhang KL, et al. Therapeutic effects of human monoclonal PSMA antibody-mediated TRIM24 siRNA delivery in PSMA-positive castration-resistant prostate cancer. *Theranostics* (2019) 9 (5):1247–63. doi: 10.7150/thno.29884
27. Li Y, Han D, Wu P, Ren J, Ma S, Zhang J, et al. Comparison of (68)Ga-PSMA-617 PET/CT with mpMRI for the detection of PCa in patients with a PSA level of 4-20 ng/ml before the initial biopsy. *Sci Rep* (2020) 10(1):10963. doi: 10.1038/s41598-020-67385-9
28. Zhang Q, Zang S, Zhang C, Fu Y, Lv X, Zhang Q, et al. Comparison of (68)Ga-PSMA-11 PET-CT with mpMRI for preoperative lymph node staging in patients with intermediate to high-risk prostate cancer. *J Transl Med* (2017) 15 (1):230. doi: 10.1186/s12967-017-1333-2
29. Jilg CA, Drendel V, Rischke HC, Beck T, Vach W, Schaal K, et al. Diagnostic Accuracy of Ga-68-HBED-CC-PSMA-Ligand-PET/CT before Salvage Lymph Node Dissection for Recurrent Prostate Cancer. *Theranostics* (2017) 7 (6):1770–80. doi: 10.7150/thno.18421
30. Vinsensia M, Chyoke PL, Hadaschik B, Holland-Letz T, Moltz J, Kopka K, et al. (68)Ga-PSMA PET/CT and Volumetric Morphology of PET-Positive Lymph Nodes Stratified by Tumor Differentiation of Prostate Cancer. *J Nucl Med Off publication Soc Nucl Med* (2017) 58(12):1949–55. doi: 10.2967/jnumed.116.185033
31. Fendler WP, Eiber M, Beheshti M, Bomanji J, Ceci F, Cho S, et al. 68Ga-PSMA PET/CT: Joint EANM and SNMMI procedure guideline for prostate cancer imaging: version 1.0. *Eur J Nucl Med Mol Imaging* (2017) 44(6):1014–24. doi: 10.1007/s00259-017-3670-z
32. Rauscher I, Maurer T, Beer AJ, Graner FP, Haller B, Weirich G, et al. Value of 68Ga-PSMA HBED-CC PET for the Assessment of Lymph Node Metastases in Prostate Cancer Patients with Biochemical Recurrence: Comparison with Histopathology After Salvage Lymphadenectomy. *J Nucl Med Off publication Soc Nucl Med* (2016) 57(11):1713–9. doi: 10.2967/jnumed.116.173492
33. Memorial Sloan Kettering Cancer Center. *Dynamic prostate cancer nomogram: coefficients*. Available at: www.mskcc.org/nomograms/prostate/pre-op/coefficients.
34. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multicentre study. *Lancet* (2020) 395 (10231):1208–16. doi: 10.1016/S0140-6736(20)30314-7
35. Tulsyan S, Das CJ, Tripathi M, Seth A, Kumar R, Bal C. Comparison of 68Ga-PSMA PET/CT and multiparametric MRI for staging of high-risk prostate cancer68Ga-PSMA PET and MRI in prostate cancer. *Nucl Med Commun* (2017) 38(12):1094–102. doi: 10.1097/mnm.0000000000000749
36. Koerber SA, Will L, Kratochwil C, Haefner MF, Rathke H, Kremer C, et al. (68)Ga-PSMA-11 PET/CT in Primary and Recurrent Prostate Carcinoma: Implications for Radiotherapeutic Management in 121 Patients. *J Nucl Med: official publication, Society of Nuclear Medicine* (2019) 60(2):234–40. doi: 10.2967/jnumed.118.211086
37. Ergul N, Yilmaz Gunes B, Yucetas U, Toktas MG, Cermik TF. 68Ga-PSMA-11 PET/CT in Newly Diagnosed Prostate Adenocarcinoma. *Clin Nucl Med* (2018) 43(12):e422–e7. doi: 10.1097/RLU.0000000000002289
38. Ferraro DA, Garcia Schuler HI, Muehlematter UJ, Eberli D, Muller J, Muller A, et al. Impact of (68)Ga-PSMA-11 PET staging on clinical decision-making in patients with intermediate or high-risk prostate cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(3):652–64. doi: 10.1007/s00259-019-04568-1
39. Cytawa W, Seitz AK, Kircher S, Fukushima K, Tran-Gia J, Schirbel A, et al. (68)Ga-PSMA I&T PET/CT for primary staging of prostate cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(1):168–77. doi: 10.1007/s00259-019-04524-z
40. Yaxley JW, Raveenthiran S, Nouhaud FX, Samaratunga H, Yaxley WJ, Coughlin G, et al. Risk of metastatic disease on (68) gallium-prostate-specific membrane antigen positron emission tomography/computed

tomography scan for primary staging of 1253 men at the diagnosis of prostate cancer. *BJU Int* (2019) 124(3):401–7. doi: 10.1111/bju.14828
41. Bostwick DG, Pacelli A, Blute M, Roche P, Murphy GP. Prostate specific membrane antigen expression in prostatic intraepithelial neoplasia and adenocarcinoma: a study of 184 cases. *Cancer* (1998) 82(11):2256–61. doi: 10.1002/(sici)1097-0142(19980601)82:11<2256::aid-cncr22>3.0.co;2-s
42. de la Riva-Perez PA, Garcia-Gomez FJ, Bujan-Lloret C, Calvo-Moron MC, Castro-Montano J. Virchow Node From Prostate Carcinoma by 18F-Choline PET/CT. *Clin Nucl Med* (2018) 43(4):271–2. doi: 10.1097/rlu.0000000000001975
43. Sterzing F, Kratochwil C, Fiedler H, Katayama S, Habl G, Kopka K, et al. (68) Ga-PSMA-11 PET/CT: a new technique with high potential for the radiotherapeutic management of prostate cancer patients. *Eur J Nucl Med Mol Imaging* (2016) 43(1):34–41. doi: 10.1007/s00259-015-3188-1
44. Shakespeare TP. Effect of prostate-specific membrane antigen positron emission tomography on the decision-making of radiation oncologists. *Radiat Oncol* (2015) 10:233. doi: 10.1186/s13014-015-0548-8
45. Roach PJ, Francis R, Emmett L, Hsiao E, Kneebone A, Hruby G, et al. The Impact of (68)Ga-PSMA PET/CT on Management Intent in Prostate Cancer: Results of an Australian Prospective Multicenter Study. *J Nucl Med Off publication Soc Nucl Med* (2018) 59(1):82–8. doi: 10.2967/jnumed.117.197160
46. Onal C, Torun N, Akyol F, Guler OC, Hurmuz P, Yildirim BA, et al. Integration of 68Ga-PSMA-PET/CT in Radiotherapy Planning for Prostate Cancer Patients. *Clin Nucl Med* (2019) 44(9):e510–e6. doi: 10.1097/RLU.0000000000002691
47. Ferraro DA, Muehlematter UJ, Garcia Schuler HI, Rupp NJ, Huellner M, Messerli M, et al. (68)Ga-PSMA-11 PET has the potential to improve patient selection for extended pelvic lymph node dissection in intermediate to high-risk prostate cancer. *Eur J Nucl Med Mol Imaging* (2019) 47(1):147–59. doi: 10.1007/s00259-019-04511-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 22 April 2021
doi: 10.3389/fonc.2021.582847

Computed-Tomography-Based Radiomics Model for Predicting the Malignant Potential of Gastrointestinal Stromal Tumors Preoperatively: A Multi-Classifier and Multicenter Study

*Minhong Wang[1], Zhan Feng[2], Lixiang Zhou[3], Liang Zhang[4], Xiaojun Hao[1] and Jian Zhai[1]**

[1] Department of Radiology, The First Affiliated Hospital of Wannan Medical College, Wuhu, China, [2] Department of Radiology, College of Medicine, The First Affiliated Hospital, Zhejiang University, Hangzhou, China, [3] Department of Pharmacy, The First Affiliated Hospital of Wannan Medical College, Wuhu, China, [4] Department of Radiology, Zhejiang Cancer Hospital, Hangzhou, China

Background: Our goal was to establish and verify a radiomics risk grading model for gastrointestinal stromal tumors (GISTs) and to identify the optimal algorithm for risk stratification.

Methods: We conducted a retrospective analysis of 324 patients with GISTs, the presence of which was confirmed by surgical pathology. Patients were treated at three different hospitals. A training cohort of 180 patients was collected from the largest center, while an external validation cohort of 144 patients was collected from the other two centers. To extract radiomics features, regions of interest (ROIs) were outlined layer by layer along the edge of the tumor contour on CT images of the arterial and portal venous phases. The dimensionality of radiomic features was reduced, and the top 10 features with importance value above 5 were selected before modeling. The training cohort used three classifiers [logistic regression, support vector machine (SVM), and random forest] to establish three GIST risk stratification prediction models. The receiver operating characteristic curve (ROC) was used to compare model performance, which was validated by external data.

Results: In the training cohort, the average area under the curve (AUC) was 0.84 ± 0.07 of the logistic regression, 0.88 ± 0.06 of the random forest, and 0.81 ± 0.08 of the SVM. In the external validation cohort, the AUC was 0.85 of the logistic regression, 0.90 of the random forest, and 0.80 of the SVM. The random forest model performed the best in both the training and the external validation cohorts and could be generalized.

Conclusion: Based on CT radiomics, there are multiple machine-learning models that can predict the risk of GISTs. Among them, the random forest algorithm had the highest prediction efficiency and could be readily generalizable. Through external validation data, we assume that the random forest model may be used as an effective tool to guide preoperative clinical decision-making.

Keywords: gastrointestinal stromal tumor, risk classification, radiomics, X-ray computer, multi-classifier

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Neree Payan, Université de Bourgogne, France
Qiang Huang, Zhejiang University, China

****Correspondence:***
Jian Zhai
1181366339@qq.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *13 July 2020*
Accepted: *19 February 2021*
Published: *22 April 2021*

Citation:
Wang M, Feng Z, Zhou L, Zhang L, Hao X and Zhai J (2021) Computed-Tomography-Based Radiomics Model for Predicting the Malignant Potential of Gastrointestinal Stromal Tumors Preoperatively: A Multi-Classifier and Multicenter Study. Front. Oncol. 11:582847. doi: 10.3389/fonc.2021.582847

INTRODUCTION

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal tumors of the digestive system, which occur in the stomach and small intestine. GISTs have a variety of biological characteristics and cannot be simply categorized as benign or malignant (1). For example, some small GISTs can progress rapidly and metastasize to the liver, while some large GISTs, even those not receiving the post-operative adjuvant treatment, present no long-term risk of recurrence or metastasis (2). Therefore, the preoperative evaluation of the malignant potential of GISTs is crucial for treatment decision-making.

Risk stratification is commonly applied to evaluate the biological behaviors and overall clinical outcome of GISTs. Currently, the most recognized criterion is the improved National Institutes of Health risk stratification standard introduced by Joensuu in 2008 (3), which is based on tumor maximum diameter and mitotic count and introduces two parameters: tumor site and tumor rupture. The risk of relapse is thereby divided into four categories: very low risk, low risk, intermediate risk, and high risk. Higher risk generally indicates a worse prognosis. Also, the introduction of imatinib mesylate has greatly changed the outcomes in high-risk GIST patients (4). The need for reliable preoperative risk stratification is of great significance for the development of treatment methods and prognostic evaluation. Most surgeries can completely remove the GISTs without first conducting a preoperative biopsy (5), which may cause tumor ulceration and bleeding, increasing the risk of tumor spread. Therefore, it is of great clinical value to explore non-invasive, reliable, and simple biomarkers for predicting the recurrence and metastasis risk of GISTs before surgery.

Previous GIST risk stratification research is largely based on analysis of computed tomography (CT) images (4, 6–9), which is likely influenced by the observer's subjective assessment. Therefore, an objective and quantitative technique is urgently needed for the accurate risk stratification of GISTs. Radiomics converts medical images into high-dimensional data that can be mined, which holds great potential for application in disease diagnosis, identification, and prognosis predictions (10–13).There are studies have examined the utility of radiomics in GIST risk stratification (14–16) and have achieved favorable results. However, most of these studies are single-center trials, whose prediction models have not been externally verified. Therefore, the generalizability of these models remains unclear. In addition, previous studies used a single classifier for modeling, due to the obvious differences in classifier algorithms (17), and such studies are unable to determine the classifier with the best performance in risk prediction.

In response to these shortcomings, we conducted a multiclassifier and multicenter GISTs radiomics study, applying the three most commonly used machine-learning classifiers in radiomics to the same cohort of data to evaluate and compare the performance of the classifiers. Also, the model was tested with independent external data to further evaluate its generalizability to provide a reference for clinical treatment decisions.

TABLE 1 | The protocols of the CT scan for the patients with GISTs.

Manufacture	Philips	SIEMENS	Philips
CT scanner	Brilliance 64	Dual source CT	Brilliance 256
Tube voltage (kV)	120	120	120
Tube current (mA)	250	200	250
Rotation time (s)	0.4	0.5	0.5
Detector collimation (mm)	64 × 0.625	128 × 0.6	64 × 0.625
Pitch	0.891	0.6	0.914
Slice thickness (mm)	5	5	5
Slice spacing (mm)	5	5	5
Matrix	512 × 512	512 × 512	512 × 512
FOV (mm)	350	300	350
Algorithm (B)	Standard	Standard	Standard

MATERIALS AND METHODS

Patients

Data from a total of 324 patients with GISTs presenting from January 1, 2016 to July 1, 2019 were collected retrospectively from three hospitals. Among them, 180 cases were analyzed from the First Affiliated Hospital of Zhejiang University School of Medicine, which was used as the training cohort, while 144 cases from another two hospitals (Zhejiang Cancer Hospital and the First Affiliated Hospital of Wannan Medical College) were used as the external validation cohort. The inclusion criteria were as follows: (1) surgical resection, negative margin, and a pathological diagnosis of GISTs, (2) abdominal enhancement CT examination within 15 days prior to surgery, and (3) pathological results with a clear risk assessment. Exclusion criteria were as follows: (1) patients receiving imatinib or other neoadjuvant therapy before surgery, and (2) those with poor CT image quality.

Clinical data, including age, gender, and tumor site, were derived from medical records. The National Institutes of Health's modified criteria were used to stratify the malignant potential of GISTs on the basis of the clinical and post-operative histological index. All patients were divided into two groups: high malignant potential group with intermediate risk and high risk; and low malignant potential group with very low risk and low risk. This study was a retrospective study, and the patient's informed consent was thereby waived, as approved by the hospital ethics committee.

CT Image Acquisition

All subjects received a default abdominal CT scan using one of the three multidetector CT (MDCT) systems with the scanning and reconstruction parameters used in daily clinical practice. See **Table 1** for the detailed information of the CT protocol. Three-phase scans were unenhanced phase, arterial phase (25–30 s after injection), and portal vein phase (55–60 s after injection). The dose of iodine contrast agent was based on the patient's weight (1 mL/kg), and the flow rate was 2.5–3.5 mL/s.

Three-Dimensional Segmentation of Tumor Images and Radiomics Feature Extraction

Both tumor segmentation and radiomics feature extraction were performed using Matlab's IBEX software package (18). Two radiologists with a depth of experience delineated the regions of interest (ROIs) layer by layer along the edge of the tumor contour on the CT images of the arterial and portal venous phases.

All images were preprocessed with image resampling (voxel size of $1 \times 1 \times 1$ mm^3) and gray value homogenization (normalized to 1–256, fixed bin number method, 256 bins) before radiomics feature extraction. The radiological feature parameters involved six major categories: histogram parameters ($n = 48$), 2.5D and 3D gray level co-occurence matrix ($n = 594$, the 2.5D feature is computed from a single matrix after merging all 2D directional matrices, the 3D feature is computed from a single matrix after merging all 3D directional), gray level adjacent difference ($n = 10$), gray level run length matrix ($n = 34$), shape and size ($n = 18$). In each stage, we retrieved 704 parameters, and a total of 1,408 parameters were collected in the two stages.

During the early stage of the study, we randomly selected images from 40 patients, and two radiologists with more than 10 years of work experience performed ROI delineation independently. The blindness method was used to analyze the reliability and repeatability between observers. The consistency was evaluated using the intra-class correlation coefficient (ICC). There is a good agreement when the ICC is > 0.75. ROI extraction of the remaining images was performed by one of the radiologists.

Feature Selection and Radiomics Model Building

Redundancy and overcorrelation in the characteristics of radiomics often lead to overfitting of the prediction model. In this study, we dimensionally reduced the radiomics features in two steps. First, multicollinearity of the features were analyzed by spearman correlation, and the correlation coefficient threshold was 0.8. Then, we used the boruta algorithm to iteratively assess the importance of features, and we removed the irrelevant features. Boruta algorithm can filter out all the characteristics related to the dependent variable and generate a ranking of importance. To achieve statistical significance, the top 10 features in importance ranking were selected for final modeling.

After dimensionality reduction of the radiomics features, the three most popular classifiers [logistic regression, support vector machine (SVM), and random forest] were applied to establish three risk stratification models for radiological prediction. We conducted holdout cross-validation for 30 times for each model in the training cohort (training: internal validation ratio is 4:1). Because each iteration is a resampling of the training cohort, each model yielded 30 different values of area under the curve (AUC), specificity, sensitivity, and accuracy, among which we used AUC as the standard to evaluate the effectiveness of the three models in the training cohort.

Subsequently, the three models were applied to the external validation cohort, and the effectiveness of the models were also evaluated through AUC, specificity, sensitivity, and accuracy.

Statistical Analysis

All statistical analysis was performed using R software (version 3.4.1; http://www.Rproject.org). We performed descriptive statistical analysis for the training and external validation cohorts, and quantitative data was described as mean ± standard deviation (SD) and qualitative data was described by frequency (percent). Qualitative variables were compared using the chi-square test. Continuous variable data was evaluated using a two-sample *t*-test or Wilcoxon test. AUC was used as the evaluation standard for the comparison of the three classification algorithms in the training cohort. The Fridman test was used for the comparison among the three algorithms, and the Nemenyi test was used in *post-hoc* analysis. Two tailed $p < 0.05$ was considered statistically significant.

RESULTS

Clinical Characteristics

In total, 324 GIST patients were included in this study, of which 150 patients had low malignant potential and 174 patients had high malignant potential. Ninety-three men and 87 women were included in the training cohort, and 64 men and 80 women were included in the external validation cohort. **Table 2** shows the baseline clinical data. Single factor analysis showed that there was no statistically significant difference between the low and the high malignant potential groups in terms of age, gender, and tumor site.

After dimension reduction by spearman correlation, we obtained 107 features, which through the dimension reduction by boruta algorithm, 25 parameters remained, from which we extracted the top 10 features, according to the built-in importance-ranking system. In the subset, parameters from the portal venous phase accounted for 80%. Morphology ranks the most important, although only one parameter was selected. See **Table 3** for a list of specific parameters and their importance.

Radiomics Model Performance

The specific performance of the three classifier prediction models is shown in **Table 4** and **Figures 1**, **2**. The Friedman test indicated that the AUC value of the three models in the training cohort was significantly different ($p < 0.001$). The Nemenyi test results show that the AUC of random forest was significantly higher than logistic regression ($p = 0.001$), significantly higher than SVM ($p = 0.0103$), and there was no significant statistics between logisitic regression and SVM ($p = 0.09$). The Friedman-Nemenyi test indicated that the AUC value of the random forest model was significantly higher than that of the other two prediction models. The random forest model achieved the most satisfactory results; the performance and generalizability were favorable. The performance of the SVM and logistic regression models were satisfactory, and the generalizability was acceptable, but the overall efficiency was not outstanding.

DISCUSSION

In this study, we built three prediction models based on CT radiomics for GIST risk stratification. After

TABLE 2 | Patient characteristics in the training and external validation cohorts.

Patient characteristics	Training cohort			External validation cohort		
	Low-malignant potential GISTs (*n* = 82)	High-malignant potential GISTs (*n* = 98)	*p*-value	Low-malignant potential GISTs (*n* = 68)	High-malignant potential GISTs *n* = 76)	*p*-value
Age (mean ± SD, years)	54.13 ± 8.31	56.71 ± 10.52	0.74	55.13 ± 8.31	57.12 ± 11.45	0.63
Gender (%)			0.15			0.77
Male	37 (45.12%)	56 (57.14%)		31 (45.59%)	33 (43.42%)	
Female	45 (54.88%)	42 (42.86%)		37 (54.41%)	43 (56.58%)	
Primary site (%)			0.65			0.19
Gastric	48 (58.53%)	53 (54.08%)		45 (66.18%)	42 (55.26%)	
Intestinal	34 (41.47%)	45 (45.92%)		23 (33.82%)	34 (44.74%)	

TABLE 3 | Texture features selection for radiomics models.

Parameters category	Parameters	Phase	Importance
Morphology	Volume	Portal venous phase	21.11
Gray level co-occurrence matrix	Variance	Portal venous phase	9.26
Gray level co-occurrence matrix	Inverse variance	Arterial phase	8.04
Gray level co-occurrence matrix	Cluster shade	Portal venous phase	7.78
Gray level adjacent difference	Contrast	Portal venous phase	7.59
Gray level co-occurrence matrix	Max probability	Arterial phase	6.17
Gray level adjacent difference	Busyness	Portal venous phase	5.39
Gray level co-occurrence matrix	Sum average	Portal venous phase	5.23
Gray level adjacent difference	Texture strength	Portal venous phase	5.15
Gray level adjacent difference	Complexity	Portal venous phase	5.14

comparing the three most commonly used machine-learning models in radiomics, we found the random forest model showed the best performance in discriminating GISTs malignant potentials, and its generalizability is outstanding.

GISTs often exhibit complex and unpredictable biological behaviors. With the development of molecular pathology research, imatinib has emerged as a first-line molecular targeted drug, which has changed the treatment of GISTs and has become a successful model for the targeted diagnosis and treatment of solid tumors. The stratification of patients based on the risk of recurrence is a key issue in managing primary GISTs. The National Comprehensive Cancer Network guidelines recommend more than 3 years of post-operative imatinib be used as an adjuvant therapy for patients with a high recurrence risk (high-risk and intermediate-risk) (19, 20), while patients with a low recurrence risk (low-risk and very low-risk) that can be cured *via* surgical resection of the tumor should not receive adjuvant therapy with imatinib (21–23). Therefore, in this study, GIST patients were classified into low and high malignant potential groups according to the risk stratification. Because the clinical characteristics of GISTs lack specificity, the preoperative diagnosis and risk stratification of GISTs mainly rely on imaging examinations. Traditional imaging evaluates the risk of GISTs by observing the size, shape, presence or absence of necrosis, ulcers, and enhancement of GISTs, and the results depend much on the professional ability and subjective experience of radiologists (4, 6–9).

The rise in the use of radiomics in recent years has resulted in imaging studies to predict GISTs recurrence risks using objective and quantitative measures. Currently, most GISTs radiomics studies focus on risk prediction, and the AUC is relatively high at ∼0.81–0.94 (15, 19, 24–27), demonstrating the superiority of radiomics over traditional methods in terms of prediction effectiveness. It also lays foundation for the future application of radiomics for GIST risk stratification. However, only one study has also conducted external data validation of the model (24). Its model efficiency was 0.87 in the training cohort and 0.85 in the external validation cohort. Although the performance of the model was not optimal, this study has published the most standardized and reliable results to date. There is no external validation for the other studies; the same data were used for the training and validation cohort, making the results less convincing (28). Studies have confirmed that equipment from different manufacturers results in differences in scanning parameter settings and post-processing reconstruction algorithms, resulting in significant differences

TABLE 4 | A performance summary of the radiomics models in the training and external validation cohorts.

	Accuracy	Sensitivity	Specificity	AUC
Logistic regression				
Training cohort	0.77 ± 0.08	0.61 ± 0.11	0.86 ± 0.10	0.84 ± 0.07
External validation cohort	0.75	0.65	0.84	0.85
Random forest				
Training cohort	0.82 ± 0.07	0.84 ± 0.10	0.73 ± 0.10	0.88 ± 0.06
External validation cohort	0.84	0.93	0.76	0.90
Support vector machine				
Training cohort	0.75 ± 0.07	0.52 ± 0.12	0.91 ± 0.08	0.81 ± 0.08
External validation cohort	0.71	0.74	0.68	0.80

Values of accuracy, sensitivity, specificity, and AUC of the three models in the training cohort are the average values after 30 holdout cross-validation, which were described as mean ± standard deviation (SD). AUC, areas under the curve.

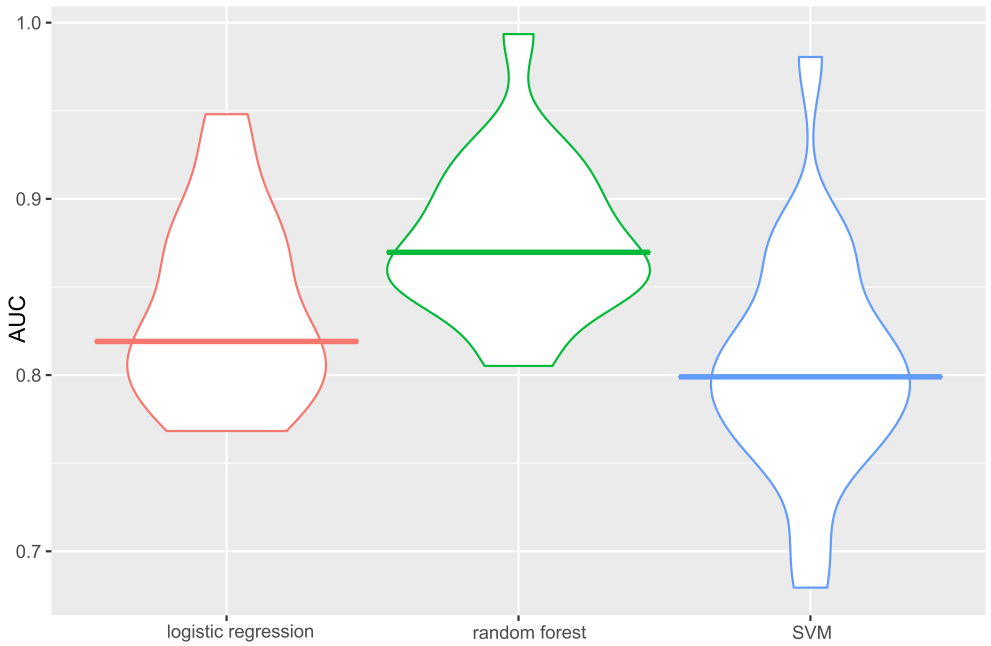


FIGURE 1 | AUC of the three classifier prediction models performance in the training cohort. The random forest model achieved the best satisfactory results. The AUC is the average AUC obtained after 30 holdout cross-validation. The horizontal line of each diagram corresponds to the average AUC. AUC, the area under the curve; SVM, support vector machine.

in the radiomics parameters (29–31). Therefore, single-center research has its limitations (32). Multicenter research can provide diverse imaging data to better interpret tumor heterogeneity, which is also in line with the development of precision medicine (33). The highlight of this research lies in its multicenter design, which uses the largest amount of data among the three hospitals as the training cohort, while the data from the other two hospitals are fused into an independent external validation cohort. We found that the AUC of the random forest model in the training cohort was 0.88 ± 0.06, which was very good in both the training cohort and the validation cohort, indicating that the generalizability of the model is excellent. Our study confirms the potential of radiomics in GISTs diagnosis and prognosis, and it proposes that the predicted models must undergo multicenter testing before providing a reliable reference for clinical decision-making (34).

Different machine-learning algorithms have their own advantages and disadvantages. The performance of an algorithm in a specific machine-learning task cannot be predicted before research. Most previous radiomics studies used a single algorithm for modeling, and no specific reason was stated for choosing the model. Currently, the most common GIST risk stratification models are logistic regression, SVM, and random forest. Logistic regression is the most commonly used classification algorithm in the medical field (35) and in GISTs imaging histology. Wang et al. (26). collected 333 GISTs cases, and the AUC of the training

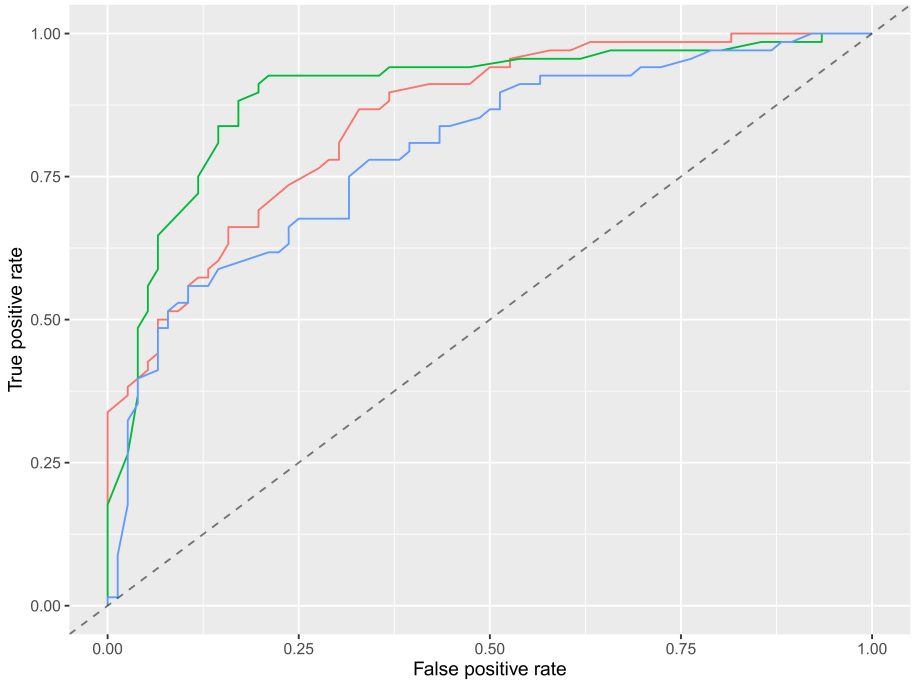


FIGURE 2 | ROC diagram of multiple models in the external validation cohort. Red is logistic regression, green is random forest, and blue is support vector machine.

cohort was 0.88. Ren et al. (27) also used logistic regression with 440 cases, and the final AUC of the training cohort was 0.93. SVM has many advantages in processing small samples and non-linear and high-dimensional data. Chen et al. chose SVM to build a prediction model, and the AUC was 0.86 in the training cohort and 0.85 in the external validation cohort. Random forest is a type of integrated machine learning, which is based on the decision tree method and can improve the prediction accuracy without significantly increasing the amount of calculation (36). Zhang et al. (19) used a random forest algorithm to predict GIST risk stratification, and achieved an AUC of 0.94 of the training cohort, which is the best performance among similar studies. These studies have their own advantages, but due to the heterogeneity between the data cohorts, the differences of the classifiers cannot be clarified. Hence, it is impossible to determine which classifier is the most suitable for stratifying the GIST risk. In this study, we conducted a multiclassification algorithm study on the same data and task and found that logistic regression and SVM performed stably, but the overall efficiency was not outstanding. Random forest performed the best in both the training and external validation cohorts, with the highest AUC and excellent generalizability, which indicated that this method is worthy of in-depth study and verification with a larger sample set and data from a multicenter study.

However, our study has the following limitations: (1) Our sample size was relatively small, and limited to Chinese people. As genetic mutations are the driving factors in the occurrence of GISTs, and the morbidity and mortality of GISTs varies among different races, it is necessary to conduct further in-depth studies on large samples of multinational and multiethnic populations, ideally in multicenter trials. (2) Because most of the previous articles suggested clinical parameters were not significant, this study used pure radiomic modeling and did not integrate clinical parameters for further analysis and comparison. (3) This study was a retrospective study, and the sample selection was biased, which requires further verification in prospective studies. (4) As the CT imaging protocols varies in different hospitals, radiomics features are affected by CT scanner parameters, such as reconstruction kernel or section thickness, thus obscuring underlying biologically important radiomics parameters. We did not process the data from multicenter with harmonization. Some features of IBEX are not compatible with IBSI (Image Biomarker Standardisation Initiative), which will affect the reproducibility of the results. (5) The algorithm of feature selection also affects the model performance. We did not compare the algorithms of dimensionality reduction; therefore, the final feature selection may not be the optimal.

In conclusion, this study predicts the risks of GISTs based on different machine-learning models of CT radiomics. After comparing the three most commonly used machine-learning algorithms in radiomics, a radiomics model of the random forest algorithm presents the most satisfactory prediction. The efficacy, optimal discrimination, strong generalizability, and confirmation in external validation data can be used as a more objective and non-invasive

technique, which has the potential to become an effective tool for clinicians to predict the risk stratification of GISTs before surgery.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional review board of the Affiliated Hospital of College of Medicine Zhejiang University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

MW and ZF proposed the conception and design of this research and analyzed and interpreted the data. ZF and LZho developed methodology. MW, ZF, LZha, and XH collected data and performed preprocessing. MW, ZF, and JZ were major contributors in writing the manuscript. All authors read and approved the final manuscript.

REFERENCES

1. Raut CP, Espat NJ, Maki RG, Araujo DM, Trent J, Williams TF, et al. Efficacy and tolerability of 5-year adjuvant imatinib treatment for patients with resected intermediate- or high-risk primary gastrointestinal stromal tumor: the PERSIST-5 clinical trial. *JAMA Oncol*. (2018) 4:e184060. doi: 10.1001/jamaoncol.2018.4060
2. Gill AJ, Angela C, Ricardo V, Adele C, Millie L, Jin R, et al. Immunohistochemistry for SDHB divides gastrointestinal stromal tumors (GISTs) into 2 distinct types. *Am J Surg Pathol*. (2010) 34:636–44. doi: 10.1097/PAS.0b013e3181d6150d
3. Joensuu H. Risk stratification of patients diagnosed with gastrointestinal stromal tumor. *Hum Pathol*. (2008) 39:1411–9. doi: 10.1016/j.humpath.2008.06.025
4. Maldonado FJ, Sheedy SP, Iyer VR, Hansel SL, Bruining DH, Mccollough CH, et al. Reproducible imaging features of biologically aggressive gastrointestinal stromal tumors of the small bowel. *Abdom Radiol*. (2017) 43:1567–74. doi: 10.1007/s00261-017-1370-6
5. Li J, Ye Y, Wang J, Zhang B, Qin S, Shi Y, et al. Chinese consensus guidelines for diagnosis and management of gastrointestinal stromal tumor. *Chin J Cancer Res*. (2017) 29:281–93. doi: 10.21147/j.issn.1000-9604.2017.04.01
6. Wang JK. Predictive value and modeling analysis of MSCT signs in gastrointestinal stromal tumors (GISTs) to pathological risk degree. *Eur Rev Med Pharmacol Sci*. (2017) 21:999–1005.
7. Zhou C, Duan X, Zhang X, Hu H, Wang D, Shen J. Predictive features of CT for risk stratifications in patients with primary gastrointestinal stromal tumour. *Eur Radiol*. (2016) 26:3086–93. doi: 10.1007/s00330-015-4172-7
8. Su Q, Wang Q, Zhang H, Yu D, Wang Y, Liu Z, et al. Computed tomography findings of small bowel gastrointestinal stromal tumors with different histologic risks of progression. *Abdom Radiol*. (2018) 43:2651–8. doi: 10.1007/s00261-018-1511-6
9. Xue A, Yuan W, Gao X, Fang Y, Shu P, Xu C, et al. Gastrointestinal stromal tumors (GISTs) with remarkable cystic change: a specific subtype of GISTs with relatively indolent behaviors and favorable prognoses. *J Cancer Res Clin Oncol*. (2019) 145:1559–68. doi: 10.1007/s00432-019-02853-y
10. Vivek V, Simone CB, Sunil K, Lin SH, Yang J, Hahn SM. The rise of radiomics and implications for oncologic management. *J Natl Cancer Inst*. (2017) 109:djx055. doi: 10.1093/jnci/djx055
11. Liang C, Huang Y, He L, Chen X, Ma Z, Dong D, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer. *Oncotarget*. (2016) 7:31401–12. doi: 10.18632/oncotarget.8919
12. Lu W, Chen W. Positron emission tomography/computerized tomography for tumor response assessment—a review of clinical practices and radiomics studies. *Transl Cancer Res*. (2016) 5:364–70. doi: 10.21037/tcr.2016.07.12
13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. (2015) 278:151169. doi: 10.1148/radiol.2015151169
14. Choi IY, Yeom SK, Cha J, Cha SH, Lee SH, Chung HH, et al. Feasibility of using computed tomography texture analysis parameters as imaging biomarkers for predicting risk grade of gastrointestinal stromal tumors: comparison with visual inspection. *Abdom Radiol*. (2019) 44:2346–56. doi: 10.1007/s00261-019-01995-4
15. Feng C, Lu F, Shen Y, Li A, Yu H, Tang H, et al. Tumor heterogeneity in gastrointestinal stromal tumors of the small bowel: volumetric CT texture analysis as a potential biomarker for risk stratification. *Cancer Imaging*. (2018) 18:46. doi: 10.1186/s40644-018-0182-4
16. Yarandi RB, Panahi MH. Noninvasive prediction model for diagnosing gastrointestinal stromal tumors using contrast-enhanced harmonic endoscopic ultrasound: Methodological issues. *Dig Liver Dis*. (2019) 51:1355–6. doi: 10.1016/j.dld.2019.04.017
17. Demiar J, Schuurmans D. Statistical comparisons of classifiers over multiple data sets. *J Mach Learn Res*. (2006) 7:1–30. doi: 10.1007/s10846-005-9016-2
18. Zhang L, Fried DV, Fave XJ, Hunter LA, Yang J, Court LE. IBEX: An open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys*. (2015) 42:1341–53. doi: 10.1118/1.4908210
19. Zhang L, Kang L, Li G, Zhang X, Ren J, Shi Z, et al. Computed tomography-based radiomics model for discriminating the risk stratification of gastrointestinal stromal tumors. *Radiol Med*. (2020) 125:465–73. doi: 10.1007/s11547-020-01138-6
20. Blay JY, Levard A. Adjuvant imatinib treatment in gastrointestinal stromal tumor: which risk stratification criteria and for how long? A case report. *Anti Cancer Drugs*. (2016) 27:71–5. doi: 10.1097/CAD.0000000000000286
21. Rutkowski P, Przybył J, Zdzienicki M. Extended adjuvant therapy with imatinib in patients with gastrointestinal stromal tumors. *Mol Diagn Ther*. (2013) 17:9–19. doi: 10.1007/s40291-013-0018-7
22. Blay JY, Rutkowski P. Adherence to imatinib therapy in patients with gastrointestinal stromal tumors. *Cancer Treat Rev*. (2014) 40:242–7. doi: 10.1016/j.ctrv.2013.07.005
23. López RL, Del Muro XG. Management of localized gastrointestinal stromal tumors and adjuvant therapy with imatinib. *Anti Cancer Drugs*. (2012) 23:S3–6. doi: 10.1097/CAD.0b013e3283559fab
24. Chen T, Ning Z, Xu L, Feng X, Han S, Roth HR, et al. Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively. *Eur Radiol*. (2019) 29:1074–82. doi: 10.1007/s00330-018-5629-2
25. Liu S, Pan X, Liu R, Zheng H, Chen L, Guan W, et al. Texture analysis of CT images in predicting malignancy risk of gastrointestinal stromal tumours. *Clin Radiol*. (2017) 73:266–74. doi: 10.1016/j.crad.2017.09.003
26. Wang C, Li H, Jiaeken Y, Huang P, Sun L, Dong F, et al. Building CT radiomics-based models for preoperatively predicting malignant potential and mitotic count of gastrointestinal stromal tumors. *Transl Oncol*. (2019) 12:1229–36. doi: 10.1016/j.tranon.2019.06.005
27. Ren C, Wang S, Zhang S. Development and validation of a nomogram based on CT images and 3D texture analysis for preoperative prediction of the malignant potential in gastrointestinal stromal tumors. *Cancer Imaging*. (2020) 20:5. doi: 10.1186/s40644-019-0284-7
28. Hunter LA, Krafft S, Stingo F, Choi H, Martel MK, Kry SF, et al. High quality machine-robust image features: identification in nonsmall

cell lung cancer computed tomography images. *Med Phys*. (2013) 40:121916. doi: 10.1118/1.4829514
29. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol*. (2015) 50:757–65. doi: 10.1097/RLI.0000000000000180
30. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep*. (2015) 5:13087. doi: 10.1038/srep13087
31. Balagurunathan Y, Gu Y, Wang H, Kumar V, Grove O, Hawkins S, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol*. (2014) 7:72–87. doi: 10.1593/tlo.13844
32. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of a method to compensate multicenter effects affecting CT radiomics. *Radiology*. (2019) 291:53–9. doi: 10.1148/radiol.2019182023
33. Zhao B, Tan Y, Tsai WY, Schwartz LH, Lu L. Exploring variability in CT characterization of tumors: a preliminary phantom study. *Transl Oncol*. (2014) 7:88–93. doi: 10.1593/tlo.13865
34. Caramella C, Allorant A, Orlhac F, Bidault F, Asselain B, Ammari S, et al. Can we trust the calculation of texture indices of CT images? A phantom study. *Med Phys*. (2018) 45:1529. doi: 10.1002/mp.12809
35. Nakasone Y, Ikeda O, Yamashita Y, Kudoh K, Shigematsu Y, Harada K. Shock index correlates with extravasation on angiographs of gastrointestinal hemorrhage: a logistics regression analysis. *Cardiovasc Intervent Radiol*. (2007) 30:861–5. doi: 10.1007/s00270-007-9131-5
36. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inform Comput Sci*. (2003) 43:1947–58. doi: 10.1021/ci034160g

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

REVIEW
published: 05 May 2021
doi: 10.3389/fonc.2021.603595

The Role of Radiomics in Lung Cancer: From Screening to Treatment and Follow-Up

Radouane El Ayachy[1,2,3]*, *Nicolas Giraud*[3,4], *Paul Giraud*[1,2,3], *Catherine Durdux*[1,2], *Philippe Giraud*[1,2], *Anita Burgun*[2,3] and *Jean Emmanuel Bibault*[1,2,3]

[1] Radiation Oncology Department, Georges Pompidou European Hospital, Assistance Publique-Hôpitaux de Paris, Université de Paris, Paris, France, [2] Cancer Research and Personalized Medicine-Integrated Cancer Research Center (SIRIC), Georges Pompidou European Hospital, Assistance Publique-Hôpitaux de Paris, Université de Paris, Paris, France, [3] INSERM UMR 1138 Team 22: Information Sciences to support Personalized Medicine, Cordeliers Research Centre, Paris Descartes University, Paris, France, [4] Radiation Oncology Department, Haut-Lévêque Hospital, CHU de Bordeaux, Pessac, France

Purpose: Lung cancer represents the first cause of cancer-related death in the world. Radiomics studies arise rapidly in this late decade. The aim of this review is to identify important recent publications to be synthesized into a comprehensive review of the current status of radiomics in lung cancer at each step of the patients' care.

Methods: A literature review was conducted using PubMed/Medline for search of relevant peer-reviewed publications from January 2012 to June 2020

Results: We identified several studies at each point of patient's care: detection and classification of lung nodules (n=16), determination of histology and genomic (n=10) and finally treatment outcomes predictions (=23). We reported the methodology of those studies and their results and discuss the limitations and the progress to be made for clinical routine applications.

Conclusion: Promising perspectives arise from machine learning applications and radiomics based models in lung cancers, yet further data are necessary for their implementation in daily care. Multicentric collaboration and attention to quality and reproductivity of radiomics studies should be further consider.

Keywords: radiomics, lung cancer, machine learning, oncology, lung cancer screening, treatment outcome and efficiency

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Antoine Girard, Centre Eugène Marquis, France
Sylvain Reuzé, GE Healthcare, France

***Correspondence:**
Radouane El Ayachy
radouane.elayachy@gmail.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *07 September 2020*
Accepted: *06 April 2021*
Published: *05 May 2021*

Citation:
El Ayachy R, Giraud N, Giraud P, Durdux C, Giraud P, Burgun A and Bibault JE (2021) The Role of Radiomics in Lung Cancer: From Screening to Treatment and Follow-Up. Front. Oncol. 11:603595. doi: 10.3389/fonc.2021.603595

Abbreviations: 18F-FDG, 18F-fluorodeoxyglucose; ALK, anaplastic lymphoma kinase; ANN, artificial neural network; AUC, area under the curve; BRAF, v-raf murine sarcoma viral oncogene homolog B1; CAD, computer aided diagnosis; CNN, convolutional neural networks; CBCT, cone beam computed tomography; CT, computed tomography; EGFR, epidermal growth factor receptor; FNR, false negative rate; FPR, false positive rate; IBSI, Image Biomarker Standardisation Initiative; ICC, intraclass correlation coefficient; KRAS, Kirsten rat sarcoma viral oncogene homolog; LASSO, Least Absolute Shrinkage and Selection Operator; ML, machine learning; MRMR, minimum redundancy maximum relevance; NILST, National Lung Screening Trial; NSCLC, non-small-cell lung carcinoma; OS, overall survival; PCA, principal component analysis; pCR, pathologic complete response; PET, positron emission tomography; PML, pixel/voxel-based machine learning; PSO, particle swarm optimization; QDA, quadratic discriminant analysis; RFC, random forest classifier; ROI: region of interest; RP, radiation pneumonitis; RILI, radiation induced lung injury; RSF, random survival forests; SBRT, stereotactic body radiation therapy; SVM, support vector machine.

INTRODUCTION

Death from lung cancer is estimated to be 1.7 millions each year worldwide, essentially due to late diagnoses (1), making it the first cause of cancer-related death in the world (2) despite recent discoveries in the field of tumor biology and new treatment strategies. The emergence of new targeted treatment focusing on specific biomolecular alterations such as EGFR (3) and ALK mutations has led to a new paradigm of cancer care, so-called "personalized" medicine, conversely to the historic "one-size-fits-all" medicine. In that regard, radiomics could also play a role in patient-specific treatment adaptations.

Common imaging interpretation, for instance with positron emission tomography (PET), Magnetic resonance imaging (MRI) or computed tomography (CT), relies on the visual analysis in terms of size, shape, signal intensity or contrast enhancement of various structures within the image.

« Radiomics », with reference to genomics, has been introduced in 2012 by Lambin et al. (4). Its aim is to extract a large number of quantitative variables from medical imaging, followed by a selection of the most informative ones in order to derive a scientific hypothesis.

Radiomics is based on the innovative approach that computerized algorithms are able to process imaging exams into more complex quantitative data. They can be applied to different imaging modalities (ultrasound, CT, PET, conventional radiology) by analyzing in a selected region of interest (ROI) the distribution of signal intensities.

Different ROI segmentation methods can be used. Manual delineation is close to daily practice, but requires a considerable amount of human time, limiting the creation of large databases, and is subject to high inter- and intra-observer variability (5–7). Automatic segmentation is thus largely preferable for reproducibility purposes, but is only applicable when there is a strong signal difference between the lesion and the adjacent tissues. This is why semi-automatic approaches are most often necessary: a software program defines a delineation which is then adjusted by the observer (8).

The extracted variables are divided into three categories (**Figure 1**): shape variables, first-order variables and second-order variables. The shape variables describe, independently of grey levels, the shape, surface area and dimensions of the ROI (example: surface area in square millimeters, sphericity,…). The first-order variables study the distribution of voxel gray level intensity values without consideration of spatial relationships. As for the second-order variables, they describe the spatial relationships between the voxels generally from matrices (example: grayscale co-occurrence matrix, size of homogeneous grayscale areas, neighborhood grayscale difference, length of grayscale ranges, grayscale dependence).

Like other high-throughput techniques, labeled "-omics" (9), radiomics aims to develop new imaging biomarkers to better understand the microbiology of cancer (10). The use of radiomics could provide additional data about the biological constitution of a tissue, predict treatment response or even offer new prognostic markers.

Radiomics thus offer several advantages due to their non-invasive character, the possibility to account for intra-tumor heterogeneity (11) by a complete analysis of the tumor, and inter-lesional heterogeneity (12) by sampling all the tumors within the

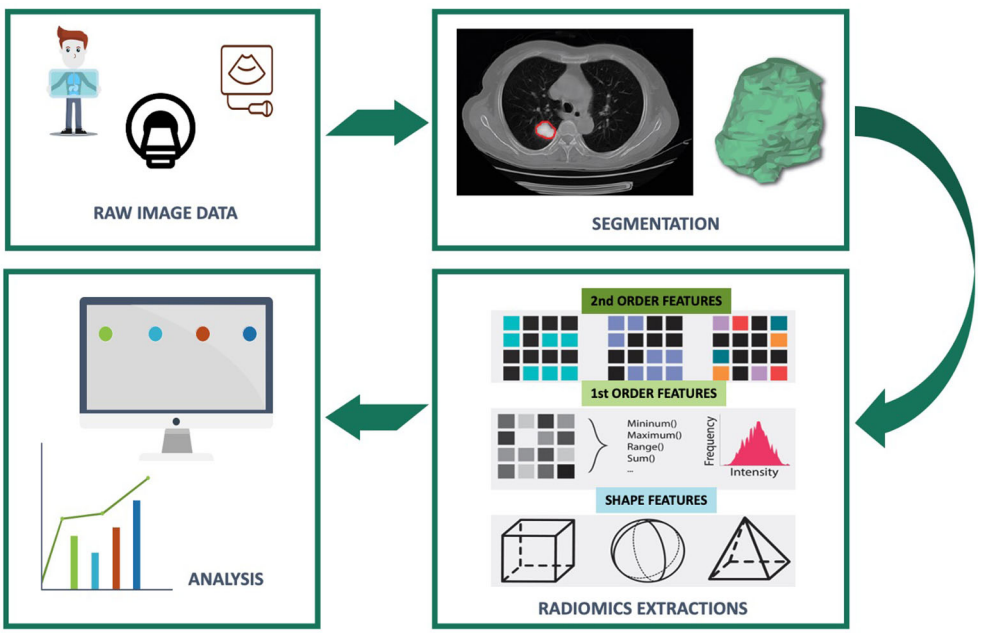


FIGURE 1 | Flowchart of radiomics feature based analysis.

same patient as well as the tumor microenvironment. They also allow monitoring temporal heterogeneity (13).

The last few decades have been paved by the advent of clinical, biological, radiological and genomic diagnostic advances offering access to a multitude of new data available for each patient as well as by the development of new therapeutics that are more targeted and personalized to each patient. Given the large amount of information generated, the major challenge in enabling personalized treatment in oncology lies in the ability to exploit this wealth of information to accurately predict the behavior and response of a tumor. Machine learning seems to be able to process and manage this huge amount of information.

In machine learning, a classification model is trained from a data set in order to "learn" (training set) the distribution of the different classes in a multidimensional variable space. In machine learning, there are several methods, each with their advantages and disadvantages (14). They are grouped into two types of classification: supervised and unsupervised.

In the supervised classification methods, individuals are labelled (e.g., benign vs. malignant) and the algorithm tries to predict this explicit variable, called the output variable, from a large number of input variables (radiomics, genomics, clinical,…).

Unsupervised methods do not use predefined output variables. The goal is to find a model that groups the most similar data together and separates the most different data, known as clustering. For example, K-means clustering generates K clusters by comparing the degree of similarity of observations, so that two individuals that are similar will have a reduced distance of dissimilarity.

One of the most used ML subset is Artificial Neural Network (ANN) (**Figure 2**). It is considered as a supervised classification model. Its variant, Deep Learning (DL), is associated with the feature extraction, directly from raw imaging data, through a series of nonlinear processing units comprising multiple layers, which tries to establish a relationship between stimuli and associated neural responses present in the brain.

The expansion of medical imaging data (14) in lung cancer offers an opportunity to explore the value of radiomics for every step of the patient's care: screening, diagnosis, staging, treatment planning, and response evaluation. The objective of this article is to benchmark radiomics applications in lung cancer at each of these steps.

MATERIALS AND METHODS

The authors conducted a literature review using PubMed/Medline in order to identify important recent publications to be synthesized into a comprehensive review of the current status of radiomics in lung cancer at each step of the patients' care. A comprehensive list of MeSH terms and keywords was included in the search: "lung cancer," "radiomics," "signature," "machine learning," as well as other associated technical ML

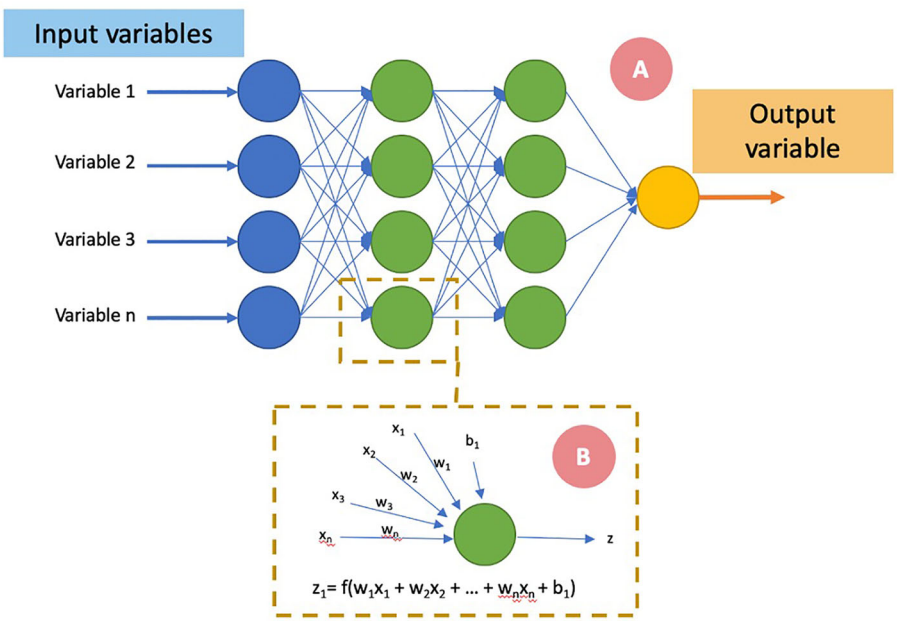


FIGURE 2 | Schematic representation of an artificial neural network. The input variables **(A)** are presented at the first neural layer (blue). The information is then passed to a succession of layers ("hidden layers," in green) and finally an output neural layer predicting the variable to be estimated. Each layer (i) consists of Ni neurons, taking their inputs from the Ni-1 neurons of the previous layer. A neuron **(B)** adds each of its inputs (xn) and multiplies them by a weight (wn). An activation function (f) allows according to a threshold the activation of the neuron and the transmission of information (z) to the next layer. An optimizer adjusts the weights and biases (b) of each neuron in order to make the neural network converge toward its state allowing it to make the best prediction.

keywords. Selected articles were published between January 2012 and June 2020, and based on relevance to the subject. The search strategy also included screening of reference lists of relevant publications. The search query returned 133 articles that were screened. We removed review articles and selected 49 studies in the final analysis.

RESULTS

Characterization of Lung Abnormalities

One of the first application of radiomics in lung cancer was tumor detection. Lung abnormality discoveries are frequent; thus, the challenge is to be able to distinguish benign lesions from malignant ones. Qualitative features such as measurements of diameter or volume of pulmonary nodules provide important information to differentiate benign from malign nodules. Notwithstanding the encouraging results of low-dose computed tomography (CT) versus (vs.) chest X-ray in lung cancer-specific mortality reduction (15), the application of low-dose CT in selected population screening remains contested (16) on account of its cost-efficiency, the high false positive rate (FPR) and the optimal schedule (1). In that setting, overdiagnosis remains a challenging issue (17). In addition, due to the lack of validated software, the volumetric assessment of the lesion is not the current standard of practice (18). 18F-fluorodeoxyglucose positron emission tomography (18F-FDG PET)/CT is a performant tool to help clinicians in the characterization of lung nodules (19) but still holds a low detection rate of small lesions (20) and delivers high radiation doses.

Recent promising strategies based on radiomics or circulating biomarkers (21) could be interesting and less invasive (22). Computer aided diagnosis (CAD) systems can help to improve radiologists' performances (23) on tumor detection and could be even further improved by radiomics.

Radiomics features could be used in traditional statistic model as linear classifier with high accuracy in predicting lung nodule malignancy (24).

Integrating radiomics, the optimal ML model to apply remains unknown. Random forest classifiers showed good performance in anticipating nodules that would become cancerous one and two years later, with accuracies of 80% (25), better than a Support Vector Machine (SVM) classifier or the recently developed McWilliams (26) and Lung-RADS (27) risk scores. Schematically, SVM models, through a kernel function, depict individuals in a 3rd dimensional space in order to find a hyperplane that classifies individuals into two groups (**Figure 3**).

Different supervised ML models can also be used together. After a feature selection by a Random Forest classifier (RFC), Wang et al. (28) found 15 radiomics features able to single out benign from malignant nodules with an accuracy of 86% through a SVM algorithm.

Some studies tried to benchmark the added value of clinical features to these radiomics features. As a matter of fact, they can improve the performance of ML methods to distinguish focal

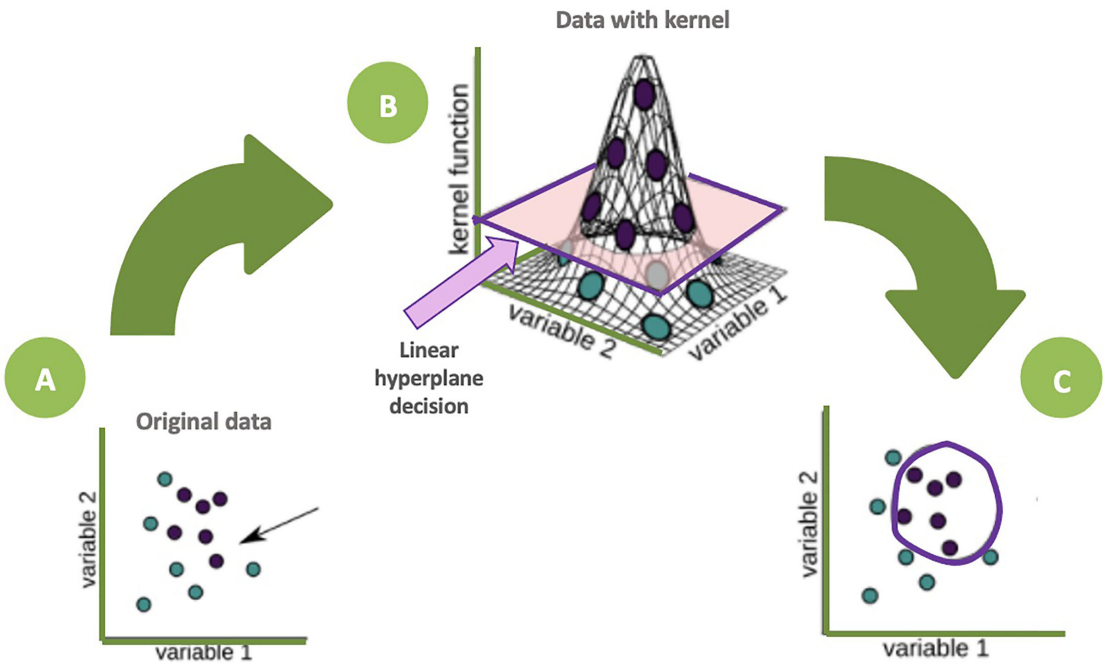


FIGURE 3 | Schematic representation of a SVM algorithm. The dots represent individuals according to two variables **(A)**, no linear classification function seems obvious. The kernel function allows a representation of the individuals in a 3rd dimension allowing the highlighting of a hyperplane which classifies the individuals in two groups **(B)**. The individuals are then projected into the initial dimensional space **(C)** with a non-linear separator (purple circle).

pneumonia from adenocarcinoma (29) or non-small-cell lung cancer (30). Clinical features addition could also produce no improvement of the model performance (24), highlighting the importance of the radiomics features.

Interestingly, some studies (31) indicate a trend toward increased performance when the surrounding parenchyma is included, revealing the importance of microenvironment.

Most studies use radiomics approaches needing 2D or 3D quantitative images features. Another category of computational strategy is Deep learning and particularly convolutional neural networks (CNN). CNN could perform prediction without needing nodule segmentation, taking directly as an input the raw imaging data. Deep learning showed good performance for differentiating lung nodule from other thoracic structures (vessels, bone, …) (32–37). Particularly in a study, Causey et al. (38) processed 1065 nodules with different malignancy scores. The model was developed with a deep CNN architecture, capable of performing classification or producing a feature vector that could then be used as input to a secondary classifier such as a RFC. The CNN classification highly performed (AUC 0.97) and was improved (AUC 0.99) when combined to handcrafted radiomics features (38) through a RFC. The main studies relating to lung nodule classification are summarized in **Table 1**.

Major hope is that characterization of lung abnormalities could potentially allow for an early diagnosis of lung cancer, even for very small nodules, aiming to considerably improve patients' prognosis.

Histology and Radio-Genomics

When a suspicious lung abnormality is detected on imaging, obtaining histological evidence of cancer is necessary. It often requires an invasive procedure, sometimes leading to technical difficulties or complications. Thus, some patients, due to their state of health, are unable to undergo a biopsy.

Radiomics provide a promising alternative in this regard. From CT exam, radiomics features could be extracted to characterize tumor histology. From two independent cohorts, a Naïve Baye's classifier achieved a high AUC (0.72; p-value = 2.3×10^{-7}) with only five features (40).

Using ANN, similar performances for the prediction of histopathology were also obtained. Raniery Ferreira et al. (41) constructed different machine learning models for histopathological pattern recognition. From a dataset of 68 malignant lung tumors with confirmed histology, they extracted radiomics features by a semi-automatically segmentation. The radial basis function-based (RBF) ANN obtained an AUC of 0.71 on histopathological pattern recognition with radiomics features. In this study, adding clinical to radiomics features provided different behaviors on the models' performances on the testing and validation sets, and did not improve the results.

This last decade, targeted treatments played a leading role in lung cancer management (42). For most of those treatments, the identification of a specific mutation requires an invasive biopsy of the tumor, not always performable thus potentially depriving these patients of highly beneficial treatment. A more recent alternative could be liquid biopsy, consisting of the search of mutations on circulating tumor cells or DNA by a blood sample. Liquid biopsy has recently demonstrated its clinical usefulness in advanced NSCLC but keeps very poor sensitivity in early stage lung tumors (43, 44). The most common gene mutations seen in non-small-cell lung carcinoma (NSCLC) are V-Ki-ras2, Kirsten rat sarcoma viral oncogene homolog (KRAS), epidermal growth factor receptor (EGFR), v-raf murine sarcoma viral oncogene homolog B1 (BRAF), and anaplastic lymphoma kinase (ALK); of these, KRAS and EGFR mutations are the most commonly detected.

Regarding the specific mutation identification, the association of clinical features with radiomics ones seems to provide added value. Zhang et al. (45) conducted a multivariate analysis using seven handcrafted radiomics and three clinical features of 180 cases. They predicted EGFR mutation with an AUC of 0.87. Another study (46) explored a multicentric CT dataset of 381

TABLE 1 | Mains studies regarding lung nodule prediction of malignancy.

Reference	Number of cases	Imaging modality	Algorithm	Segmentation	Feature types	No of features	Validation	Results
Hawkins et al. (25)	598	CT	RFC	Semi-automatically segmented	Shape ++, 1st order	23	Cross-validation	AUC 0.83 at 1 year
Balagurunathan et al. (24)	479 (244 for Training)	CT	Linear classifier	Semi-automatically segmented	Shape, 1st order, 2nd order	4	Split sample	AUC 0.83
Wang et al. (28)	593 (400 for Training)	CT	SVM	Semi-automatically segmented	Shape, 1st order, 2nd order	15	Split sample	Accuracy 86%
Chen et al. (39)	72	CT	SVM	Manually segmented	Shape, 1st order, 2nd order	4	Cross-validation	Accuracy 84%
Dilger et al. (31)	50	CT	ANN	Manually segmented + surrounding lung parenchyma	Shape, 1st order, 2nd order	5	Cross-validation	AUC 0.938
Causey et al. (38)	1065	CT	CNN + RFC	Semi-Automatic + manually segmented radiomics	Deep features	NE	Split sample	AUC 0.99

ANN, artificial neural network; AUC, area under the curve; CNN, convolutional neural network; CT, computed tomography; NE, not evaluable; PSO, particle swarm optimization; RFC, random forest classifier; SVM, support vector machine.

patients who underwent surgical resection. The 20 remaining radiomics features using a RFC outperformed good prediction in discriminating between EGFR+ and EGFR- tumors (AUC 0.69). A clinical model of EGFR status (AUC 0.70) was combined to significantly improve prediction accuracy (AUC 0.75). The highest performing signature was capable of distinguishing between EGFR+ and KRAS+ cases (AUC 0.80) and, when combined with a clinical model (AUC 0.81), substantially improved its performance (AUC 0.86). One study by Zhao et al. (47) aimed at predicting EGFR mutation status and subtypes, in particular the two most common ones (exon 19 deletion and exon 21 L858R mutations). A radiomics score (R-score) based on 11 radiomics features was calculated for each lesion. Using a radiomics-based model and a combined radiomics and clinical model, the respective AUC values in the validation cohort were 0.73 and 0.76.

Deep learning methods have also been explored in prediction of genomic alterations. Using a CNN-based approach, Wang et al. (48), by training a network on 14926 CT images from 603 patients, achieved encouraging predictive performance on a validation cohort of 241 patients (AUC 0.81). For applying the deep learning model, a cubic region of interest (ROI) containing the entire tumor was manually selected. The first 20 convolutional layers were trained using transfer learning by 1.28 million natural images from the ImageNet dataset avoiding as much as possible an overfitting and the last four convolutional layers were trained using CT images from lung adenocarcinoma tumors in the independent test cohort. Authors used a method to visualize tumor region that was most related to EGFR mutation status.

While these studies focused on CT-based radiomics, another imaging modality commonly used in oncology is PET-CT. In PET-based radiomics, radiomics features could detect EGFR mutation status with good performance. Zhang et al. (49) developed a radiomics signature made of 10 features (PET and CT radiomics features) trained on 175 patients. The model showed a significant ability to discriminate between EGFR mutation and EGFR wild type in the validation set (AUC 0.85), which was improved when combined with clinical variables (AUC 0.87).

Yamamoto et al. (50) aimed instead at predicting the ALK status using visual qualitative CT features combined with clinical parameters. Their predictive model had a good performance in both the training and the validation set. Another study including clinical and radiomics variables extracted from PET and CT (51) from 539 patients with confirmed lung adenocarcinomas investigated the potential of differentiating the ALK/ROS1/RET fusion-positive and fusion-negative adenocarcinomas, building a model that resulted in 73% sensitivity and 70% specificity with seven features.

The main studies dealing with histologic and radio-genomics prediction are summarized in **Table 2**.

Treatment Outcome

Radiomics could play a role in predicting the prognosis and the treatment response, in order to adapt treatment strategies individually with view of personalized medicine. The main studies relating to this subject are summarized in **Table 3**.

Radiotherapy

In locally advanced lung cancer, radiotherapy, often associated with systemic therapies, is the standard option. A specific radiation option of lung cancer treatment is stereotactic body radiation therapy (SBRT), in inoperable patients presenting with a small local lesion (66). Radiosensitivity varies to a great extent across tumor types and also between patients bearing the same type of tumor. Biomarkers predicting the clinical outcome after radiotherapy are already available, but their levels of evidence are heterogeneous (67).

Radiomics features could be leveraged to predict different outcomes that conventional imaging metrics cannot predict in SBRT patients (68).

Several studies tried to predict different clinical endpoints such as local control and/or disease free survival and/or overall survival (52–58, 61) with good accuracy. Some others attempted to predict radiation induced toxicity (69), in particular to differentiate local failure from radiation induced lung injury (RILI) (59, 70).

Many of those studies outperformed different models concomitantly. Those studies revealed that a same feature selection technique and/or a same classifier model could considerably perform differently in distinct cohorts, suggesting a dependency not on the endpoint but on the study population.

The number of selected features is also notably heterogeneous between the studies from two (52) to fifteen radiomics features (53). After different feature selection methods, the texture features (i.e. second-order radiomics features) seemed to be the more correlated to clinical endpoints (53, 55, 71). Aiming to reduce the number of radiomics features, Diassaux et al. (52) found, in a multicentric study including 87 patients with an independent test set, a radiomics signature combining one PET feature and one CT feature predicting local control with an accuracy of 98%. They used ComBat harmonization method (72) on radiomics features to handle the differences of imaging acquisition. This method was initially used in gene expression microarray data to deal with the "batch effect," i.e., the source of variations in measurements caused by handling of samples by different laboratories, tools and technicians. The advantage of this technique is that it allows a correction to be applied directly to the extracted radiomic variables as opposed to the images before extraction, making it easier to analyze retrospective and multicentric data.

In radiation oncology, total dose and space dose distribution are carefully evaluated for each patient during treatment planning. In that way, a study (62) queried the lung CT-derived feature space to identify radiation sensitivity parameters that can predict treatment failure and hence guide the individualization of radiotherapy dose. The authors input pre-therapy lung CT images into Deep Profiler, a multitask deep neural network that has radiomics incorporated into the training process. Then, they combined these data with clinical variables to derive *i*Gray, an individualized radiation dose that results in an

TABLE 2 | Mains studies regarding histology and radio-genomic characterization.

Reference	Application	Number of cases	Imaging modality	Algorithm	Segmentation	Feature types	No of features	Validation	Results
Histology subtypes									
Wu et al. (40)	Prediction of histology subtype	350 (198 for Training)	CT	Naïve Baye's classifier	Manually segmented	Shape, 1st order, 2nd order	5	Independent	AUC 0,72
Raniery Ferreira et al. (41)	Prediction of histology subtype	68 (52 for Training)	CT	RBF-based ANN	Semi-Automatically segmented	Shape, 1st order, 2nd order	100	Sample split	AUC 0,71
Genomic alterations									
Zhang et al. (45)	Prediction of EGFR mutation	180 (140 for Training)	CT	multivariate analysis	Manually segmented	Clinical, Shape, 1st order, 2nd order	7	Sample split	AUC 0,87
Velazquez et al. (46)	Prediction of EGFR and KRAS mutation	381 (190 for Training)	CT	RFC	Manually segmented	Clinical, Shape, 1st order, 2nd order	25	Independent	AUC 0,86
Zhao et al. (47)	Prediction of EGFR subtype	637 (322 for Training)	CT	multivariate analysis	Manually segmented	Clinical, Shape, 1st order, 2nd order	11	Sample split	AUC 0,76
Wang et al. (48)	Prediction of EGFR mutation	843 (603 for Training)	CT	CNN	Manual segmentation	Deep features	NE	Independent	AUC 0,81
Zhang et al. (49)	Prediction of EGFR mutation	248 (175 for Training)	PET, CT	Logistic regression	Semi-Automatically segmented	Clinical, Shape, 1st order, 2nd order	13	Sample split	AUC 0,87
Yoon et al. (51)	Prediction of ALK status	539	PET, CT	Logistic regression	Semi-Automatically segmented	Clinical, Shape, 1st order, 2nd order	7	Cross validation	sensitivity and specificity, 0.73 and 0.70, respectively

ALK, anaplastic lymphoma kinase; ANN, artificial neural network; AUC, area under the curve; CNN, convolutional neural network; CT, computed tomography; EGFR, epidermal growth factor receptor, KRAS, Kirsten rat sarcoma viral oncogene homolog, NE, not evaluable; PET, positron emission tomography; RBF, radial basis function; SVM, support vector machine.

estimation of failure probability below 5% at 24 months. Thus, it would seem that a reduction in the irradiation dose could have been proposed in 23.3% of patients.

Integration of reported dosimetric features from the dose distribution in the irradiated lung calculated in the planning CT, showed to be predictive of radiation pneumonitis (73). Liang et al. (60) used the "dosiomics" method, which attempts to extract the spatial features from dose distribution, for the occurrence of grade 2 or more RP prediction.

To assist the physician during treatment planning, visualization of high-risk tumor spot of treatment failure could be very convenient. In a study (74), the authors visualized which regions in the patient images predicted low survival probability. From such observations, the heat map visualization has the potential to identify regions at high risk for tumor progression or recurrence that could be utilized for the purpose of assisting patient-tailored treatment planning in the future.

During radiation therapy treatment and follow-up, patients are subject to several imaging procedures. Like blood circulating biomarker changes during treatment could be predictive of the effectiveness of some treatments (75), the question of radiomics features modification has been studies, called "delta radiomics." It aims to analyze radiomics features' evolution through time and treatments based of evaluations obtained from longitudinal scans. Some studies demonstrated that delta radiomics seem to be more robust than radiomics features with the potential of using delta features for early assessment of treatment response and developing tailored therapies (76). A study focusing on 107 patients with stage III NSCLC (77) tried to evaluate the impact of radiomics features changes due to radiation therapy and their values at the end of treatment on tumor response. All of the radiomics features changed significantly during radiation therapy. For local recurrence, pretreatment imaging features were not prognostic, while texture-strength measured at the end of treatment significantly stratified high- and low-risk patients.

Another study focused on Cone Beam CT (CBCT), commonly used in radiotherapy for patient's precise setup, In this study, delta radiomics revealed to be predictive of overall survival in locally advanced lung cancer in a preliminary study with 23 patients (78).

In a study (79) including 268 patients with stage III NSCLC and using different CT at different timepoints of the treatment (pre-treatment, at 1, 3, and 6 months of follow-up), a deep learning networks was built to predict clinical outcomes of patients. Model performance was enhanced with each additional follow-up scan into the CNN model (2-year overall survival: AUC 0.74, $p < 0.05$).

In terms of toxicity prediction, Moran et al. (71) in a study with 14 patients who underwent SBRT tried to demonstrate the

TABLE 3 | Main studies evaluating radiomics in prediction of treatment outcomes in lung cancer.

Reference	Application	Number of cases	Imaging modality	Feature selection method	Model algorithm	Segmentation	Feature type	No. of features	Validation	Results
Radiotherapy										
Dissaux et al. (52)	Local control after SBRT	87 (64 for Training)	CT – PET/ CT	Univariate analysis	Multivariate regression	Semi-automatically + manually	1st order, 2nd order	2 (PET)	Independent set	Accuracy 0.91
Huynh et al. (53)	Outcomes after SBRT	113	CT	PCA	Concordance index	Manually	Clinical	15	Cross-validation	C-index of 0.33 for OS(q = 0.0016)
Zhang et al. (54)	Outcomes after SBRT	112	CT	PCA	RFC	Manually	1st order, 2nd order	NA	NA	OS: AUC 0,77
Yu et al. (55)	Outcome of stage I NSCLC	442 (147 for Training)	CT	Random Survival Forest	Multivariate regression	Manually	1st order, 2nd order	2	Independent set	OS: log-rank p=0.0173; HR 1.02, p= 0.0438
Hawkins et al. (56)	Outcome of NSCLC	81	CT	Relief-f	Decision tree	Manually	Shape, 1st order, 2nd order	5	Cross-validation	Accuracy 0.78
Aerts et al. (57)	OS of NSCLC and H&N cancer	1019 (474 for Training)	CT	Univariate analysis	Multivariate regression	Manually	Shape, 1st order, 2nd order	4	Independent set	C-index 0.65
Hosny et al. (58)	OS outcome of stage I and II NSCLC	1194 (786 for Training)	CT	NE	CNN	Manually	Deep features	NE	Independent set	AUC 0.71 and 0.70 for radiotherapy and surgery sets
Mattonen et al. (59)	Differentiate early recurrence from RILI post SBRT	45	CT at 3 months post SBRT	LOOCV	SVM	Semi-automatically	1st order, 2nd order	5	Cross-validation	AUC 0.85
Liang et al. (60)	Prediction of radiation pneumonitis	70	CT with dose distribution	Multivariate regression	Multivariate regression	Automatically	2nd order	2	None	AUC 0,78
Coroller et al. (61)	Predict pathological response after chemoradiation	127	CT	PCA	Multivariate regression	Manually	Clinical, Shape, 1st order, 2nd order	10	Cross-validation	AUC 0.68
Lou et al. (62)	Local control after SBRT	944 (849 for Training)	CT	NE	CNN	Manually	Deep features, clinical (dose)	NE	Independent set	C-index 0.77
Systemic treatment										
Khorrami et al. (63)	Response to 1st line chemotherapy	125 (53 for Training)	CT	LASSO	QDA	Manually	Shape, 2nd order	7	Split sample	AUC 0.77
Kim et al. (64)	Response to 1st line EGFR TKI	48	CT	Univariate analysis	Multivariate regression	Manually	Clinical, Shape, 1st order, 2nd order	5	None	C-index 0.77
Sun et al. (65)	Outcome anti-PD-1 and anti-PD-L1 treatment	272 (135 for Training)	CT	Elastic-net regularized regression	Elastic-net regularized regression	Semi-automatically	Location, technical, Shape, 1st order, 2nd order	8	Independent set	OS : HR 0.52; p=0.0022

AUC, area under the curve; CNN, convolutional neural network; CT, computed tomography; LASSO, least absolute shrinkage and selection operator; LOOCV, leave-one-out cross validation; NE, not evaluable; OS, overall survival; PCA, Principle Component Analysis; PET, positron emission tomography; QDA, Quadratic discriminant analysis; RFC, Random Forest Classifier; SVM, support vector machine.

potential of CT-based radiomics on 3, 6 and 9 month post-SBRT CT to distinguish moderate/severe lung injury from none/mild lung injury. Texture features outperformed the first-order features in differentiating lung injury severity levels.

After Systemic Treatments

While early-stage lung cancer patients with large tumors (stage IB-IIA) who have undergone surgery are likely to receive adjuvant chemotherapy (68), inoperable patients or patients presenting locally advanced lung cancer often have co-morbidities that limit their tolerance to systemic treatment. Consequently, systemic treatments cannot be a generalizable recommendation for all patients. The advent of immunotherapy and targeted therapies over the last decade in the management of metastatic lung cancer has led to important clinical results with a very acceptable safety profile (80, 81). It is therefore more than necessary to be able to determine in advance which patients are at risk of not responding to therapy and thus

allow either an intensification of the therapeutic strategy or of the therapeutic sequence or, conversely, avoid harmful therapies without benefit to the patient.

Radiomics showed good hope to be able to respond to this issue. In a study (63) including patients who were treated with front-line platinum-based chemotherapy, the combination of the top seven discriminating features outperformed an accuracy of 0.77 in prediction of tumor response. A significant correlation with both time to progression and overall survival for patients with NSCLC was also found.

Radiomics models could identify responders to EGFR tyrosine kinase inhibitor (TKIs) such as Gefitinib from the change in features between the pre-treatment and 3 weeks post-treatment CT. In a study conducted by Aerts et al. (82) including 47 patients, one delta-radiomics feature was significantly predictive (AUC 0.74) of Gefitinib response.

Pretreatment contrast-enhanced CT and first follow-up CT after initiation of EGFR TKIs were retrospectively analyzed in 48 NSCLC patients (64).

A recent promising treatment is immune checkpoint blocker (i.e. immunotherapy). The choice of patients who would benefit most from this treatment remains unclear and it is necessary to identify the good responders. Radiomics should have a role to play in this purpose (83).

Similarly, PET and PET/CT have been used for the prediction of treatment response. Radiomics signature was successfully validated to discriminate immune phenotype and predict survival and response to anti-PD-1 or PD-L1 immunotherapy (65, 84, 85).

Regarding the treatment sides effects, radiomics has been proven to be able to predict pneumonitis following immunotherapy (86), allowing closer surveillance for at-risk patients or even impacting the therapeutic choice.

DISCUSSION

This last decade, studies about radiomics drastically increased in different domains of oncology (87) with significant improvements. The new paradigm of precision medicine supports the research of new biomarkers and thus a lot of studies tried to explore radiomics in various applications with promising results which could have a huge impact on clinical routine. Machine and deep learning algorithms provide powerful modeling tools to explore the big amount of image data available, especially in oncology, to bring to light underlying complex biological mechanisms, and make personalized precision cancer diagnosis and treatment planning possible.

We could imagine a CAD, based on imaging, that directly establishes the nature of a lung lesion, its genomic alterations and provide guidance to physicians to choose the best therapeutic options that fit the most for each patient. Thus, the time between diagnosis and treatment initiation could be considerably reduced as well as the invasiveness of the procedures in patients who are, in most cases, very fragile. Patients could be offered therapeutics that are as effective as low in toxicity. For instance, SBRT, which is a first-choice treatment option for patients with stage I lung cancer who have surgical contraindications (88), could be proposed more broadly on the condition that the patients who can benefit from it could be accurately identified.

Then, in analogy to genomic signatures in breast cancer (89), therapeutic de-escalation may be possible when treatment would be identified as bringing no over gain.

Traditionally, the radiomics features being extracted are hand-crafted. Feature-based methods require a segmentation of the region of interest through a manual, semiautomated, or automatic methods. Then, hundreds or even thousands of radiomics features are extracted. Thus, feature selection and extraction are crucial steps that aim at obtaining the optimal feature representation that correlates most with the endpoint and correlates least between each other. Hand-crafted features suffer from the tedious designing process and may not faithfully capture the underlying imaging information. Semiautomatic segmentation could improve the stability of radiomics features (8) and fully automatic segmentation tools could be as accurate as manual segmentation by medical experts (90). With the development of deep learning based on multilayer neural networks, particularly CNN, the extraction of machine learnt features is becoming widely applicable. In deep learning, the processes of data representation and prediction are performed jointly (91). Pixel/voxel-based ML (PML) emerged in medical image analysis (92), which use pixel/voxel values in images directly instead of features calculated from segmented objects as input information; thus, feature calculation or segmentation is not required. Because the PML can avoid errors caused by inaccurate feature calculation and segmentation, the performance of the PML can potentially be higher than common classifiers. Moreover, the data representation removes the feature selection portion eliminating associated statistical bias in the process. The peritumoral space around the tumor may also provide valuable information over the visible tumor features for patient risk stratification due to cancer metastasis as demonstrated in a study carried by Dou et al. on 200 patients (93). A SVM classifier predicted distant failure with an accuracy of 0.83 thanks to analysis of the peritumoral space radiomics features from PET images of 48 NSCLC patients and 52 cervical cancer patients (94), arguing the fact that information around the tumor could provide better accuracy. PML are generally taking into account the peritumoral space. In a study evaluating a CNN based model (58), the visual mapping demonstrated that tissue within and beyond the tumor were both crucial for characterization and eventual prediction. CAD could be so able to highlight specific spot to overtreat.

It is clear that to this day, daily clinical radiomics applications remains very limited (95). At the present time, no clinical application of radiomics is available. Many factors could explain this situation (96).

First, the overall scientific quality and reporting of radiomics studies is insufficient. Scientific improvements need to be made to feature reproducibility, analysis of clinical utility, and open science categories. The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

(TRIPOD) checklist (97) was adapted to radiomics studies by Park et al. (98) after finding very poor results of his analysis of multiple studies in term of radiomics quality scores and adherence to the TRIPOD checklist. It intends to improve the transparency of a prediction model study's reporting regardless of the study methods. It is a checklist of 22 items considered important for good reporting of studies developing or validating multivariable prediction models. The items relate to the title and abstract, background and objectives, methods, results, discussion, and other information. The TRIPOD Statement covers studies that report solely development, both development and external validation, and solely external validation (with or without model updating) of a diagnostic or prognostic prediction model. Recently, a Checklist for Artificial Intelligence in Medical Imaging was proposed (99). In the batch of radiomics studies, few ones are able to provide clear details of the models and the selected predictors.

Moreover, reproducibility of radiomics features should be carefully explored. For instance, differences on imaging acquisition modalities could greatly influence radiomics features (100). Thus, harmonizing acquisition parameters between studies is a crucial step for future texture analysis (101). There is a real need for the harmonization of features to allow consistent findings in radiomics multicenter studies. Two main approaches could be considered to address this issue: harmonizing images and harmonizing radiomic features. The first one focuses on the harmonization issue in imaging and usually looks upon standardization of acquisition protocols and reconstruction settings, such as guidelines already available for PET/CT imaging (102). This approach should not be enough. Recently, techniques based on generative adversarial networks (103) have also been developed. Heterogeneous images are translated to match the statistical properties of a standard dataset, such as a template reference image. The second approach focuses on the issue in the feature area by either using prior feature selection based on their robustness, keeping only features insensitive to multicenter variability, or by keeping all features and harmonizing their statistical properties so they can be pooled during the modeling step. In this regard, different methods could be considered, such as normalization or batch-effect correction using the ComBat method.

One of the other challenges of imaging research is enhancing global collaboration and sharing trial data (104). Big and standardized clinical data will make radiomics clinically applicable (105). Access to big data is needed, as medical images are dispersed in different hospitals or data centers. Data sharing among institutes and hospitals is important for radiomics, although it presents complex logistical problems. The Cancer Imaging Archive (TCIA) provides a good example of data sharing with a large portion of clinical information (106).

To perform generalizable models, it will be mandatory to develop them by involving multiple centers and to improve national and international collaboration (107).

Patient medical records are a great source of data. Some studies have shown the added value of clinical features combined with clinical ones. We can also hypothesize that the adjunction of genomic data and radiomics features from different imaging modalities could permit to get closer to a more personalized medicine.

The field is certainly high on promise and relatively low on data and proof, with the need of prospective validation (108). For clinical application, higher evidence levels are important. Prospective, multicenter, randomized controlled trials studies are needed.

One critical aspect of the radiomics workflow that remains relatively unexamined is the implementation of the software platforms used to calculate radiomics features. Some studies have demonstrated features variability from different software platforms (109, 110). The Image Biomarker Standardisation Initiative (IBSI), an international collaboration, was developed to help standardize radiomics feature calculation and has provided a framework to deliver practical solutions (111). The IBSI has made recommendations concerning feature calculation, standardized feature definition and nomenclature. A study (112) demonstrated the benefits of standardizing feature calculation platforms according to the IBSI with greater statistical reliability, but only when calculation settings were also harmonized.

Another point which should carefully assessed is the integration of radiomics software in the job process. Optimization, effectiveness and utility should be evaluated. As digital assistants (113), software programs designed to interact with people in a conversational manner, radiomics based software impact on clinical routine need assessment. By the same way, human factors (114) should be more consider as human factor interventions are known to have great potential to contribute to efficient Healthcare Information Technology design. Human factors and human-centered design play a critical role in ensuring that health IT is well designed and fits with clinical and patient workflows. The gaps between stakeholders, particularly vendors, researchers, clinicians, healthcare organization administrations, and purchasers, need also to be reduced.

CONCLUSION

Radiomics in this last decade shows good ability to be considered as a potential new biomarker at different steps of the patient's care in lung cancer. More multicentric prospective studies are still needed to evaluate the application of radiomics in daily practice. Deep learned radiomics should replace the traditional handcrafted radiomics for more efficiency on large datasets and more reproducibility.

AUTHOR CONTRIBUTIONS

RE wrote the manuscript. JB helped for article selection and manuscript redaction. JB, AB, NG, and PaG helped for technical review of the manuscript for machine learning and radiomics aspects. CD and PhG performed a review of the manuscript on the clinical aspects. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Patz EF, Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, et al. Overdiagnosis in Low-Dose Computed Tomography Screening for Lung Cancer. *JAMA Intern Med* (2014) 174(2):269–74. doi: 10.1001/jamainternmed.2013.12738
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2018. *CA Cancer J Clin* (2018) 68(1):7–30. doi: 10.3322/caac.21442
3. Lee CK, Brown C, Gralla RJ, Hirsh V, Thongprasert S, Tsai C-M, et al. Impact of EGFR Inhibitor in non-Small Cell Lung Cancer on Progression-Free and Overall Survival: A Meta-Analysis. *J Natl Cancer Inst* (2013) 105 (9):595–605. doi: 10.1093/jnci/djt072
4. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer Oxf Engl 1990* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
5. Zhao B, Tan Y, Tsai W-Y, Qi J, Xie C, Lu L, et al. Reproducibility of Radiomics for Deciphering Tumor Phenotype With Imaging. *Sci Rep* (2016) 6:23428. doi: 10.1038/srep23428
6. Balagurunathan Y, Kumar V, Gu Y, Kim J, Wang H, Liu Y, et al. Test-Retest Reproducibility Analysis of Lung CT Image Features. *J Digit Imaging* (2014) 27(6):805–23. doi: 10.1007/s10278-014-9716-x
7. Rios Velazquez E, Aerts HJWL, Gu Y, Goldgof DB, De Ruysscher D, Dekker A, et al. A Semiautomatic CT-based Ensemble Segmentation of Lung Tumors: Comparison With Oncologists' Delineations and With the Surgical Specimen. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2012) 105(2):167–73. doi: 10.1016/j.radonc.2012.09.023
8. Parmar C, Velazquez ER, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust Radiomics Feature Quantification Using Semiautomatic Volumetric Segmentation. *PloS One* (2014) 9(7):e102107. doi: 10.1371/journal.pone.0102107
9. Chakraborty S, Hosen MI, Ahmed M, Shekhar HU. Onco-Multi-OMICS Approach: A New Frontier in Cancer Research. *BioMed Res Int* (2018) 2018:9836256. doi: 10.1155/2018/9836256
10. O'Connor JPB, Rose CJ, Waterton JC, Carano RAD, Parker GJM, Jackson A. Imaging Intratumor Heterogeneity: Role in Therapy Response, Resistance, and Clinical Outcome. *Clin Cancer Res Off J Am Assoc Cancer Res* (2015) 21 (2):249–57. doi: 10.1158/1078-0432.CCR-14-0990
11. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N Engl J Med* (2012) 366(10):883–92. doi: 10.1056/NEJMoa1113205
12. Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational Implications of Tumor Heterogeneity. *Clin Cancer Res Off J Am Assoc Cancer Res* (2015) 21(6):1258–66. doi: 10.1158/1078-0432.CCR-14-1429
13. Janiszewska M. The Microcosmos of Intratumor Heterogeneity: The Space-Time of Cancer Evolution. *Oncogene* (2020) 39(10):2031–9. doi: 10.1038/s41388-019-1127-5
14. Bibault J-E, Giraud P, Burgun A. Big Data and Machine Learning in Radiation Oncology: State of the Art and Future Prospects. *Cancer Lett* (2016) 382(1):110–7. doi: 10.1016/j.canlet.2016.05.033
15. Liang M, Tang W, Xu DM, Jirapatnakul AC, Reeves AP, Henschke CI, et al. Low-Dose CT Screening for Lung Cancer: Computer-aided Detection of Missed Lung Cancers. *Radiology* (2016) 281(1):279–88. doi: 10.1148/radiol.2016150063
16. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET, et al. Lung Cancer Probability in Patients With CT-detected Pulmonary Nodules: A Prespecified Analysis of Data From the NELSON Trial of Low-Dose CT Screening. *Lancet Oncol* (2014) 15 (12):1332–41. doi: 10.1016/S1470-2045(14)70389-4
17. Carter SM, Barratt A. What is Overdiagnosis and Why Should We Take it Seriously in Cancer Screening? *Public Health Res Pract 26 juill* (2017) 27(3). doi: 10.17061/phrp2731722
18. Zhao YR, van Ooijen PMA, Dorrius MD, Heuvelmans M, de Bock GH, Vliegenthart R, et al. Comparison of Three Software Systems for Semi-Automatic Volumetry of Pulmonary Nodules on Baseline and Follow-Up CT Examinations. *Acta Radiol Stockh Swed 1987* (2014) 55(6):691–8. doi: 10.1177/0284185113508177
19. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. Fdg PET/CT: EANM Procedure Guidelines for Tumour Imaging: Version 2.0. *Eur J Nucl Med Mol Imaging* (2015) 42:328–54. doi: 10.1007/s00259-014-2961-x
20. Salaün P-Y, Abgral R, Malard O, Querellou-Lefranc S, Quere G, Wartski M, et al. Good Clinical Practice Recommendations for the Use of PET/CT in Oncology. *Eur J Nucl Med Mol Imaging* (2020) 47(1):28–50. doi: 10.1007/s00259-019-04553-8
21. Hofman P. Liquid Biopsy for Early Detection of Lung Cancer. *Curr Opin Oncol* (2017) 29(1):73–8. doi: 10.1097/CCO.0000000000000343
22. Wilson R, Devaraj A. Radiomics of Pulmonary Nodules and Lung Cancer. *Transl Lung Cancer Res* (2017) 6(1):86–91. doi: 10.21037/tlcr.2017.01.04
23. Rubin GD, Lyo JK, Paik DS, Sherbondy AJ, Chow LC, Leung AN, et al. Pulmonary Nodules on Multi–Detector Row CT Scans: Performance Comparison of Radiologists and Computer-aided Detection. *Radiology* (2005) 234(1):274–83. doi: 10.1148/radiol.2341040589
24. Balagurunathan Y, Schabath MB, Wang H, Liu Y, Gillies RJ. Quantitative Imaging Features Improve Discrimination of Malignancy in Pulmonary Nodules. *Sci Rep* (2019) 9(1):1–14. doi: 10.1038/s41598-019-44562-z
25. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting Malignant Nodules From Screening Cts. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer* (2016) 11(12):2120–8. doi: 10.1016/j.jtho.2016.07.002
26. Talwar A, Rahman NM, Kadir T, Pickup LC, Gleeson F. A Retrospective Validation Study of Three Models to Estimate the Probability of Malignancy in Patients With Small Pulmonary Nodules From a Tertiary Oncology Follow-Up Centre. *Clin Radiol fevr* (2017) 72(2):177.e1–8. doi: 10.1016/j.crad.2016.09.014
27. Godoy MCB, Odisio EGLC, Truong MT, de Groot PM, Shroff GS, Erasmus JJ. Pulmonary Nodule Management in Lung Cancer Screening: A Pictorial Review of Lung-RADS Version 1.0. *Radiol Clin North Am* (2018) 56(3):353–63. doi: 10.1016/j.rcl.2018.01.003
28. Wang J, Liu X, Dong D, Song J, Xu M, Zang Y, et al. Prediction of Malignant and Benign of Lung Tumor Using a Quantitative Radiomic Method. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Embc)* (2016). p. 1272–5. doi: 10.1109/EMBC.2016.7590938
29. Zhang T, Yuan M, Zhong Y, Zhang Y-D, Li H, Wu J-F, et al. Differentiation of Focal Organising Pneumonia and Peripheral Adenocarcinoma in Solid Lung Lesions Using Thin-Section CT-based Radiomics. *Clin Radiol* (2019) 74(1):78.e23–30. doi: 10.1016/j.crad.2018.08.014
30. Wu W, Pierce LA, Zhang Y, Pipavath SNJ, Randolph TW, Lastwika KJ, et al. Comparison of Prediction Models With Radiological Semantic Features and Radiomics in Lung Cancer Diagnosis of the Pulmonary Nodules: A Case-Control Study. *Eur Radiol* (2019) 29(11):6100–8. doi: 10.1007/s00330-019-06213-9
31. Dilger SK, Uthoff J, Judisch A, Hammond E, Mott SL, Smith BJ, et al. Improved Pulmonary Nodule Classification Utilizing Quantitative Lung Parenchyma Features. *J Med Imaging* (2015) 2(4):041004. doi: 10.1117/1.JMI.2.4.041004
32. da Silva GLF, Valente TLA, Silva AC, de Paiva AC, Gattass M. Convolutional Neural Network-Based PSO for Lung Nodule False Positive Reduction on CT Images. *Comput Methods Programs BioMed* (2018) 162:109–18. doi: 10.1016/j.cmpb.2018.05.006
33. Li W, Cao P, Zhao D, Wang J. Pulmonary Nodule Classification With Deep Convolutional Neural Networks on Computed Tomography Images. *Comput Math Methods Med* (2016) 2016:6215085. doi: 10.1155/2016/6215085
34. Shen W, Zhou M, Yang F, Yang C, Tian J. Multi-Scale Convolutional Neural Networks for Lung Nodule Classification. *Inf Process Med Imaging Proc Conf* (2015) 24:588–99. doi: 10.1007/978-3-319-19992-4_46
35. Kumar D, Wong A, Clausi DA. Lung Nodule Classification Using Deep Features in CT Images. In: *2015 12th Conference on Computer and Robot Vision* (2015). p. 133–8. doi: 10.1109/CRV.2015.25
36. Golan R, Jacob C, Denzinger J. Lung Nodule Detection in CT Images Using Deep Convolutional Neural Networks. In: *2016 International Joint Conference on Neural Networks (Ijcnn)* (2016). p. 243–50. doi: 10.1109/IJCNN.2016.7727205

37. Ciompi F, Chung K, van Riel SJ, Setio AAA, Gerke PK, Jacobs C, et al. Towards Automatic Pulmonary Nodule Management in Lung Cancer Screening With Deep Learning. *Sci Rep* (2017) 7(1):46479. doi: 10.1038/srep46479
38. Causey JL, Zhang J, Ma S, Jiang B, Qualls JA, Politte DG, et al. Highly Accurate Model for Prediction of Lung Nodule Malignancy With CT Scans. *Sci Rep* (2018) 8(1):9286. doi: 10.1038/s41598-018-27569-w
39. Chen C-H, Chang C-K, Tu C-Y, Liao W-C, Wu B-R, Chou K-T, et al. Radiomic Features Analysis in Computed Tomography Images of Lung Nodule Classification. *PloS One* (2018) 13(2):e0192002. doi: 10.1371/journal.pone.0192002
40. Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, et al. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front Oncol* (2016) 6:71. doi: 10.3389/fonc.2016.00071
41. Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, Fabro AT, de Azevedo-Marques PM. Radiomics-Based Features for Pattern Recognition of Lung Cancer Histopathology and Metastases. *Comput Methods Programs BioMed* (2018) 159:23–30. doi: 10.1016/j.cmpb.2018.02.015
42. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ, Wu Y-L, et al. Lung Cancer: Current Therapies and New Targeted Treatments. *Lancet Lond Engl* (2017) 389(10066):299–311. doi: 10.1016/S0140-6736(16)30958-8
43. Santarpia M, Liguori A, D'Aveni A, Karachaliou N, Gonzalez-Cao M, Daffinà MG, et al. Liquid Biopsy for Lung Cancer Early Detection. *J Thorac Dis* (2018) 10(Suppl 7):S882–97. doi: 10.21037/jtd.2018.03.81
44. Ye M, Li S, Huang W, Wang C, Liu L, Liu J, et al. Comprehensive Targeted Super-Deep Next Generation Sequencing Enhances Differential Diagnosis of Solitary Pulmonary Nodules. *J Thorac Dis avr* (2018) 10(Suppl 7):S820–9. doi: 10.21037/jtd.2018.04.09
45. Zhang L, Chen B, Liu X, Song J, Fang M, Hu C, et al. Quantitative Biomarkers for Prediction of Epidermal Growth Factor Receptor Mutation in Non-Small Cell Lung Cancer. *Transl Oncol* (2017) 11(1):94–101. doi: 10.1016/j.tranon.2017.10.012
46. Velazquez ER, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. *Cancer Res* (2017) 77(14):3922–30. doi: 10.1158/0008-5472.CAN-17-0122
47. Zhao W, Wu Y, Xu Y, Sun Y, Gao P, Tan M, et al. The Potential of Radiomics Nomogram in Non-invasively Prediction of Epidermal Growth Factor Receptor Mutation Status and Subtypes in Lung Adenocarcinoma. *Front Oncol* (2019) 9:1485. doi: 10.3389/fonc.2019.01485
48. Wang S, Shi J, Ye Z, Dong D, Yu D, Zhou M, et al. Predicting EGFR Mutation Status in Lung Adenocarcinoma on Computed Tomography Image Using Deep Learning. *Eur Respir J* (2019) 53(3). doi: 10.1183/13993003.00986-2018
49. Zhang J, Zhao X, Zhao Y, Zhang J, Zhang Z, Wang J, et al. Value of Pre-Therapy 18F-FDG PET/CT Radiomics in Predicting EGFR Mutation Status in Patients With non-Small Cell Lung Cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(5):1137–46. doi: 10.1007/s00259-019-04592-1
50. Yamamoto S, Korn RL, Oklu R, Migdal C, Gotway MB, Weiss GJ, et al. ALK Molecular Phenotype in non-Small Cell Lung Cancer: CT Radiogenomic Characterization. *Radiology* (2014) 272(2):568–76. doi: 10.1148/radiol.14140789
51. Yoon HJ, Sohn I, Cho JH, Lee HY, Kim J-H, Choi Y-L, et al. Decoding Tumor Phenotypes for ALK, ROS1, and RET Fusions in Lung Adenocarcinoma Using a Radiomics Approach. *Med (Baltimore)* (2015) 94(41):e1753. doi: 10.1097/MD.0000000000001753
52. Dissaux G, Visvikis D, Da-Ano R, Pradier O, Chajon E, Barillot I, et al. Pretreatment 18F-FDG PET/CT Radiomics Predict Local Recurrence in Patients Treated With Stereotactic Body Radiotherapy for Early-Stage non-Small Cell Lung Cancer: A Multicentric Study. *J Nucl Med Off Publ Soc Nucl Med* (2020) 61(6):814–20. doi: 10.2967/jnumed.119.228106
53. Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-Based Radiomic Analysis of Stereotactic Body Radiation Therapy Patients With Lung Cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2016) 120(2):258–66. doi: 10.1016/j.radonc.2016.05.024
54. Zhang Y, Oikonomou A, Wong A, Haider MA, Khalvati F. Radiomics-Based Prognosis Analysis for Non-Small Cell Lung Cancer. *Sci Rep* (2017) 7:46349. doi: 10.1038/srep46349
55. Yu W, Tang C, Hobbs BP, Li X, Koay EJ, Wistuba II, et al. Development and Validation of a Predictive Radiomics Model for Clinical Outcomes in Stage I non-Small Cell Lung Cancer. *Int J Radiat Oncol Biol Phys* (2018) 102(4):1090–7. doi: 10.1016/j.ijrobp.2017.10.046
56. Hawkins SH, Korecki JN, Balagurunathan Y, Gu Y, Kumar V, Basu S, et al. Predicting Outcomes of Nonsmall Cell Lung Cancer Using Ct Image Features. *IEEE Access* (2014) 2:1418–26. doi: 10.1109/ACCESS.2014.2373335
57. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach. *Nat Commun* (3) juin 20145(1):4006. doi: 10.1038/ncomms5006
58. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep Learning for Lung Cancer Prognostication: A Retrospective Multi-Cohort Radiomics Study. *PloS Med* (2018) 15(11):e1002711. doi: 10.1371/journal.pmed.1002711
59. Mattonen SA, Palma DA, Johnson C, Louie AV, Landis M, Rodrigues G, et al. Detection of Local Cancer Recurrence After Stereotactic Ablative Radiation Therapy for Lung Cancer: Physician Performance Versus Radiomic Assessment. *Int J Radiat Oncol Biol Phys* (2016) 94(5):1121–8. doi: 10.1016/j.ijrobp.2015.12.369
60. Liang B, Yan H, Tian Y, Chen X, Yan L, Zhang T, et al. Dosiomics: Extracting 3d Spatial Features From Dose Distribution to Predict Incidence of Radiation Pneumonitis. *Front Oncol* (2019) 9:269. doi: 10.3389/fonc.2019.00269
61. Coroller TP, Agrawal V, Narayan V, Hou Y, Grossmann P, Lee SW, et al. Radiomic Phenotype Features Predict Pathological Response in non-Small Cell Lung Cancer. *Radiother Oncol* (2016) 119(3):480–6. doi: 10.1016/j.radonc.2016.04.004
62. Lou B, Doken S, Zhuang T, Wingerter D, Gidwani M, Mistry N, et al. An Image-Based Deep Learning Framework for Individualising Radiotherapy Dose: A Retrospective Analysis of Outcome Prediction. *Lancet Digit Health* (2019) 1(3):e136–47. doi: 10.1016/S2589-7500(19)30058-5
63. Khorrami M, Khunger M, Zagouras A, Patil P, Thawani R, Bera K, et al. Combination of Peri- and Intratumoral Radiomic Features on Baseline Ct Scans Predicts Response to Chemotherapy in Lung Adenocarcinoma. *Radiol Artif Intell* (2019) 1(2):180012. doi: 10.1148/ryai.2019180012
64. Kim H, Park CM, Keam B, Park SJ, Kim M, Kim TM, et al. The Prognostic Value of CT Radiomic Features for Patients With Pulmonary Adenocarcinoma Treated With EGFR Tyrosine Kinase Inhibitors. *PloS One* (2017) 12(11):e0187500. doi: 10.1371/journal.pone.0187500
65. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A Radiomics Approach to Assess Tumour-Infiltrating CD8 Cells and Response to anti-PD-1 or anti-PD-L1 Immunotherapy: An Imaging Biomarker, Retrospective Multicohort Study. *Lancet Oncol* (2018) 19(9):1180–91. doi: 10.1016/S1470-2045(18)30413-3
66. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, et al. Early and Locally Advanced non-Small-Cell Lung Cancer (NSCLC): ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up. *Ann Oncol Off J Eur Soc Med Oncol* (2017) 28(suppl_4):iv1–21. doi: 10.1093/annonc/mdx222
67. Bibault J-E, Fumagalli I, Ferté C, Chargari C, Soria J-C, Deutsch E. Personalized Radiation Therapy and Biomarker-Driven Treatment Strategies: A Systematic Review. *Cancer Metastasis Rev* (2013) 32(3–4):479–92. doi: 10.1007/s10555-013-9419-7
68. Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-Based Radiomic Analysis of Stereotactic Body Radiation Therapy Patients With Lung Cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2016) 120(2):258–66. doi: 10.1016/j.radonc.2016.05.024
69. Krafft SP, Rao A, Stingo F, Briere TM, Court LE, Liao Z, et al. The Utility of Quantitative CT Radiomics Features for Improved Prediction of Radiation Pneumonitis. *Med Phys* (2018) 45(11):5317–24. doi: 10.1002/mp.13150
70. Mattonen SA, Palma DA, Haasbeek CJA, Senan S, Ward AD. Early Prediction of Tumor Recurrence Based on CT Texture Changes After Stereotactic Ablative Radiotherapy (SABR) for Lung Cancer. *Med Phys mars* (2014) 41(3):033502. doi: 10.1118/1.4866219
71. Moran A, Daly ME, Yip SSF, Yamamoto T. Radiomics-Based Assessment of Radiation-induced Lung Injury After Stereotactic Body Radiotherapy. *Clin Lung Cancer* (2017) 18(6):e425–31. doi: 10.1016/j.cllc.2017.05.014

72. Orlhac F, Boughdad S, Philippe C, Stalla-Bourdillon H, Nioche C, Champion L, et al. A Postreconstruction Harmonization Method for Multicenter Radiomic Studies in PET. *J Nucl Med Off Publ Soc Nucl Med* (2018) 59 (8):1321–8. doi: 10.2967/jnumed.117.199935
73. Chen S, Zhou S, Yin F-F, Marks LB, Das SK. Investigation of the Support Vector Machine Algorithm to Predict Lung Radiation-Induced Pneumonitis. *Med Phys* (2007) 34(10):3808–14. doi: 10.1118/1.2776669
74. Baek S, He Y, Allen BG, Buatti JM, Smith BJ, Tong L, et al. Deep Segmentation Networks Predict Survival of non-Small Cell Lung Cancer. *Sci Rep* (2019) 9(1):17286. doi: 10.1038/s41598-019-53461-2
75. Thery L, Meddis A, Cabel L, Proudhon C, Latouche A, Pierga J-Y, et al. Circulating Tumor Cells in Early Breast Cancer. *JNCI Cancer Spectr* (2019) 3 (2):pkz026. doi: 10.1093/jncics/pkz026
76. Nardone V, Reginelli A, Guida C, Belfiore MP, Biondi M, Mormile M, et al. Delta-Radiomics Increases Multicentre Reproducibility: A Phantom Study. *Med Oncol Northwood Lond Engl* (31) mars 202037(5):38. doi: 10.1007/s12032-020-01359-9
77. Fave X, Zhang L, Yang J, Mackin D, Balter P, Gomez D, et al. Delta-Radiomics Features for the Prediction of Patient Outcomes in non-Small Cell Lung Cancer. *Sci Rep* (03) 20177(1):588. doi: 10.1038/s41598-017-00665-z
78. Shi L, Rong Y, Daly M, Dyer B, Benedict S, Qiu J, et al. Cone-Beam Computed Tomography-Based Delta-Radiomics for Early Response Assessment in Radiotherapy for Locally Advanced Lung Cancer. *Phys Med Biol* (10) 202065(1):015009. doi: 10.1088/1361-6560/ab3247
79. Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, et al. Deep Learning Predicts Lung Cancer Treatment Response From Serial Medical Imaging. *Clin Cancer Res* (2019) 25(11):3266–75. doi: 10.1158/1078-0432.CCR-18-2495
80. Ettinger DS, Wood DE, Aggarwal C, Aisner DL, Akerley W, Bauman JR, et al. Nccn Guidelines Insights: Non-Small Cell Lung Cancer, Version 1.2020. *J Natl Compr Cancer Netw JNCCN* (2019) 17(12):1464–72. doi: 10.6004/jnccn.2019.0059
81. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab Versus Chemotherapy for PD-L1-Positive non-Small-Cell Lung Cancer. *N Engl J Med* (2016) 375(19):1823–33. doi: 10.1056/NEJMoa1606774
82. Aerts HJWL, Grossmann P, Tan Y, Oxnard GR, Rizvi N, Schwartz LH, et al. Defining a Radiomic Response Phenotype: A Pilot Study Using Targeted Therapy in NSCLC. *Sci Rep* (20) sept 20166(1):33860. doi: 10.1038/srep41197
83. Limkin EJ, Sun R. Radiomics to Predict Response to Immunotherapy: An Imminent Reality? *Future Oncol* (2020) 16(23):1673–6. doi: 10.2217/fon-2020-0015
84. Mu W, Qi J, Lu H, Schabath M, Balagurunathan Y, Tunali I, et al. Radiomic Biomarkers From PET/CT Multi-Modality Fusion Images for the Prediction of Immunotherapy Response in Advanced non-Small Cell Lung Cancer Patients. (2018) 0575:105753S. doi: 10.1117/12.2293376
85. Jiang M, Sun D, Guo Y, Guo Y, Xiao J, Wang L, et al. Assessing PD-L1 Expression Level by Radiomic Features From PET/CT in Nonsmall Cell Lung Cancer Patients: An Initial Result. *Acad Radiol* (2020) 27(2):171–9. doi: 10.1016/j.acra.2019.04.016
86. Colen RR, Fujii T, Bilen MA, Kotrotsou A, Abrol S, Hess KR, et al. Radiomics to Predict Immunotherapy-Induced Pneumonitis: Proof of Concept. *Invest New Drugs* (2018) 36(4):601–7. doi: 10.1007/s10637-017-0524-2
87. Bibault J-E, Xing L, Giraud P, El Ayachy R, Giraud N, Decazes P, et al. Radiomics: A Primer for the Radiation Oncologist. *Cancer Radiother J Soc Francaise Radiother Oncol* (2020) 24(5):403–10. doi: 10.1016/j.canrad.2020.01.011
88. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, et al. Early and Locally Advanced non-Small-Cell Lung Cancer (NSCLC): ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up†. *Ann Oncol* (2017) 28:iv1–21. doi: 10.1093/annonc/mdx222
89. Sestak I. Risk Stratification in Early Breast Cancer in Premenopausal and Postmenopausal Women: Integrating Genomic Assays With Clinicopathological Features. *Curr Opin Oncol* (2019) 31(1):29–34. doi: 10.1097/CCO.0000000000000490
90. Echegaray S, Gevaert O, Shah R, Kamaya A, Louie J, Kothary N, et al. Core Samples for Radiomics Features That are Insensitive to Tumor Segmentation: Method and Pilot Study Using CT Images of Hepatocellular Carcinoma. *J Med Imaging Bellingham Wash* (2015) 2(4):041011. doi: 10.1117/1.JMI.2.4.041011
91. LeCun Y, Bengio Y, Hinton G. Deep Learning. *Nature* (2015) 521 (7553):436–44. doi: 10.1038/nature14539
92. Suzuki K. Pixel-Based Machine Learning in Medical Imaging. *Int J BioMed Imaging* (2012) 2012:792079. doi: 10.1155/2012/792079
93. Dou TH, Coroller TP, van Griethuysen JJM, Mak RH, Aerts HJWL. Peritumoral Radiomics Features Predict Distant Metastasis in Locally Advanced NSCLC. *PloS One* (2018) 13(11):e0206108. doi: 10.1371/journal.pone.0206108
94. Hao H, Zhou Z, Li S, Maquilan G, Folkert MR, Iyengar P, et al. Shell Feature: A New Radiomics Descriptor for Predicting Distant Failure After Radiotherapy in non-Small Cell Lung Cancer and Cervix Cancer. *Phys Med Biol* (2018) 63(9):095007. doi: 10.1088/1361-6560/aabb5e
95. Ather S, Kadir T, Gleeson F. Artificial Intelligence and Radiomics in Pulmonary Nodule Management: Current Status and Future Applications. *Clin Radiol* (2020) 75(1):13–9. doi: 10.1016/j.crad.2019.04.017
96. Pinto Dos Santos D, Dietzel M, Baessler B. A Decade of Radiomics Research: Are Images Really Data or Just Patterns in the Noise? *Eur Radiol* (2021) ;31 (1):1–4. doi: 10.1007/s00330-020-07108-w
97. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *BMJ* (2015) 350:g7594. doi: 10.1136/bmj.g7594
98. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, et al. Quality of Science and Reporting of Radiomics in Oncologic Studies: Room for Improvement According to Radiomics Quality Score and TRIPOD Statement. *Eur Radiol* (2020) 30(1):523–36. doi: 10.1007/s00330-019-06360-z
99. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (Claim): A Guide for Authors and Reviewers. *Radiol Artif Intell* (1) mars 20202(2):e200029. doi: 10.1148/ryai.2020200029
100. Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, Castro-García M, Villas MV, Mansilla Legorburo F, et al. Radiomics of CT Features May be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. *Radiology* (2018) 288(2):407–15. doi: 10.1148/radiol.2018172361
101. Espinasse M, Pitre-Champagnat S, Charmettant B, Bidault F, Volk A, Balleyguier C, et al. Ct Texture Analysis Challenges: Influence of Acquisition and Reconstruction Parameters: A Comprehensive Review. *Diagnostics* (2020) 10(5):258. doi: 10.3390/diagnostics10050258
102. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, et al. Feasibility of State of the Art PET/CT Systems Performance Harmonisation. *Eur J Nucl Med Mol Imaging* (2018) 45(8):1344–61. doi: 10.1007/s00259-018-3977-4
103. Choe J, Lee SM, Do K-H, Lee G, Lee J-G, Lee SM, et al. Deep Learning-based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses. *Radiology* (2019) 292 (2):365–73. doi: 10.1148/radiol.2019181960
104. Dewey M, Bosserdt M, Dodd JD, Thun S, Kressel HY. Clinical Imaging Research: Higher Evidence, Global Collaboration, Improved Reporting, and Data Sharing Are the Grand Challenges. *Radiology* (2019) 291(3):547–52. doi: 10.1148/radiol.2019181796
105. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics* (2019) 9(5):1303–22. doi: 10.7150/thno.30309
106. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (Tcia): Maintaining and Operating a Public Information Repository. *J Digit Imaging* (2013) 26(6):1045–57. doi: 10.1007/s10278-013-9622-7
107. Foy J-P, Durdux C, Giraud P, Bibault J-E. Re: The Rise of Radiomics and Implications for Oncologic Management. *JNCI J Natl Cancer Inst* (2018) 110 (11):1275–6. doi: 10.1093/jnci/djy037
108. Topol EJ. High-Performance Medicine: The Convergence of Human and Artificial Intelligence. *Nat Med* (2019) 25(1):44. doi: 10.1038/s41591-018-0300-7
109. Liang Z-G, Tan HQ, Zhang F, Rui Tan LK, Lin L, Lenkowicz J, et al. Comparison of Radiomics Tools for Image Analyses and Clinical Prediction

in Nasopharyngeal Carcinoma. *Br J Radiol* (2019) 92(1102):20190271. doi: 10.1259/bjr.20190271
110. Foy JJ, Robinson KR, Li H, Giger ML, Al-Hallaq H, Armato SG. Variation in Algorithm Implementation Across Radiomics Software. *J Med Imaging Bellingham Wash* (2018) 5(4):044505. doi: 10.1117/1.JMI.5.4.044505
111. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
112. Fornacon-Wood I, Mistry H, Ackermann CJ, Blackhall F, McPartlin A, Faivre-Finn C, et al. Reliability and Prognostic Value of Radiomic Features are Highly Dependent on Choice of Feature Extraction Platform. *Eur Radiol* (2020) 30(11):6241–50. doi: 10.1007/s00330-020-06957-9
113. Bibault J-E, Chaix B, Nectoux P, Pienkowski A, Guillemasé A, Brouard B. Healthcare Ex Machina: Are Conversational Agents Ready for Prime Time in Oncology? *Clin Transl Radiat Oncol* (2019) 16:55–9. doi: 10.1016/j.ctro.2019.04.002
114. Pelayo S, Senathirajah Y. Section Editors for the IMIA Yearbook Section on Human Factors and Organizational Issues. Human Factors and Sociotechnical Issues. *Yearb Med Inform* (2019) 28(1):78–80. doi: 10.1055/s-0039-1677927

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

MINI REVIEW
published: 13 May 2021
doi: 10.3389/fonc.2021.640985

Role of Cardiac Imaging in the Diagnosis of Immune Checkpoints Inhibitors Related Myocarditis

Stéphane Ederhy[1,2], Joe-Elie Salem[2,3,4,5], Laurent Dercle[6], Abrar Saqif Hasan[7], Marion Chauvet-Droit[1], Pascal Nhan[1], Samy Ammari[8,9], Bruno Pinna[2], Alban Redheuil[10], Samia Boussouar[10], Stephane Champiat[11], Laurie Soulat-Dufour[1] and Ariel Cohen[1,2]*

[1] Department of Cardiology, Sorbonne Université, AP-HP, Saint-Antoine Hospital, Paris, France, [2] UNICO-GRECO APHP.Sorbonne Cardio-Oncology Program, Sorbonne Université, Paris, France, [3] Sorbonne Université, INSERM CIC-1901, AP-HP.Sorbonne, Department of Pharmacology, Pitié-Salpêtrière Hospital, Paris, France, [4] Unité INSERM UMRS-ICAN 1166, Unité de recherche sur les maladies cardiovasculaires, du métabolisme et de la nutrition, Sorbonne Universités, Paris, France, [5] Division of Medicine and Pharmacology, Cardio-oncology program, Vanderbilt University Medical Center, Nashville, TN, United States, [6] Department of Radiology, New York Presbyterian, Columbia University Irving Medical Center, New York, NY, United States, [7] Department of Internal Medicine, Montefiore/Albert Einstein College of Medicine, Bronx, NY, United States, [8] Radiology Department, Gustave Roussy Cancer Campus, Villejuif, France, [9] BIOMAPS, UMR1281, INSERM.CEA.CNRS, Université Paris-Saclay, Paris, France, [10] LIB Biomedical Imaging Laboratory INSERM, CNRS, ICT Cardiothoracic Imaging Unit & Radiology Department, ICAN Institute of Cardiometabolism and Nutrition, Pitié-Salpêtrière Hospital (AP-HP), Sorbonne Université, Paris, France, [11] Drug Development Department (DITEP), Institut Gustave Roussy, Villejuif, France

OPEN ACCESS

Edited by:
Pilar López-Larrubia, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Carlo Gabriele Tocchetti, University of Naples Federico II, Italy
Anju Nohria, Brigham and Women's Hospital and Harvard Medical School, United States

***Correspondence:**
Stéphane Ederhy
stephane.ederhy@aphp.fr
orcid.org/0000-0002-0792-2521

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *12 December 2020*
Accepted: *24 March 2021*
Published: *13 May 2021*

Citation:
Ederhy S, Salem J-E, Dercle L, Hasan AS, Chauvet-Droit M, Nhan P, Ammari S, Pinna B, Redheuil A, Boussouar S, Champiat S, Soulat-Dufour L and Cohen A (2021) Role of Cardiac Imaging in the Diagnosis of Immune Checkpoints Inhibitors Related Myocarditis. Front. Oncol. 11:640985. doi: 10.3389/fonc.2021.640985

Immune checkpoint inhibitors (ICI) have constituted a paradigm shift in the management of patients with cancer. Their administration is associated with a new spectrum of immune-related toxicities that can affect any organ. In patients treated with ICI, cardiovascular toxicities, particularly myocarditis, occur with a low incidence (<1%) but with a high fatality rate (30–50%). ICI-related myocarditis has been attributed to an immune infiltration, comprising of T-cells that are positive for CD3+, CD4+, CD8+, and macrophages that are positive for CD68. The diagnosis remains challenging and is made based on clinical syndrome, an electrocardiogram (ECG), biomarker data, and imaging criteria. In most clinical scenarios, endomyocardial biopsy plays a pivotal role in diagnosis, while cardiac magnetic resonance imaging (cMRI) has limitations that should be acknowledged. In this review, we discuss the role of medical imaging in optimizing the management of ICI related myocarditis, including diagnosis, prognostication, and treatment decisions.

Keywords: myocarditis, immune checkpoint inhibitor, cancer, cardiac magnetic resonance imaging, cardiotoxicity

INTRODUCTION

Immune checkpoint inhibitors (ICIs) have modified the management of patients with cancer and have improved their prognosis and survival for many tumor types including melanoma, lymphoma, kidney, and lung malignancies (1). From a mechanistic point of view, ICIs are monoclonal antibodies that antagonize the pathways for programmed cell death receptor 1 (PD-1), programmed cell death ligand 1 (PD-L1), cytotoxic T-lymphocyte–associated protein 4 (CTLA-4) and could activate the immune system against cancer cells (2–4).

ICI administration can lead to immune-related adverse events (irAE) that can potentially affect all organs (5, 6). Among them, ICI related myocarditis is a rare adverse event that has an estimated incidence between 0.09 to 1.14% (7, 8). The incidence of myocarditis is higher in patients treated with ICI administered in combinations (e.g., anti-PD-1 and anti-CTLA-4) (0.27%) than in monotherapy with ICI (0.09%) (5). An early onset characterizes ICI related myocarditis, generally occurring within the first three injections (9).

ICI-related myocarditis is associated with poor outcomes, since the fatality rate has been reported to range between 30 to 50% (3, 9). Hence, early diagnosis and management is crucial. ICI-related myocarditis is frequently associated with other irAE, such as myositis (in 25.4%), myasthenia-like syndromes (ptosis, diplopia, respiratory failure; in 10.7%), and hepatitis (in 10.7%) (9). There is a growing body of evidence indicating that ICI-related myocarditis should be considered as a new entity and etiology of acute myocarditis, that differs from other etiologies by several aspects, including clinical presentation, ECG, and cardiac magnetic resonance imaging (cMRI) features (10). In this review, we discuss medical imaging's role in optimizing the management of ICI related myocarditis, including diagnosis, prognostication, treatment decision, and follow-up.

DIAGNOSTIC CRITERIA FOR ICI-RELATED MYOCARDITIS

According to the ESC 2013 position statement, the diagnosis of acute myocarditis relies on a combination of a suggestive clinical presentation, first-line tests such as ECG, biomarkers including inflammatory markers, viral antibodies, serum cardiac antibodies, and cardiac biomarkers (Troponin I or T, BNP, or NT pro BNP), transthoracic echocardiography (TTE), and cMRI. It also includes second-line tests, represented by coronary angiography (CA), to exclude coronary artery disease and endomyocardial biopsy (EMB) (11). This diagnostic workup could be applied to document ICI-related myocarditis, but it should also be adapted to the specific context of cancer patients treated with ICI, as specified in a recent statement (10).

CLINICAL PRESENTATION

The clinical presentation in a patient admitted for acute myocarditis (AM) includes a wide range of symptoms. On one hand, some patients are pauci symptomatic on presentation with chest pain and/or palpitations. On the other hand, there are more severe scenarios with cardiogenic shock (11). Discussing out of the context of cancer patients treated with ICI, five possible clinical scenarios could be identified in acute myocarditis patients. They include acute coronary like syndrome, new onset or worsening heart failure, chronic heart failure, and life-threatening conditions including arrhythmia, sudden cardiac arrest, and cardiogenic shock with impaired left ventricular systolic dysfunction (11).

In cancer patients treated with ICI, the clinical presentation of ICI-related myocarditis has recently been described by Pradhan et al. in a review gathering data from 88 published cases. They found that the most commonly reported symptoms in ICI-related myocarditis were dyspnea (49%), weakness (25%), chest pain (17%), syncope (9%), fever (6%), and cough (4%) (12). Multiple recent publications show that ICI-related myocarditis could have several other presentations. There is a growing body of evidence suggesting that patients presenting with an irAE could have a final diagnosis of AM without any evidence of clinical cardiac manifestations and that some patients could have isolated troponin rise and/or ECG modifications (13).

ECG FINDINGS

ECG is generally abnormal, but it should be noticed that ECG features are neither specific nor sensitive enough for the diagnosis of AM. Main ECG findings that could be encountered in patients with acute myocarditis are the following: 1st, 2nd, and 3rd degree atrioventricular block, bundle branch block, ST/T wave modifications, ST elevation (generally concave and non-mirror), non-ST elevation, T wave inversion, sinus arrest, ventricular tachycardia or fibrillation, asystole, atrial fibrillation, intraventricular conduction delay, abnormal Q waves, premature beats, and SVT (14). However, a normal ECG does not rule out the diagnosis of acute myocarditis. In a study of 77 patients with acute myocarditis, ECGs were normal in 32% of the patients, ST elevation was found in 57%, inverted T wave in 9%, and left bundle branch block in 3% (15). In a similar fashion, for a study of 65 patients with biopsy-proven myocarditis, ST-abnormalities were detected in 69% of the patients, bundle-branch-block in 26%, Q-waves in 8%, atrial fibrillation was present in 6%, and AV-Block in 3% of the patients (16).

In cancer patients with ICI-related myocarditis, Pradhan showed that 91% of ECGs were found to be abnormal and that there was a broad spectrum of abnormal findings. ST-elevation was reported in 32% and ST-depression in 4% cases. Various degrees of heart block were found in 51%, with complete AV block involving 66% of them. Ventricular tachycardia or fibrillation were noted in 35% (12).

TRANSTHORACIC ECHOCARDIOGRAPHY

Transthoracic echocardiography (TTE) represents the first line of imaging when AM is clinically suspected. It has the advantage of being non-invasive, non-ionizing, versatile, and available at bedside. TTE provides information on cardiac geometry, morphology, and function. TTE findings suggestive of acute myocarditis encompass segmental wall motion abnormalities, increased LV wall thickness, global hypokinesia, particularly in fulminant myocarditis, and pericardial effusion. The current role of speckle tracking imaging is not clearly established in this context. A normal TTE does not rule out the diagnosis of acute myocarditis (17).

Out of the context of immunotherapy, in a case series of 41 patients with histologically proven myocarditis, left ventricular dysfunction was noticed in 69%, right ventricular dysfunction in 23%, wall motion abnormality in 64%, left ventricular "pseudo hypertrophy" in 20%, and ventricular thrombi in 15% (15). Felker described the TTE features in the vast majority of patients admitted for fulminant myocarditis, as showing an increased septal thickness and normal LV dimension. In contrast, those with non-fulminant myocarditis had an increased diastolic dimension with a normal septal thickness (18).

In cancer patients with ICI-related myocarditis, there are scarce data on TTE findings. In a recent review of the literature describing TTE findings in 53 myocarditis cases, 23% of TTE were classified as normal, and 32.5% of TTE examinations reported a normal LVEF (12).

CARDIAC MAGNETIC RESONANCE AND ICI-RELATED MYOCARDITIS

cMRI as TTE allows to define cardiac geometry, morphology, function and add important information by allowing myocardial tissue characterization, particularly in the context of inflammation related to myocarditis. The combination of markers of edema and inflammation increases the probability of AM. According to the updated 2018 Lake Louise criteria, at least one T2 based criterion (a regional or global increase of myocardial T2 relaxation time or increased signal intensity in T2-weighted CMR images) with at least one T1 based criterion (increase myocardial T1, extracellular volume, or late gadolinium enhancement) should be analyzed and combined to improve the diagnostic accuracy of cMRI for the diagnosis of AM (19, 20).

In cancer patients with ICI-related myocarditis, the most extensive description of cMRI findings was made by Zhang (21). LGE was found in 48% of all the cases, was predominantly distributed at the anteroseptal, inferoseptal, inferior, and inferolateral segments. Myocardial edema, as assessed with T2 weighted STIR was found in only 28% of cases. Forty-three patients had neither elevated T2 nor LGE. The predominant LGE pattern was subendocardial/transmural in 6.1%, subepicardial in 26.5%, mid-myocardial in 49%, and diffuse in 18.4%. Native T1 value was comparable (18). However, it should be noticed that a normal cMRI with a normal T2, T1 values and no LGE does not rule out ICI myocarditis proven with BEM (21, 22).

ENDOMYOCARDIAL BIOPSY AND ACUTE MYOCARDITIS

Acute myocarditis, according to the 2013 ESC position statement, is defined as an inflammatory disease of the myocardium diagnosed by histological evidence of inflammatory infiltrates defined as ≥14 leucocytes/mm², including up to 4 monocytes/mm² with the presence of CD 3 positive T-lymphocyte ≥7 cells/mm², within the myocardium associated with myocyte degeneration and necrosis of non-ischemic origin (11).

Endomyocardial biopsy is necessary to achieve a diagnosis of certainty and identify its cause. Causes of myocarditis include infectious myocarditis (bacterial, spirochetal, viral), immune-mediated myocarditis (allergens, alloantigens, autoantigens), and toxic myocarditis (drugs, heavy metals, hormones, physical agents, miscellaneous) (11). It is also essential to bear in mind that cancer patients developing myocarditis could have other possible etiologies, including radiotherapy, anthracycline, or other viral infections. (23).

In cancer patients with ICI-related myocarditis, postmortem histopathological analysis of heart and skeletal biopsies in two patients treated with combination therapy (ipilimumab and nivolumab) revealed the myocardium as necrotized and associated with an intense, patchy, lymphocytic infiltrate. The infiltrate comprised T-cells positive for CD3+, CD4+, and CD8+, and macrophages positive for CD68. PD-L1 was expressed on myocytes' membranous surface, on infiltrating CD8+ T-cells, but was not expressed on skeletal muscle or tumor (7).

Of note, in expert centers, the complication rate of EMB is low and is influenced by operator experience, the volume of procedures and a learning curve. Incidence of peri-procedural complications (perforation, tamponade, embolization) after left ventricular EMB and right ventricular was low and comparable (0.33% for left ventricular EMB and 0.45% for right ventricular EMB) with no death occurring following EMB (24).

IMAGING AND PROGNOSIS IN ICI-RELATED MYOCARDITIS

There is scarce data in the context of ICI-related myocarditis on prognostic markers. Transthoracic echocardiography, mainly left ventricular ejection fraction (LVEF), global longitudinal strain (GLS), and abnormal findings on CMR were evaluated as potential prognostic markers in patients with ICI-related myocarditis (25).

In cancer patients with ICI-related myocarditis, Pradhan et al. in a review of 88 cases, published that LVEF was not a predictor of poor outcomes (12). Awadalla et al. showed that MACE's risk was higher with a lower GLS among patients with either a reduced or preserved left ventricular ejection fraction. After adjustment for ejection fraction, each percent reduction in GLS was associated with a 1.5-fold increase in MACE among patients with a reduced left ventricular ejection fraction (HR 1.5, 95% CI 1.2–1.8) and a 4.4-fold increase with a preserved left ventricular ejection fraction (HR 4.4, 95% CI 2.4–7.8) (25). Regarding CMR parameters, the presence of LGE, LGE pattern, and elevated T2-weighted short TI inversion recovery were not associated with MACE (21).

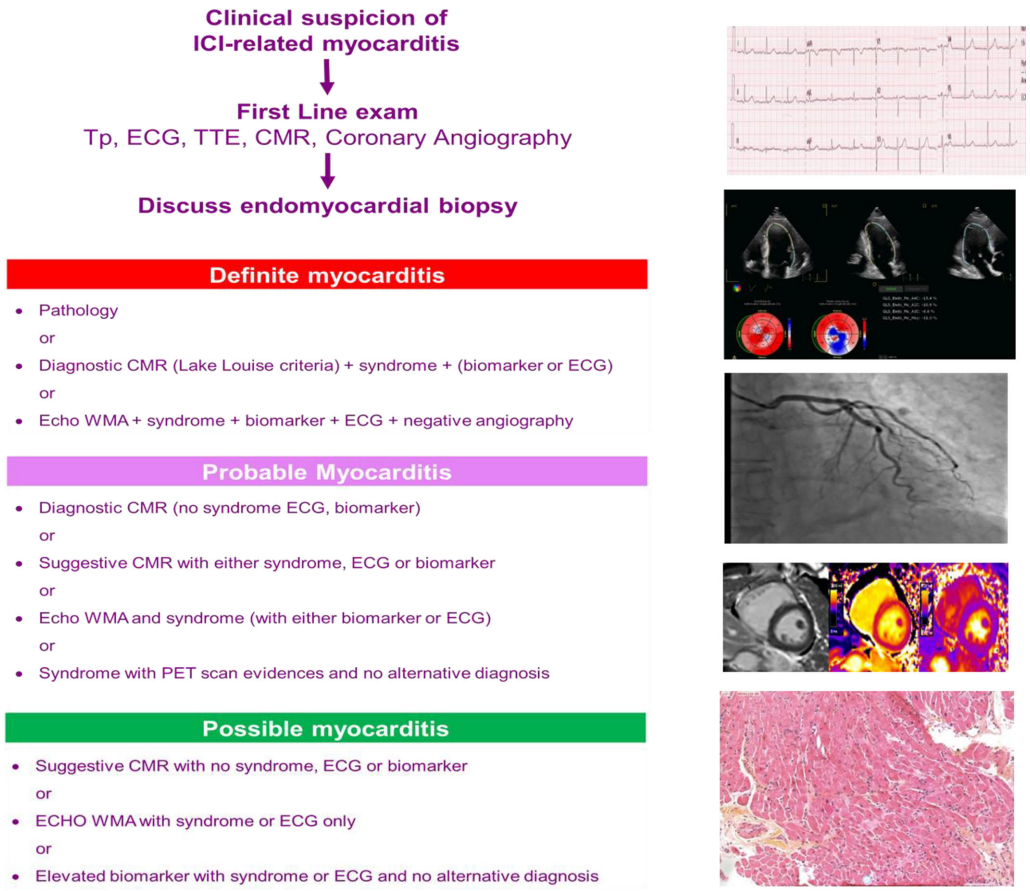


FIGURE 1 | CMR, cardiac magnetic resonance; ECG, electrocardiogram; EMB, endomyocardial biopsy; ICI, immune checkpoint inhibitor; Tn, troponin; TTE, transthoracic echocardiography.

TOWARDS A PRAGMATIC APPROACH ON HOW TO DIAGNOSE ICI-RELATED MYOCARDITIS

The diagnosis of ICI-related myocarditis relies on combining a clinical syndrome, ECG, troponin measurement, and imaging criteria. Most recent ESMO and ASCO guidelines dealing with the management of ICI-related myocarditis do not describe any diagnostic workup (26, 27). However, recent position statements from the European heart failure association suggest, despite clear evidences, to obtain references values before ICI initiation based on an echocardiogram, an electrocardiogram, and biomarkers measurement (troponin and a natriuretic peptide) (28, 29).

In symptomatic patients, Bonaca intended to describe particularly in the context of immunotherapy, the diagnostic criteria that should be used both in everyday clinical practice and in clinical trials (10).

Figure 1 depicts a diagnostic workup adapted from Bonaca et al. and proposes that when the diagnosis of ICI-related myocarditis is suspected, patients should be evaluated with at least one ECG and a troponin measurement and should be rapidly referred to a cardio-oncology unit that can confirm or exclude the diagnosis of ICI-associated myocarditis (10).

CONCLUSIONS

The incidence of ICI-associated myocarditis is low (below 1%) but could be underestimated since it is not systematically screened. It is critical to diagnose this irAE at an early stage since it is associated with a fatality rate between 30 to 50%. Current strategies usually rely on a suspicion by the patient's oncologist and a confirmation by a cardiologist or cardio-oncologist based on CMR or endomyocardial biopsies. The diagnosis of ICI-related myocarditis remains challenging and the main objective is to make an early diagnosis since no predictive markers are currently available to identify patients prone to develop ICI related myocarditis.

AUTHOR CONTRIBUTIONS

Writing: SE, J-ES, LD, AC. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Champiat S, Lambotte O, Barreau E, Belkhir R, Berdelou A, Carbonnel F, et al. Management of Immune Checkpoint Blockade Dysimmune Toxicities: A Collaborative Position Paper. *Ann Oncol* (2016) 27:559–74. doi: 10.1093/annonc/mdv623
2. Geraud A, Gougis P, Vozy A, Anquetil C, Allenbach Y, Romano E, et al. Clinical Pharmacology and Interplay of Immune Checkpoint Agents: A Yin-Yang Balance. *Annu Rev Pharmacol Toxicol* (2020). doi: 10.1146/annurev-pharmtox-022820-093805. doi: 10.1146/annurev-pharmtox-022820-093805
3. Varricchi G, Galdiero MR, Tocchetti CG. Cardiac Toxicity of Immune Checkpoint Inhibitors: Cardio-Oncology Meets Immunology. *Circulation* (2017) 136(21):1989–92. doi: 10.1161/CIRCULATIONAHA.117.029626
4. Hu JR, Florido R, Lipson EJ, Naidoo J, Ardehali R, Tocchetti CG, et al. Cardiovascular Toxicities Associated With Immune Checkpoint Inhibitors. *Cardiovasc Res* (2019) 115:854–68. doi: 10.1093/cvr/cvz026
5. Wang DY, Salem JE, Cohen JV, Chandra S, Menzer C, Ye F, et al. Fatal Toxic Effects Associated With Immune Checkpoint Inhibitors: A Systematic Review and Meta-Analysis. *JAMA Oncol* (2018) 4:1721–8. doi: 10.1001/jamaoncol.2018.3923
6. Michot JM, Lappara A, Le Pavec J, Simonaggio A, Collins M, De Martin E, et al. The 2016-2019 ImmunoTOX Assessment Board Report of Collaborative Management of Immune-Related Adverse Events, an Observational Clinical Study. *Eur J Cancer* (2020) 130:39–50. doi: 10.1016/j.ejca.2020.02.010
7. Johnson DB, Balko JM, Compton ML, Chalkias S, Gorham J, Xu Y, et al. Fulminant Myocarditis With Combination Immune Checkpoint Blockade. *N Engl J Med* (2016) 375:1749–55. doi: 10.1056/NEJMoa1609214
8. Mahmood SS, Fradley MG, Cohen JV, Nohria A, Reynolds KL, Heinzerling LM, et al. Myocarditis in Patients Treated With Immune Checkpoint Inhibitors. *J Am Coll Cardiol* (2018) 71:1755–64. doi: 10.1016/S0735-1097(18)31240-3
9. Salem JE, Manouchehri A, Moey M, Lebrun-Vignes B, Bastarache L, Pariente A, et al. Cardiovascular Toxicities Associated With Immune Checkpoint Inhibitors: An Observational, Retrospective, Pharmacovigilance Study. *Lancet Oncol* (2018) 19:1579–89. doi: 10.1016/S1470-2045(18)30608-9
10. Bonaca MP, Olenchock BA, Salem JE, Wiviott SD, Ederhy S, Cohen A, et al. Myocarditis in the Setting of Cancer Therapeutics: Proposed Case Definitions for Emerging Clinical Syndromes in Cardio-Oncology. *Circulation* (2019) 140:80–91. doi: 10.1161/CIRCULATIONAHA.118.034497
11. Caforio ALP, Adler Y, Agostini C, Allanore Y, Anastasakis A, Arad M, et al. Diagnosis and Management of Myocardial Involvement in Systemic Immune-Mediated Diseases: A Position Statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Disease. *Eur Heart J* (2017) 38(35):2649–62. doi: 10.1093/eurheartj/ehx321
12. Pradhan R, Nautiyal A, Singh S. Diagnosis of Immune Checkpoint Inhibitor-Associated Myocarditis: A Systematic Review. *Int J Cardiol* (2019) 296:113–21. doi: 10.1016/j.ijcard.2019.07.025
13. Lyon AR, Yousaf N, Battisti NML, Moslehi J, Larkin J. Immune Checkpoint Inhibitors and Cardiovascular Toxicity. *Lancet Oncol* (2018) 19(9):e447–58. doi: 10.1016/S1470-2045(18)30457-1
14. Di Bella G, Florian A, Oreto L, Napolitano C, Todaro MC, Donato R, et al. Electrocardiographic Findings and Myocardial Damage in Acute Myocarditis Detected by Cardiac Magnetic Resonance. *Clin Res Cardiol* (2012) 101(8):617–24. doi: 10.1007/s00392-012-0433-5
15. Lee Chuy K, Oikonomou EK, Postow MA, Callahan MK, Chapman PB, Shoushtari AN, et al. Myocarditis Surveillance in Patients With Advanced Melanoma on Combination Immune Checkpoint Inhibitor Therapy: The Memorial Sloan Kettering Cancer Center Experience. *Oncologist* (2019) 24:e196–7. doi: 10.1634/theoncologist.2019-0040
16. Deluigi CC, Ong P, Hill S, Wagner A, Kispert E, Klingel K, et al. ECG Findings in Comparison to Cardiovascular MR Imaging in Viral Myocarditis. *Int J Cardiol* (2013) 165(1):100–6. doi: 10.1016/j.ijcard.2011.07.090
17. Pinamonti B, Alberti E, Cigalotto A, Dreas L, Salvi A, Silvestri F, et al. Echocardiographic Findings in Myocarditis. *Am J Cardiol* (1988) 62(4):285–91. doi: 10.1016/0002-9149(88)90226-3
18. Felker GM, Boehmer JP, Hruban RH, Hutchins GM, Kasper EK, Baughman KL, et al. Echocardiographic Findings in Fulminant and Acute Myocarditis. *J Am Coll Cardiol* (2000) 36(1):227–32. doi: 10.1016/S0735-1097(00)00690-2
19. Friedrich MG, Sechtem U, Schulz-Menger J, Holmvang G, Alakija P, Cooper LT, et al. International Consensus Group on Cardiovascular Magnetic Resonance in Myocarditis. Cardiovascular Magnetic Resonance in Myocarditis: A Jacc White Paper. *J Am Coll Cardiol* (2009) 53(17):1475–87. doi: 10.1016/j.jacc.2009.02.007
20. Ferreira VM, Schulz-Menger J, Holmvang G, Kramer CM, Carbone I, Sechtem U, et al. Cardiovascular Magnetic Resonance in Nonischemic Myocardial Inflammation: Expert Recommendations. *J Am Coll Cardiol* (2018) 72(24):3158–76. doi: 10.1016/j.jacc.2018.09.072
21. Zhang L, Awadalla M, Mahmood SS, Nohria A, Hassan MZO, Thuny F, et al. Cardiovascular Magnetic Resonance in Immune Checkpoint Inhibitor-Associated Myocarditis. *Eur Heart J* (2020) 41:1733–43. doi: 10.1093/eurheartj/ehaa051
22. Ederhy S, Fenioux C, Cholet C, Rouvier P, Redheuil A, Cohen A, et al. Immune Checkpoint Inhibitor Myocarditis With Normal Cardiac Magnetic Resonance Imaging: Importance of Cardiac Biopsy and Early Diagnosis. *Can J Cardiol* (2020) S0828-282X(20):31189–2. doi: 10.1016/j.cjca.2020.12.022
23. Mirabel M. Late-Onset Giant Cell Myocarditis Due to Enterovirus During Treatment With Immune Checkpoint Inhibitors, Jacc. *CardioOncology* (2020) 2(3):511–4. doi: 10.1016/j.jaccao.2020.05.022
24. Chimenti C, Frustaci A. Contribution and Risks of Left Ventricular Endomyocardial Biopsy in Patients With Cardiomyopathies: A Retrospective Study Over a 28-Year Period. *Circulation* (2013) 128(14):1531–41. doi: 10.1161/CIRCULATIONAHA.13.001414
25. Awadalla M, Mahmood SS, Groarke JD, Hassan MZO, Nohria A, Rokicki A, et al. Global Longitudinal Strain and Cardiac Events in Patients With Immune Checkpoint Inhibitor-Related Myocarditis. *J Am Coll Cardiol* (2020) 75:467–78. doi: 10.1016/j.jacc.2019.11.049
26. Curigliano G, Lenihan D, Fradley M, Ganatra S, Barac A, Blaes A, et al. Management of Cardiac Disease in Cancer Patients Throughout Oncological Treatment: ESMO Consensus Recommendations. *Ann Oncol* (2020) 31:171–90. doi: 10.1016/j.annonc.2019.10.023
27. Brahmer JR, Lacchetti C, Schneider BJ, Atkins MB, Brassil KJ, Caterino JM, et al. Management of Immune-Related Adverse Events in Patients Treated With Immune Checkpoint Inhibitor Therapy: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol* (2018) 36:1714–68. doi: 10.1200/JCO.2017.77.6385
28. Lyon AR, Dent S, Stanway S, Earl H, Brezden-Masley C, Cohen-Solal A, et al. Baseline Cardiovascular Risk Assessment in Cancer Patients Scheduled to Receive Cardiotoxic Cancer Therapies: A Position Statement and New Risk Assessment Tools From the Cardio-Oncology Study Group of the Heart Failure Association of the European Society of Cardiology in Collaboration With the Internationalcardio-Oncologysociety. *Eur J Heart Fail* (2020) 22(11):1945–60. doi: 10.1002/ejhf.1920
29. Čelutkienė J, Pudil R, López-Fernández T, Grapsa J, Nihoyannopoulos P, Bergler-Klein J, et al. Role of Cardiovascular Imaging in Cancer Patients Receiving Cardiotoxic Therapies: A Position Statement on Behalf of the Heart Failure Association (HFA), the European Association of Cardiovascular Imaging (EACVI) and the Cardio-Oncology Council of the European Society of Cardiology(ESC). *Eur J Heart Fail* (2020) 22(9):1504–24. doi: 10.1002/ejhf.1957

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 27 May 2021
doi: 10.3389/fonc.2021.638185

Uncontrolled Confounders May Lead to False or Overvalued Radiomics Signature: A Proof of Concept Using Survival Analysis in a Multicenter Cohort of Kidney Cancer

Lin Lu[1†], Firas S. Ahmed[1†], Oguz Akin[2], Lyndon Luk[1], Xiaotao Guo[1], Hao Yang[1], Jin Yoon[1], A. Aari Hakimi[3], Lawrence H. Schwartz[1] and Binsheng Zhao[1]*

[1] Department of Radiology, Columbia University Irving Medical Center, New York, NY, United States, [2] Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, United States, [3] Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, United States

OPEN ACCESS

Edited by:
Fatima-Zohra Mokrane, Université Toulouse III Paul Sabatier, France

Reviewed by:
Isaac Shiri, Geneva University Hospitals (HUG), Switzerland
Cyril Jaudet, Centre François Baclesse, France

***Correspondence:**
Firas S. Ahmed
fsa2101@cumc.columbia.edu

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *05 December 2020*
Accepted: *06 April 2021*
Published: *27 May 2021*

Citation:
Lu L, Ahmed FS, Akin O, Luk L, Guo X, Yang H, Yoon J, Hakimi AA, Schwartz LH and Zhao B (2021) Uncontrolled Confounders May Lead to False or Overvalued Radiomics Signature: A Proof of Concept Using Survival Analysis in a Multicenter Cohort of Kidney Cancer. Front. Oncol. 11:638185. doi: 10.3389/fonc.2021.638185

Purpose: We aimed to explore potential confounders of prognostic radiomics signature predicting survival outcomes in clear cell renal cell carcinoma (ccRCC) patients and demonstrate how to control for them.

Materials and Methods: Preoperative contrast enhanced abdominal CT scan of ccRCC patients along with pathological grade/stage, gene mutation status, and survival outcomes were retrieved from The Cancer Imaging Archive (TCIA)/The Cancer Genome Atlas—Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) database, a publicly available dataset. A semi-automatic segmentation method was applied to segment ccRCC tumors, and 1,160 radiomics features were extracted from each segmented tumor on the CT images. Non-parametric principal component decomposition (PCD) and unsupervised hierarchical clustering were applied to build the radiomics signature models. The factors confounding the radiomics signature were investigated and controlled sequentially. Kaplan–Meier curves and Cox regression analyses were performed to test the association between radiomics signatures and survival outcomes.

Results: 183 patients of TCGA-KIRC cohort with available imaging, pathological, and clinical outcomes were included in this study. All 1,160 radiomics features were included in the first radiomics signature. Three additional radiomics signatures were then modelled in successive steps removing redundant radiomics features first, removing radiomics features biased by CT slice thickness second, and removing radiomics features dependent on tumor size third. The final radiomics signature model was the most parsimonious, unbiased by CT slice thickness, and independent of tumor size. This final radiomics signature stratified the cohort into radiomics phenotypes that are different by cancer-specific and recurrence-free survival; HR (95% CI) = 3.0 (1.5–5.7), $p < 0.05$ and HR (95% CI) = 6.6 (3.1–14.1), $p < 0.05$, respectively.

Conclusion: Radiomics signature can be confounded by multiple factors, including feature redundancy, image acquisition parameters like slice thickness, and tumor size. Attention to and proper control for these potential confounders are necessary for a reliable and clinically valuable radiomics signature.

Keywords: radiomics, quality control, machine learning, TCGA, The Cancer Imaging Archive (TCIA), clear cell renal cell cancer

INTRODUCTION

Tumor radiomics is a rapidly evolving field aiming to link tumor imaging phenotypes to pathological and clinical outcomes in a quantitative and non-invasive way (1). Radiomics generally converts medical image data into a large-scale and mineable set of imaging features, termed radiomics features, that characterize tumor imaging phenotypes (2). Radiomics signatures, essentially constellations of radiomics features, have shown to be helpful in plenty of medical tasks (3), including predicting malignancy in lung nodules at lung cancer screening CT scans (4), predicting genomic alteration on lung cancer imaging (5), predicting tumor recurrence and patients' survival (6), and assessing response to treatment (7, 8).

Radiomics signature models have been developed by cancer researchers but their usefulness is usually difficult to replicate at other institutions or cohorts. This is mostly due to challenges encountered in the construction of a radiomics signature models attributed to radiomics feature redundancy and image quality differences (resulting from differences in image acquisition/technical parameters or from scanner vender differences). Another challenge facing useful radiomics signature is the need to provide new information independent of already known and established prognosticators, especially tumor size which is retrieved from routine clinical imaging without the need to run radiomics image analysis (9). Feature redundancy is a challenge to replicate and consolidate radiomics signatures. Two research teams, Lu et al. (10) and Berenguer et al. (11), independently pointed out that radiomics feature sets, which usually contain several hundreds to a thousand radiomics features, could actually be summarized into dozens of representative features. The variations in image acquisition parameters, *e.g.* thin/thick slice thickness and sharp/smooth reconstruction kernels, etc., could produce images of different qualities (12), which might impede generalization of radiomics signatures. For instance, the performance of radiomics signature developed using CT images of thin slice thickness decreased when applied on CT images with thicker in the predicting the risk of malignancy of lung nodule (13) and cancer-related genomic mutation status (14). Finally, including tumor size measurement (unidimensional, bidimensional and three dimensional) within radiomics features creates confusion about the usefulness of the texture based radiomics; it raises the question whether the prognostic or predictive radiomics signature effect is mainly driven by tumor size which is readily available through routine medical imaging without the need for radiomic analysis. Association between radiomics signature and well-established clinical factors (*e.g.*, tumor size or patient's age), may lead to overvalued radiomics signatures; this is because the predictive value of radiomics signature may be exaggerated by radiomics' association with these important clinical factors (15).

Several approaches were proposed for establishing reproduceable and generalizable radiomics studies including radiomics reporting guidelines, such as Radiomics Quality Score (RQS) (9), The Image Biomarker Standardization Initiative (IBSI) (16, 17), and recently harmonization algorithms (18), such as Combat. Although these studies have demonstrated that radiomics signature could be impacted by multiple clinical and technical factors, there is still suboptimal awareness of this confounding potential and lack of consensus on how to control for such confounding. For example, within the RQS, although imaging protocol was suggested to be reported, it does not provide a reliable statistical method to control the confounding effect from imaging protocol and does not alarm that confounding effect of imaging protocol could lead to fake result. In IBSI, its main focus is on standardizing implementation parameters for radiomics feature extraction instead of controlling confounding effect. For those harmonization algorithms, like Combat, although they showed promising potential on removing confounding effect, however, there is limitation on application on new data. For example, when new data were added, the new data have to be combined with original data and the harmonization has to be re-established on the entire combined database (19).

Therefore, in this study, we designed multiple radiomics signature models to show the effect of uncontrolled confounders which may lead to false/overvalued radiomics signature among patients with clear cell renal cell carcinoma (ccRCC). The reason for using radiomics analysis on ccRCC as an example is that, ccRCC is the predominant pathological subtype (85%) in renal adenocarcinomas which account for 90% of kidney cancers because of its variable course (20, 21). The prediction of survival outcomes for ccRCC patients still remain challenging (22–25), due to the variation in ccRCC's growth pattern, with some tumor showing an indolent growth pattern while others exhibiting aggressive behaviors including local recurrence after resection and distant metastases (26, 27).

METHOD

We aimed to conduct this study in The Cancer Genome Atlas—Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) cohort data (28) which is a publicly available dataset from multiple medical

institutions in the US. The TCGA-KIRC project house the pathological, clinical, and imaging data for patients with clear cell renal cell carcinoma (ccRCC).

Compared to single-center data, the TCGA-KIRC was a multicenter data therefore was more heterogenous in terms of tumor's pathological stage and grade as well as the image acquisition parameters. The Cancer Imaging Archive (TCIA) (29) represent a repository of clinical imaging for patients/tumors included in the TCGA cohort housing de-identified clinical imaging and provide a great resource for researchers to conduct and validate their imaging related studies.

The overview of our study design is presented in **Figure 1**. Our study design followed the basic radiomics phases, which included data collection, feature extraction, modeling, and outcome analysis (9, 17). The highlights of the study are the following. First, all the used data are publicly available in the TCIA, so that other researchers can easily and reliably replicate our results. Second, multiple factors that might affect radiomics analysis (9) were investigated, including feature redundancy (*e.g.* correlation among features), image acquisition parameters (slice thickness was the main CT parameter impacting radiomic signature), and signature's dependency to tumor size (a previously validated prognostic factor). Four radiomics signatures were successively built throughout our study that included: 1) entire radiomics feature set, 2) radiomics feature set after dimension reduction (i.e. excluding redundant features), 3) radiomics feature set after further exclusion of radiomics features affected by CT scan slice thickness, and 4) radiomics feature set after further exclusion of tumor size related features. Third, to address the over-fitting problem, non-parametric principal component decomposition (PCD) for dimension reduction and unsupervised hierarchical clustering for pattern discovery were used. Fourth, the radiomics signatures associations with clinical outcomes (OS, PFS, and RFS) were tested using Kaplan–Meier's analysis. Finally, supplementary analyses were conducted to illustrate how impacting factors can affect the radiomics signatures.

Patient Data Collection

The data we used were downloaded from the TCGA-KIRC project, which is publicly available in the TCIA dataset (https://wiki.cancerimagingarchive.net/display/Public/TCGA-KIRC, accessed in August 2016). It contained 267 ccRCC patients collected from multiple medical centers nationwide. The downloaded content contained both presurgical contrast-enhanced abdominal CT scans and clinical information such as pathological stage, grade, gene mutation status, and patient's survival outcomes. For the survival outcome, patients were censored at their last follow-up date if: they were alive (overall survival-OS), alive or dead from non-ccRCC related illness

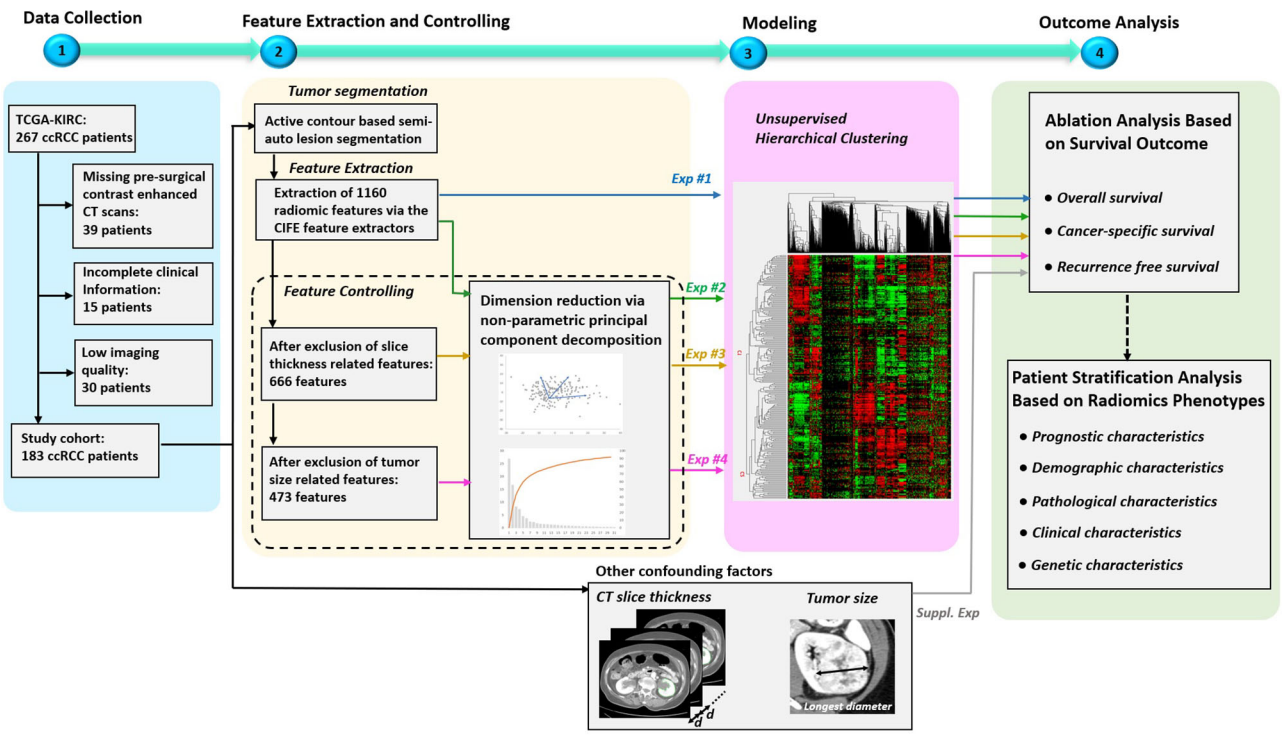


FIGURE 1 | Overview of the study design. The study mainly consisted of four key parts: 1) patient data collection, 2) feature extraction and controlling, 3) Modeling, and 4) Outcome analysis. Specially, four experiments were designed to evaluate the effects on radiomics signatures built by radiomics feature sets under four different controlling levels. In addition, supplementary experiments were performed to explore the association between outcomes and the confounding factors, such as CT slice thickness and tumor size.

(cancer-specific survival-CSS), alive without tumor recurrence (recurrence-free survival-RFS). Somatic gene mutation status of our patients was retrieved from TCGA official website and adjudicated at Memorial Sloan Kettering Cancer Center (MSKCC) based on collaborative clinical TCGA (cTCGA) consortium data (30). Genetic information included gene mutation status of VHL, PBRM1, SETD2, BAP1, and KDM5C genes. All patients' informed consents and institutional review boards' approvals were obtained as part of the TCGA/TCIA efforts, and all demographic and imaging data were de-identified to comply with the Health Insurance Portability and Accountability Act (HIPAA).

In our study, 183 out of 267 patients were included in the final segmentation and analysis (as shown in **Figure 1**) based solely on the availability of pre-surgical contrast-enhanced abdominal CT scans where ccRCC was depicted and segmented to generate radiomics features.

Tumor Segmentation and Feature Extraction

A fellowship trained abdominal radiologist performed ccRCC segmentations. The radiologist was blinded to the various study endpoints (pathology at surgery and patient outcomes). For image analysis we used a MatLab (MathWorks, Natick, Massachusetts) based dedicated software application to visualize and segment the tumor from each patient's CT scan. This semi-automated algorithm, combining the region-based active contours and a level set approach, was used in a slice-by-slice fashion (i.e. the entire ccRCC tumor was segmented) (31). The initial step for tumor segmentation required the radiologist to manually select a region-of-interest (ROI) that roughly enclosed the tumor region on a single CT slice. Rough boundary localization of the tumor was then automatically generated by the software algorithm and propagated to consecutive slices, serving as an initial ROI for subsequent segmentations on the neighboring images. The final tumor segmentation boundaries were then verified and fine-tuned by the radiologist (32). The total volume of the tumor (created by adding all segmentations from all slices) was then utilized to generate the radiomics features of each individual tumor.

A total of 1,160 radiomics features, i.e., quantitative imaging features, were extracted from each segmented tumor *via* the Columbia Image Feature Extractor (CIFE) (33) which has been successfully applied in many radiomics studies (34–36). More details of the CIFE, as well as its comparison with two other open-source feature extractors, the IBEX (37) and Pyradiomics (38), can be founder at (33). Three preprocesses were performed before the feature extraction, 1) a modified soft tissue CT window was adopted with level of 50 HU and width of 175 HU, 2) voxel resolution was resampled to $0.5 \times 0.5 \times 0.5$ mm^3 and 3) image was discretized into 64 bins.

Principal Component Decomposition (PCD)

In this study, we introduced an unsupervised method, PCD (39), for feature dimension reduction. On contrary to its supervised counterpart, an unsupervised method focuses more on the intrinsic characteristics of features and is not easily affected by the overfitting problem. PCD belongs to a type of non-parametric transformation that is able to convert a set of possibly correlated features into a set of linearly uncorrelated variables. Such uncorrelated variables are called principal components and are ranked by their corresponding variance, which is their contribution to feature variability in the data. Hence, in the set of resulting principal components, the first principal component has the largest variance, and each succeeding component in turn has a smaller variance. We hypothesized that if a principal component had a larger variance, it would contain more information, so that the dimension reduction could be fulfilled by selecting a compact set of principal components that had the large variance while excluding a large number of principal components with small variances (i.e. excluding those with the least input to the data). In this study, Matlab version 9.5 was used. Principal components that summed up to 99% contribution to the total variance were selected as the new representative features.

Exclusion of Slice Thickness Related Features

The main CT scan parameter in this cohort that affected (was associated with) radiomics features values was the CT scan slice thickness (among other parameters including CT scan voltage (kVp), vender and reconstruction algorithm). Because this is a multi-institutional cohort with different imaging protocols, we aimed to remove the potentially confounding effect of slice thickness from the radiomics signature to be built. The identification of radiomics features dependent on slice thickness involved three steps: First, the patients were distributed into two groups: one with thin CT slices (i.e. ≤3 mm) and one with thick slices (i.e. >3 mm) (The selection of 3 mm as an cutoff is based on clinical practice (40)). Second, C-index (41) was calculated for each feature based on the slice thickness group labels. The C-index, in this model, provided a measure of how good a radiomics feature could fit a binary outcome (groups of slice thickness ≤3 mm and >3 mm). In other words, we attempted to measure how much of the radiomics feature was explained by the CT slice thickness. Generally, for C-index, values below 0.5 indicate poor fitting, values over 0.7 indicate good fitting, values over 0.8 indicate strong fitting, and a value of 1 means perfect fitting. In radiomics signature model #3, we excluded all radiomics features whose C-index was >0.8 in order to remove the radiomics features that are heavily influenced/biased by CT slice thickness.

Exclusion of Tumor Size Related Features

In this project, we aimed to build a radiomics signature that deliver new prognostic information, independent of tumor size which has long been known as an important prognosticator. The correlation between tumor size and the radiomics features were measured by Pearson's linear correlation coefficients (also called Pearson's R). In our study, tumor size was obtained by measuring the longest diameter across the tumor's cross-sectional region, as shown in **Figure 1**. The features that have strong positive or negative correlation with tumor size (Pearson's R >0.7 or <−0.7,

$p < 0.05$) were excluded from radiomics signature model #4 as tumor size dependent features.

Unsupervised Clustering

Unsupervised hierarchical clustering was used to identify the clusters of ccRCC tumors based on input radiomics features. We identified two major clusters of ccRCC tumors in this study. During the clustering, the distance between two clusters in the feature space was measured by 'cosine' distance. The unsupervised clustering method was intrinsically an iteration process based on the similarity among radiomics features. At each iteration of the clustering, two of the most similar clusters were combined into one cluster and then acted as one cluster for the next iteration. A cluster node in the clustering tree could be one individual radiomics features or several radiomics features. A detailed description of unsupervised hierarchical clustering can be found in our previous publication (10).

Association Between Confounding Factors and Survival Outcomes

In this study, a direct association between confounding factors and survival outcomes was also studied. The two confounding factors were CT slice thickness and tumor size. The information of CT slice thickness was retrieved from DICOM attributes tagged as (0018,0050). Patients were assigned to two subgroups with slice thickness ≤3 mm (74 patients) and >3 mm (109 patients). With respect to tumor size, patients were assigned to two subgroups with tumor size less than or equal to the median size value (60 patients) and greater than the median value (123 patients).

Statistical Analysis

Unsupervised clustering and principle component analyses were used to stratify the cohort into two groups/phenotypes. The association of this radiomics clustering/phenotypic binary classification was tested primarily with survival outcomes (OS, CSS and RFS) using Kaplan–Meier curves and Cox-regression models. Secondarily, the radiomics cluster's association with other patient's and tumor's characteristics (including demographic characteristics (age, gender and race), pathological characteristics (tumor grade), American Joint Committee on Cancer tumor, node, metastasis staging (AJCC TNM staging), and genetic characteristics (VHL, PBRM1, SETD2, BAP1, and KDM5C) using Chi-Square and T-test when appropriate. P-values smaller than 0.05 indicated statistical significance. All statistical analyses were performed using Matlab 2020a.

RESULTS

Patient Characteristics and CT Examination

A total of 183 patients were included in our study according to the inclusion and the exclusion criteria. The patient characteristics are presented in **Table 1**. Patients' average age was 60 years (± standard deviation (std) of 12). Majority of patients were men (66%) and white (96%). The mean ± std of tumor size was 6.4 ± 3.2 cm. The minimum and maximum of tumor size were 1.5 and 15.5 cm, respectively. The cohort was close to be evenly split between early stage (52% had stage I) and advanced stage (48% has stages II–IV). The CT scan characteristics are presented in **Table 2**. Most of the patients were scanned by the same vender CT scanner (GE Medical System, 85%) but with different slice thicknesses; 60% had thin CT slices scans while 40% had thick CT slices scans. A more detailed CT characteristics were provided in **Supplement S2**.

TABLE 1 | Patient characteristics.

Patient characteristics	Total patients (n = 183)
Age, year	59.9 (± 11.7)
Gender	
Female	62 (34%)
Male	121 (66%)
Race	
White	176 (96%)
Others	7 (4%)
Tumor grade	
G1	1 (1%)
G2	72 (39%)
G3	79 (43%)
G4	31 (17%)
AJCC TNM staging	
Stage I	96 (52%)
Stage II	14 (8%)
Stage III	48 (26%)
Stage IV	25 (14%)
Distant Metastasis	
M0	160 (87%)
M1	23 (13%)
VHL mutation	
Positive	100 (55%)
Negative	71 (38%)
Not available	12 (7%)
PBRM1 mutation	
Positive	52 (28%)
Negative	119 (65%)
Not available	12 (7%)
SETD2 mutation	
Positive	14 (8%)
Negative	157 (86%)
Not available	12 (7%)
BAP1 mutation	
Positive	16 (9%)
Negative	155 (85%)
Not available	12 (7%)
KDM5C mutation	
Positive	8 (4%)
Negative	163 (89%)
Not available	12 (7%)

Values are presented as n (%) for categorical variables and mean (± std) for continuous variables.

Ablation Analysis Based on Survival Outcome

As shown in **Figure 1**, radiomics-based analysis consisted of four experiments. The results of the four corresponding experiments are presented in **Table 3**.

In experiment #1, all the features were used to create a radiomics signature without any exclusion. In this situation, the radiomics signature was not associated with any of the survival outcomes (OS, CSS and RFS, p-value >0.05). In experiment #2, the redundant radiomics features were excluded leaving in only the redundancy-

controlled radiomics set yielding a radiomics signature that was significantly associated with OS (HR (95% CI) = 1.8 (1.0–3.3), p-value <0.05), CSS (HR (95% CI) = 2.0 (1.0–4.1), p-value <0.05), and RFS (HR (95% CI) = 2.6 (1.1–6.2), p-value <0.05). The radiomics signature in experiment #2 included striking fewer radiomics features (89/1160 = 7.6%). However, a correlation analysis showed that the second radiomics signature had a high correlation with CT scan slice thickness (p-value <0.001) which is an image acquisition parameter that should not be associated with clinical outcomes. This may be attributed to selection bias inherent in retrospective multi-institutional cohort studies when technical parameters are different between institutions, along with inter-institutional differences in tumor stage, grade, or aggressiveness tumors. Thus, in experiment #3, we used a radiomics feature set used in experiment #2 but after further exclusion of CT slice thickness dependent radiomics features. The third radiomics signature continued to be significantly associated with OS, CSS, and RFS (all p-value <0.05) with even higher magnitude of association; HR (95% CI) increasing to 2.6 (1.5–4.4), 13.7 (7.1–26.5), and 8.0 (3.8–17.0), respectively. The high HRs on predicting patients' outcomes indicated that the third radiomics signature was a powerful prognostic signature, especially on predicting CSS. However, experiment #3 radiomics signature continued to be associated with tumor size which is an information readily available through routine clinical imaging without the need for complex radiomics analysis. In order to render this radiomics signature independent of ccRCC tumor size, in experiment #4 we further excluded radiomics features (from the set used in experiment#3 model) that are highly correlating with tumor size (C-Index = 0.877) to yield a tumor-size independent radiomics signature. Final results showed that the well-controlled radiomics signature from experiment #4 was significantly associated with CSS (HR (95% CI) = 3.0 (1.5–5.7), p-value <0.05) and RFS (HR (95% CI) = 6.6 (3.1–14.1), p-value <0.05) for ccRCC patients, but was not significantly associated with OS (p = 0.06).

In addition, the associations between confounding factors with survival outcomes were also studied. As shown in **Supplement S1 Figure 1**, there was a significant association between CT slice thickness and patient's OS (HR (95% CI) = 2.0 (1.2–3.5), p-value <0.01), CSS (HR (95% CI) = 2.0 (1.0–4.0), p-value <0.01) and RFS (HR (95% CI) = 3.6 (1.6–8.0), p-value <0.01), respectively. In **Supplement S1 Figure 2**, there was expected significant association between tumor size and patient's OS (HR (95% CI) = 2.9 (1.2–5.6), p-value <0.01), CSS (HR (95% CI) = 6.0 (3.0–12.2), p-value <0.01) and RFS (HR (95% CI) = 5.0 (2.3–11.4), p-value <0.01). These two association studies revealed that the patient data in the TCGA-KIRC project were indeed factor-biased data within which real imaging phenotypical signals were suppressed.

TABLE 2 | CT scan characteristics.

CT scan characteristics	Total patients (n = 183)
Scanner manufacturer	
GE Medical System	156 (85%)
SIEMENS	24 (13%)
Philips	3 (2%)
CT slice thickness	
thin section (≤3 mm)	109 (60%)
thick section (>3 mm)	74 (40%)
overall	3.63 ± 1.51, 1.25, 7.5
Current-time product (mAs)	324 ± 124, 101, 686
Pixel spacing (mm)	0.81 ± 0.10, 0.59, 0.97
Voltage (kVp)	
120	172 (94%)
130 or 140	11 (6%)

Values are presented as frequency (%) for categorical variables and mean ± std, minimum and maximum for continuous variables.

Radiomics Phenotypes

The radiomics feature set used in experiment #4 was our final set to be implemented in constructing final most parasomnias, scanning parameter-independent, and tumor size-independent radiomics signature model classifying the study cohort into two major phenotypes; referred hereafter as radiomics phenotype I (RAD1) and radiomics phenotype II (RAD2). Demographics, pathological characteristics, clinical parameters, and gene mutation status are presented in **Table 4**. There was no statistically significant difference between RAD1 and RAD2 clusters, except in regard the AJCC staging; almost three quarters of patients with RAD1 radiomic signature had stage I tumor while less than half of patients in RAD2 cluster had stage I disease (72% vs 40%, p-value <0.01). No significant difference between the two radiomics phenotypes in terms of gene mutation status as can be seen in **Table 4**.

RAD1 radiomics phenotype included 71 patients and RAD2 phenotype included 112 patients. RAD1 cluster was reflective of the less aggressive ccRCC, in comparison to RAD2 cluster, consistently associated with lower AJCC cancer stage and with better cancer-specific and recurrence-free survival as reflected in **Figure 2**. In terms of overall survival, RAD1 tended to have better survival also but the association was not statistically significant. The most striking divergence of survival is noticed in the recurrence-free survival; Cox-regression hazard ratio of RAD2 vs. RAD1 was HR (95% CI) = 6.6 (3.1–14.1), p-value <0.05.

DISCUSSION

In this study, we demonstrated a proof of concept to remove redundant, CT slice thickness-dependent (biased), and tumor-size dependent radiomics features towards building a concise radiomics signature in patients with ccRCC. Furthermore, we demonstrated that the final most parsimonious radiomics signature model stratified this multi-institutional cohort into two major radiomics phenotypes that are significantly different by AJCC staging, CSS, and RFS. However, the radiomics signature model was not associated with genetic mutation status nor with any other available patient or tumor characteristic. In this study we have demonstrated how radiomics models can be negatively impacted by confounders

TABLE 3 | Results of the four designed experiments.

Experiment		Feature Exclusion and Dimension Reduction				Survival Outcome			Supplementary Experiment	
#	Purpose	CT Slice Thickness	Tumor Size	Principal Component Analysis	Num of Feature Dimensions	OS (HR (95%CI) and log-rank test)	CSS (HR (95%CI) and log-rank test)	RFS (HR (95%CI) and log-rank test)	Correlation to CT Slice Thickness (Chi-square)	Correlation to Tumor Size (C-Statistic)
1	**Study All Features**				1,160	1.02 (0.59–1.75) 0.929	1.04 (0.53–2.01) 0.905	1.17 (0.55–2.51) 0.674	**<0.001**	0.628
2	**Study Redundancy Effect**			×	89	1.79 (0.98–3.29) **0.033**	1.95 (0.93–4.08) **0.043**	2.63 (1.11–6.21) **0.009**	**<0.001**	0.605
3	**Study Scanning Parameter Effect**	×		×	86	2.58 (1.51–4.42) **0.002**	13.72 (7.12–26.5) **<0.001**	7.98 (3.76–16.9) **<0.001**	0.872	**0.877**
4	**Study Tumor Size Effect**	×	×	×	81	1.74 (1.01–2.99) 0.0582	2.95 (1.52–5.72) **0.007**	6.59 (3.09–14.1) 1<0.00	0.188	0.667

The bold values represent p <0.05 indicates significance. C-index >0.8 indicates high correlation.

(like CT slice thickness) and tumor size (a previously proven prognosticator) leading to false/overvalued signatures. As shown in **Figure 1**, the addition of the controlling procedure after the feature extraction is a supplement to the existing standard radiomics guideline (9), and should be helpful for medical image analysis whose data were usually relatively small and with high heterogeneity of imaging protocol.

Our analysis framework followed an ablation study paradigm, *i.e.*, investigating factors sequentially, from feature redundancy to imaging parameters (CT slice thickness) to tumor size in order to evaluate the effect of each factor to the final radiomics signature. For the feature redundancy, we have shown that there was a large redundancy existing within the radiomics feature set. A raw feature set containing 1,160 features could be efficiently represented by only 89 dimensions of principal components, which represent 7.6% of the original radiomics features. The reason for the existence of redundancy is because feature extractors, including other widely used extractors (20) (e.g., PyRadiomics (38) and IBEX (37)), were based on a number of basic feature extraction algorithms (*e.g.*, Wavelet features (42), Gray-Level Co-occurrence Matrix (GLCM) features (43), etc.), which contained multiple tunable parameters aiming to extract the features in multi-scales for the sake of not missing any valuable image patterns (33). Thus, it is highly recommended that the removal of feature redundancy be the first step when initializing a radiomics analysis. It's also notable that unsupervised machine-learning methods, such as non-parametric principal component decomposition (39) and unsupervised hierarchical clustering (10), were recommended for redundancy removal and radiomics signature building. Compared to the supervised methods, the unsupervised machine-learning methods generally have a lower risk on overfitting the problem, because little or no prior knowledge is needed for the learning parameters.

For image acquisition parameters, our study showed that ccRCC patients imaged with different slice thicknesses were associated with significantly different survival outcomes, which is not biologically plausible and certainly is attributed to inherent bias in retrospective studies. The patients with thicker CT scan slices thickness were of more aggressive tumors when compared to patients with thinner CT scan slices (See **Supplement S1 Figure 1(A)**, thick vs. thin slice thickness group was of HR of recurrence (95% CI) = 3.6 (1.6–8.0), p <0.01). We believe this apparent association is because institutions that contributed ccRCC to the TCGA and TCIA with thicker CT slices happened to be contributing ccRCC tumors with more aggressive behavior (i.e. larger tumors with higher stage of disease). If the effect of slice thickness on radiomics features is not attended to and controlled for, we would have committed an error by producing a radiomics signature that is dependent on the slice thickness of the CT scan and therefore completely false.

ccRCC tumor size has long been identified as an important prognosticator and it is easily measured on routine abdominal imaging without the need for advanced processing or radiomics. Our study demonstrated that tumor size-dependent radiomics features may exaggerate the clinical utility of radiomics and may mask the real/tumor size-independent radiomics clinical utility. Size independent radiomics features are reflective of tumor textural heterogeneity will ultimately provide additional prognostic information separate from tumor size measurement which is routinely implemented clinically (*e.g.*, clinical staging for kidney cancer (44), RECIST 1.1 (45)). In this study, we introduced a method to remove the effect of tumor size from the radiomics signature models built to yield a size-independent radiomics signature with more valuable input into the tumor internal environment.

In summary, there were two main findings in our work: in retrospective multi-institutional imaging data with heterogenous techniques, image acquisition parameters could lead to false radiomics signatures while size-dependent radiomics may yield overvalued clinical utility of radiomics signature. Unfortunately, there is still suboptimal awareness of these two pitfalls in radiomics literature, although some researchers have tried to establish quality assurance criteria for radiomics study (46).

TABLE 4 | Demographic, clinical, pathological, and genetic characteristics of the final radiomics phenotypes.

Patient characteristics		Radiomics Phenotype I (Low-risk, n = 71)	Radiomics Phenotype II (High-risk, n = 112)	p
Age, year		62 (± 12)	59 (± 11)	0.148
Gender				0.886
	Female	24 (34%)	38 (34%)	
	Male	47 (66%)	74 (66%)	
Race				0.330
	White	68 (96%)	108 (96%)	
	Others	3 (4%)	4	
Tumor grade				0.227
	G1	1 (1%)	0 (0%)	
	G2	31 (44%)	41 (37%)	
	G3	31 (44%)	48 (43%)	
	G4	8 (11%)	23 (21%)	
AJCC TNM staging				<0.01**
	Stage I	51 (72%)	45 (40%)	
	Stage II	1 (1%)	13 (12%)	
	Stage III	13 (18%)	35 (31%)	
	Stage IV	6 (8%)	19 (17%)	
Distant Metastasis				0.267
	M0	65 (92%)	95 (85%)	
	M1	6 (8%)	17 (15%)	
VHL mutation				0.068*
	Positive	43 (68%)	57 (53%)	
	Negative	20 (32%)	51 (47%)	
	Not available	8 (–)	4 (–)	
PBRM1 mutation				0.207
	Positive	15 (24%)	37 (34%)	
	Negative	48 (76%)	71 (66%)	
	Not available	8 (–)	4 (–)	
SETD2 mutation				0.843
	Positive	6 (10%)	8 (7%)	
	Negative	57 (90%)	100 (93%)	
	Not available	8 (–)	4 (–)	
BAP1 mutation				0.746
	Positive	8 (13%)	8	
	Negative	55 (87%)	100	
	Not available	8 (–)	4 (–)	
KDM5C mutation				
	Positive	6 (10%)	2 (2%)	0.055*
	Negative	57 (90%)	106 (98%)	
	Not available	8 (–)	4 (–)	

*Values are presented as n (%) for categorical variables and mean (± std) for continuous variables. **indicates high significance with p<0.05, and *indicates weak significance with a p-value between 0.05 and 0.10.*

For image acquisition parameters, most of previous studies on the effect of image acquisition parameters focused on studying feature reproducibility and model generalization. Such studies could only result in conclusions that the heterogeneity in image acquisition parameters decreased the reproducibility and performance of radiomics signatures. Our study, for the first time, showed that the effect of image acquisition parameters could be severe enough to lead to false biologically unplausible association. Attention to and control for imaging acquisition parameters that are influential of radiomics is of crucial importance in retrospective studies, especially multi-institutional ones.

For tumor size, the risk of overvalued radiomics signature induced by tumor-size-effect was mainly caused by a large portion of radiomics features that were basically 'mixture' features, which characterized tumor size and intratumor imaging pattern simultaneously, such as Gray-Level Run Length Matrix (GLRLM) (47) and Gray-Level Size Zone Matrix (GLSZM) (48). The contribution weights between tumor size and image patterns to the final feature value were variable depending on the specific tumor phenotypes. Unfortunately, in most of radiomics feature extraction packages, such as PyRadiomics, IBEX, etc., the effect of size on 'mixture' features was not well studied (33). Thus, the size effect in 'mixture' feature may be easily overlooked and may lead to an over-valued size-dependent radiomics signature. The prognostic information from such a signature will overlap with the prognostic information already retrieved by measuring tumor size. For example, Mattea et al. (15) tested a radiomics signature previously shown to have predictive values on survival outcome among head and neck cancer patients, but eventually, this radiomic signature was found to be a surrogate for tumor size.

The limitations of our work include the following points. First, the number of patients in the TCGA-KIRC project was

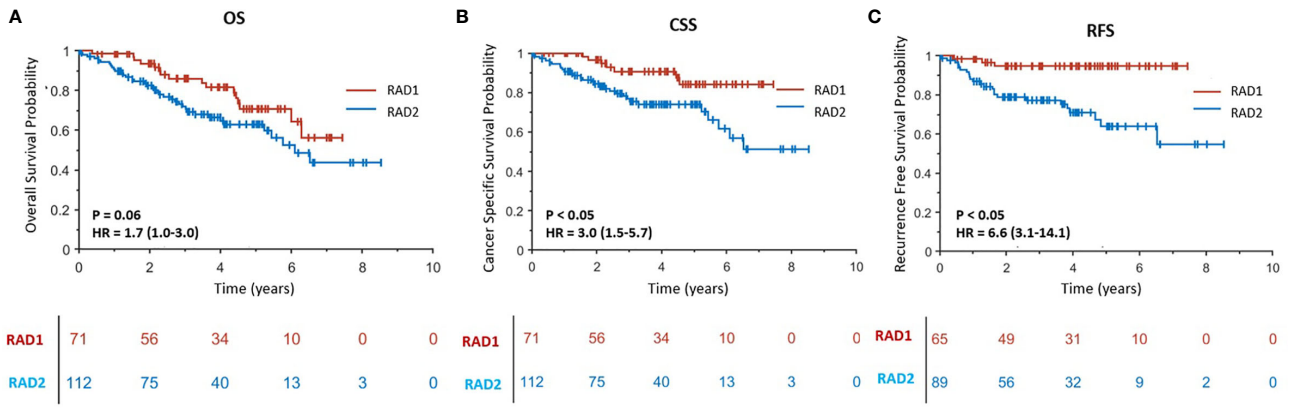


FIGURE 2 | **(A–C)** are Kaplan–Meier curves displaying the association between radiomics phenotypes and patients' OS, CSS and RFS, respectively. RAD1 and RAD2 represent radiomics phenotypes I and II.

relatively small and there were no further external data to validate our final radiomics signature. Second, except for slice thickness, other image acquisition parameters, such as scanning mode and reconstruction kernel, were not studied because most of the CT scanning (85%, see **Table 1**) were performed on scanners manufactured by the same vender with similar smooth reconstruction kernels. Fortunately, since the TCIA is a rapidly developing open research community supported by National Cancer Institute, it is promising that more and more projects such as the TCGA-KIRC will be created/improved and available in near future. Third, using PCD for dimension reduction could lead to losing of features' original physical quantification and make the modeling a black box which is difficult for interpret. Finally, in this study, we only used single-modal imaging and unsupervised machine learning algorithms for modeling, however, as the data increase in future, we could investigate multimodal imaging and supervised machine learning algorithms which have shown promising results in recent years (49–53).

CONCLUSION

In this paper, we demonstrated that a radiomics signature could be negatively impacted by multiple factors, including radiomics redundancy from large-scale feature extraction, biases from image acquisition parameters, and underlying dependency to established clinical prognosticator (tumor size). Proper attention to and control for these pitfalls are needed to guarantee a reliable, reproducible, and clinically relevant radiomics signature. Our work used the prediction of survival outcomes in ccRCC patients as an example. In our study, the final most concise, slice thickness independent, and tumor size-independent radiomics signature stratified multi-institutional retrospective cohort of ccRCC into two distinct phenotypes that are significantly different in tumor stage, CSS and RFS.

DATA AVAILABILITY STATEMENT

The data we used were downloaded from the TCGA-KIRC project, which is publicly available in the TCIA dataset (https://wiki.cancerimagingarchive.net/display/Public/TCGA-KIRC).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by dataset downloaded from The Cancer Genome Atlas Program supported by National Cancer Institute. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LLu: Conceptualization, methodology, formal analysis, and writing (original draft). FA: Conceptualization, methodology, data curation, and writing (extensive review and editing). OA: Resource, data curation and funding acquisition. LLuk: Data curation. XG: Software. HY: Data preparation. JY: Writing (review and editing). AH: Data curation. LS: Conceptualization, discussion, and funding acquisition. BZ: Conceptualization, discussion, and funding acquisition. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.638185/full#supplementary-material

REFERENCES

1. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5644
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
3. Li R, Xing L, Napel S, Rubin DL. *Radiomics and Radiogenomics: Technical Basis and Clinical Applications*. Boca Raton, Florida: CRC Press (2019).
4. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting Malignant Nodules From Screening Ct Scans. *J Thorac Oncol* (2016) 11(12):2120–8. doi: 10.1016/j.jtho.2016.07.002
5. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. *Cancer Res* (2017) 77(14):3922–30. doi: 10.1158/0008-5472.CAN-17-0122
6. Oikonomou A, Khalvati F, Tyrrell PN, Haider MA, Tarique U, Jimenez-Juan L, et al. Radiomics Analysis At PET/CT Contributes to Prognosis of Recurrence and Survival in Lung Cancer Treated With Stereotactic Body Radiotherapy. *Sci Rep* (2018) 8(1):4003. doi: 10.1038/s41598-018-22357-y
7. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A Radiomics Approach to Assess Tumour-Infiltrating CD8 Cells and Response to anti-PD-1 or anti-PD-L1 Immunotherapy: An Imaging Biomarker, Retrospective Multicohort Study. *Lancet Oncol* (2018) 19(9):1180–91. doi: 10.1016/S1470-2045(18)30413-3
8. He B, Dong D, She Y, Zhou C, Fang M, Zhu Y, et al. Predicting Response to Immunotherapy in Advanced non-Small-Cell Lung Cancer Using Tumor Mutational Burden Radiomic Biomarker. *J Immunother Cancer* (2020) 8(2): e000550. doi: 10.1136/jitc-2020-000550
9. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141
10. Lu L, Ehmke RC, Schwartz LH, Zhao B. Assessing Agreement Between Radiomic Features Computed for Multiple Ct Imaging Settings. *PLoS One* (2016) 11(12):e0166550. doi: 10.1371/journal.pone.0166550
11. Berenguer R, Pastor-Juan MDR, Canales-Vazquez J, Castro-Garcia M, Villas MV, Mansilla Legorburo F, et al. Radiomics of CT Features May be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. *Radiology* (2018) 288(2):407–15. doi: 10.1148/radiol.2018172361
12. Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L, et al. Reproducibility of Radiomics for Deciphering Tumor Phenotype With Imaging. *Sci Rep* (2016) 6:23428. doi: 10.1038/srep23428
13. He L, Huang Y, Ma Z, Liang C, Liang C, Liu Z. Effects of Contrast-Enhancement, Reconstruction Slice Thickness and Convolution Kernel on the Diagnostic Performance of Radiomics Signature in Solitary Pulmonary Nodule. *Sci Rep* (2016) 6:34921. doi: 10.1038/srep34921
14. Li Y, Lu L, Xiao M, Dercle L, Huang Y, Zhang Z, et al. Ct Slice Thickness and Convolution Kernel Affect Performance of a Radiomic Model for Predicting Egfr Status in Non-Small Cell Lung Cancer: A Preliminary Study. *Sci Rep* (2018) 8(1):17913. doi: 10.1038/s41598-018-36421-0
15. Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of Radiomic Signature Development: The Need for Safeguards. *Radiother Oncol* (2019) 130:2–9. doi: 10.1016/j.radonc.2018.10.027
16. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
17. Vallieres M, Zwanenburg A, Badic B, Cheze Le Rest C, Visvikis D, Hatt M. Responsible Radiomics Research for Faster Clinical Translation. *J Nucl Med* (2018) 59(2):189–93. doi: 10.2967/jnumed.117.200501
18. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of A Method to Compensate Multicenter Effects Affecting CT Radiomics. *Radiology* (2019) 291(1):53–9. doi: 10.1148/radiol.2019182023
19. Da-Ano R, Visvikis D, Hatt M. Harmonization Strategies for Multicenter Radiomics Investigations. *Phys Med Biol* (2020) 65(24):24TR02. doi: 10.1088/1361-6560/aba798
20. Chow WH, Dong LM, Devesa SS. Epidemiology and Risk Factors for Kidney Cancer. *Nat Rev Urol* (2010) 7(5):245–57. doi: 10.1038/nrurol.2010.46
21. Storkel S, van den Berg E. Morphological Classification of Renal Cancer. *World J Urol* (1995) 13(3):153–8. doi: 10.1007/BF00184870
22. Nazari M, Shiri I, Hajianfar G, Oveisi N, Abdollahi H, Deevband MR, et al. Noninvasive Fuhrman Grading of Clear Cell Renal Cell Carcinoma Using Computed Tomography Radiomic Features and Machine Learning. *Radiol Med* (2020) 125(8):754–62. doi: 10.1007/s11547-020-01169-z
23. Yang R, Wu J, Sun L, Lai S, Xu Y, Liu X, et al. Radiomics of Small Renal Masses on Multiphasic CT: Accuracy of Machine Learning-Based Classification Models for the Differentiation of Renal Cell Carcinoma and Angiomyolipoma Without Visible Fat. *Eur Radiol* (2020) 30(2):1254–63. doi: 10.1007/s00330-019-06384-5
24. Ursprung S, Beer L, Bruining A, Woitek R, Stewart GD, Gallagher FA, et al. Radiomics of Computed Tomography and Magnetic Resonance Imaging in Renal Cell Carcinoma-a Systematic Review and Meta-Analysis. *Eur Radiol* (2020) 30(6):3558–66. doi: 10.1007/s00330-020-06666-3
25. Nazari M, Shiri I, Zaidi H. Radiomics-Based Machine Learning Model to Predict Risk of Death Within 5-Years in Clear Cell Renal Cell Carcinoma Patients. *Comput Biol Med* (2021) 129:104135. doi: 10.1016/j.compbiomed.2020.104135
26. Gulati S, Martinez P, Joshi T, Birkbak NJ, Santos CR, Rowan AJ, et al. Systematic Evaluation of the Prognostic Impact and Intratumour Heterogeneity of Clear Cell Renal Cell Carcinoma Biomarkers. *Eur Urol* (2014) 66(5):936–48. doi: 10.1016/j.eururo.2014.06.053
27. Tang PA, Vickers MM, Heng DY. Clinical and Molecular Prognostic Factors in Renal Cell Carcinoma: What We Know So Far. *Hematol Oncol Clin North Am* (2011) 25(4):871–91. doi: 10.1016/j.hoc.2011.04.003
28. Akin O, Elnajjar P, Heller M, Jarosz R, Erickson B, Kirk S, et al. *Radiology Data From the Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma [TCGA-KIRC] Collection*. The Cancer Imaging Archive (2016).
29. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J Digit Imaging* (2013) 26(6):1045–57. doi: 10.1007/s10278-013-9622-7
30. Hakimi AA, Ostrovnaya I, Reva B, Schultz N, Chen YB, Gonen M, et al. Adverse Outcomes in Clear Cell Renal Cell Carcinoma With Mutations of 3p21 Epigenetic Regulators BAP1 and SETD2: A Report by MSKCC and the KIRC TCGA Research Network. *Clin Cancer Res* (2013) 19(12):3259–67. doi: 10.1158/1078-0432.CCR-12-3886
31. Guo X, Schwartz LH, Zhao B. Automatic Liver Segmentation by Integrating Fully Convolutional Networks Into Active Contour Models. *Med Phys* (2019) 46(10):4455–69. doi: 10.1002/mp.13735
32. Ahmed FS, Akin O, Shaish H, Luk L, Guo X, Yang H, et al. Nonenhancing Component of Clear Cell Renal Cell Carcinoma on Computed Tomography Correlates With Tumor Necrosis and Stage and Serves as a Size-Independent Prognostic Biomarker. *J Comput Assist Tomogr* (2019) 43(4):628–33. doi: 10.1097/RCT.0000000000000877
33. Lu L, Sun SH, Yang H, E L, Guo P, Schwartz LH, et al. Radiomics Prediction of EGFR Status in Lung Cancer-Our Experience in Using Multiple Feature Extractors and The Cancer Imaging Archive Data. *Tomography* (2020) 6 (2):223–30. doi: 10.18383/j.tom.2020.00017
34. Dercle L, Lu L, Schwartz LH, Qian M, Tejpar S, Eggleton P, et al. Radiomics Response Signature for Identification of Metastatic Colorectal Cancer Sensitive to Therapies Targeting EGFR Pathway. *J Natl Cancer Inst* (2020) 112(9):902–12. doi: 10.1093/jnci/djaa017
35. Dercle L, Fronheiser M, Lu L, Du S, Hayes W, Leung DK, et al. Identification of Non–Small Cell Lung Cancer Sensitive to Systemic Cancer Therapies Using Radiomics. *Clin Cancer Res* (2020) 26(9):2151–62. doi: 10.1158/1078-0432.CCR-19-2942
36. Lu L, Wang D, Wang L, E L, Guo P, Li Z, et al. A Quantitative Imaging Biomarker for Predicting Disease-Free-Survival-Associated Histologic Subgroups in Lung Adenocarcinoma. *Eur Radiol* (2020) 30(7):3614–23. doi: 10.1007/s00330-020-06663-6
37. Zhang L, Fried DV, Fave XJ, Hunter LA, Yang J, Court LE. IBEX: An Open Infrastructure Software Platform to Facilitate Collaborative Work in Radiomics. *Med Phys* (2015) 42(3):1341–53. doi: 10.1118/1.4908210
38. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.CAN-17-0339

39. Abdi H, Williams LJ. Principal Component Analysis. *Wiley Interdiscip Rev: Comput Stat* (2010) 2(4):433–59. doi: 10.1002/wics.101
40. Kodama F, Fultz PJ, Wandtke JC. Comparing Thin-Section and Thick-Section CT of Pericardial Sinuses and Recesses. *AJR Am J Roentgenol* (2003) 181 (4):1101–8. doi: 10.2214/ajr.181.4.1811101
41. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei L. On the C-statistics for Evaluating Overall Adequacy of Risk Prediction Procedures With Censored Survival Data. *Stat Med* (2011) 30(10):1105–17. doi: 10.1002/sim.4154
42. Mallat SG. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans Pattern Anal Mach Intell* (1989) 11(7):674–93. doi: 10.1109/34.192463
43. Haralick RM, Shanmugam K, Dinstein IH. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, Cybernetics* (1973) (6):610–21. doi: 10.1109/TSMC.1973.4309314
44. Edge SB. *Ajcc Cancer Staging Manual. 8th Ed*. New York, NY: Springer (2017).
45. Schwartz LH, Litiere S, de Vries E, Ford R, Gwyther S, Mandrekar S, et al. Recist 1.1-Update and Clarification: From the RECIST Committee. *Eur J Cancer* (2016) 62:132–7. doi: 10.1016/j.ejca.2016.03.081
46. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
47. Tang X. Texture Information in Run-Length Matrices. *IEEE Trans image Process* (1998) 7(11):1602–9. doi: 10.1109/83.725367
48. Thibault G, Fertil B, Navarro C, Pereira S, Cau P, Levy N, et al. Shape and Texture Indexes Application to Cell Nuclei Classification. *Int J Pattern Recognit Artif Intell* (2013) 27(01):1357002. doi: 10.1142/S0218001413570024
49. Shiri I, Maleki H, Hajianfar G, Abdollahi H, Ashrafinia S, Hatt M, et al. Next-Generation Radiogenomics Sequencing for Prediction of EGFR and KRAS Mutation Status in NSCLC Patients Using Multimodal Imaging and Machine Learning Algorithms. *Mol Imaging Biol* (2020) 22(4):1132–48. doi: 10.1007/s11307-020-01487-8
50. Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ. Radiomic Machine-Learning Classifiers for Prognostic Biomarkers of Head and Neck Cancer. *Front Oncol* (2015) 5:272. doi: 10.3389/fonc.2015.00272
51. Hajianfar G, Shiri I, Maleki H, Oveisi N, Haghparast A, Abdollahi H, et al. Noninvasive O6 Methylguanine-Dna Methyltransferase Status Prediction in Glioblastoma Multiforme Cancer Using Magnetic Resonance Imaging Radiomics Features: Univariate and Multivariate Radiogenomics Analysis. *World Neurosurg* (2019) 132:e140–61. doi: 10.1016/j.wneu.2019.08.232
52. Leger S, Zwanenburg A, Pilz K, Lohaus F, Linge A, Zophel K, et al. A Comparative Study of Machine Learning Methods for Time-to-Event Survival Data for Radiomics Risk Modelling. *Sci Rep* (2017) 7(1):13206. doi: 10.1038/s41598-017-13448-3
53. Rastegar S, Vaziri M, Qasempour Y, Akhash MR, Abdalvand N, Shiri I, et al. Radiomics for Classification of Bone Mineral Loss: A Machine Learning Study. *Diagn Interv Imaging* (2020) 101(9):599–610. doi: 10.1016/j.diii.2020.01.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 27 May 2021
doi: 10.3389/fonc.2021.664304

Targeting Treatment Resistance in Head and Neck Squamous Cell Carcinoma – Proof of Concept for CT Radiomics-Based Identification of Resistant Sub-Volumes

Marta Bogowicz[1†], Matea Pavic[1†], Oliver Riesterer[1,2], Tobias Finazzi[1], Helena Garcia Schüler[1], Edna Holz-Sapra[1], Leonie Rudofsky[1], Lucas Basler[1], Manon Spaniol[1], Andreas Ambrusch[1], Martin Hüllner[3], Matthias Guckenberger[1] and Stephanie Tanadini-Lang[1]*

[1] Department of Radiation Oncology, University Hospital Zurich, University of Zurich, Zurich, Switzerland, [2] Centre for Radiation Oncology KSA-KSB, Cantonal Hospital Aarau, Aarau, Switzerland, [3] Department of Nuclear Medicine, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Purpose: Radiomics has already been proposed as a prognostic biomarker in head and neck cancer (HNSCC). However, its predictive power in radiotherapy has not yet been studied. Here, we investigated a local radiomics approach to distinguish between tumor sub-volumes with different levels of radiosensitivity as a possible target for radiation dose intensification.

Materials and Methods: Of 40 patients (n=28 training and n=12 validation) with biopsy confirmed locally recurrent HNSCC, pretreatment contrast-enhanced CT images were registered with follow-up PET/CT imaging allowing identification of controlled (GTVcontrol) vs non-controlled (GTVrec) tumor sub-volumes on pretreatment imaging. A bi-regional model was built using radiomic features extracted from pretreatment CT in the GTVrec and GTVcontrol to differentiate between those regions. Additionally, concept of local radiomics was implemented to perform detection task. The original tumor volume was divided into sub-volumes with no prior information on the location of recurrence. Radiomic features from those sub-volumes were then used to detect recurrent sub-volumes using multivariable logistic regression.

Results: Radiomic features extracted from non-controlled regions differed significantly from those in controlled regions (training AUC = 0.79 CI 95% 0.66 - 0.91 and validation AUC = 0.88 CI 95% 0.72 – 1.00). Local radiomics analysis allowed efficient detection of non-controlled sub-volumes both in the training AUC = 0.66 (CI 95% 0.56 – 0.75) and validation cohort 0.70 (CI 95% 0.53 – 0.86), however performance of this model was inferior to bi-regional model. Both models indicated that sub-volumes characterized by higher heterogeneity were linked to tumor recurrence.

OPEN ACCESS

Edited by:
Laurent Dercle, Columbia University Irving Medical Center, United States

Reviewed by:
Chirag Kamal Ahuja, Post Graduate Institute of Medical Education and Research (PGIMER), India
Swapnil Ulhas Rane, Research and Education in Cancer, India

***Correspondence:**
Marta Bogowicz
marta.bogowicz@usz.ch

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *04 February 2021*
Accepted: *06 April 2021*
Published: *27 May 2021*

Citation:
Bogowicz M, Pavic M, Riesterer O, Finazzi T, Garcia Schüler H, Holz-Sapra E, Rudofsky L, Basler L, Spaniol M, Ambrusch A, Hüllner M, Guckenberger M and Tanadini-Lang S (2021) Targeting Treatment Resistance in Head and Neck Squamous Cell Carcinoma – Proof of Concept for CT Radiomics-Based Identification of Resistant Sub-Volumes. Front. Oncol. 11:664304. doi: 10.3389/fonc.2021.664304

Conclusion: Local radiomics is able to detect sub-volumes with decreased radiosensitivity, associated with location of tumor recurrence in HNSCC in the pre-treatment CT imaging. This proof of concept study, indicates that local CT radiomics can be used as predictive biomarker in radiotherapy and potential target for dose intensification.

Keywords: local radiomics, radioresistance, head and neck cancer, predictive biomarker, contrast-enhanced CT, tumor recurrence

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) accounts for approximately 4% of all malignancies across Europe and the USA (1, 2). For locally advanced HNSCC, the standard of care is definitive radiotherapy, whenever possible combined with concurrent chemotherapy. Despite advances in treatment using modern radiotherapy delivery techniques, local recurrences still occur in up to 50% of patients and pose the predominant pattern of failure in HNSCC (3). Therapeutic options for recurrent HNSCC are mainly palliative and comprise salvage surgery, re-irradiation and systemic therapy – however, outcome is poor, and treatment often is associated with significant morbidity (4). Therefore, improvement of primary radiochemotherapy is essential.

Several early clinical trials have studied the role of radiation dose intensification to the primary tumor and metastatic lymph nodes in order to improve local control for HNSCC (5–7). However, concern still exists about excessive acute and late toxicities of this approach and therefore no large randomized trial has been conducted so far. In addition, intensification of radiotherapy by using altered fractionation schemes only lead to very modest improvement of outcome (8). Thus, a dose of 70 Gy delivered over 7 weeks to the entire tumor is widely considered a standard in patients undergoing chemoradation for HNSCC (9). Heterogeneity within a tumor has been recognized as a characteristic of head and neck cancer (10), with some sub-regions being more resistant to radiochemotherapy. Consequently, the strategy of treatment intensification to these sub-volumes could lead to better outcomes in terms of local control and subsequently overall survival, without a significant increase in treatment-related toxicities. Recent advancement in radiation technology in principle allows such an escalation of radiation dose to tumor sub-volumes. However, identification of these sub-volumes is a crucial step within this therapeutic concept.

In contrast to selective biopsy specimens obtained from a small area of the tumor, the use of imaging as biomarker has the advantage to analyze the entire three dimensional tumor volume. The feasibility of delivering a dose boost, so-called dose painting, to tumor sub-volumes has been previously demonstrated, mostly based on functional, metabolism-related (fluorodeoxyglucose) or hypoxia-related (fluoromisonidazole) positron emission tomography (PET) imaging (11–13). In recent years, high-throughput, multidimensional and quantitative images analysis (radiomics) revealed that relevant biological information can be extracted also from routinely acquired, easy to perform morphological imaging, such as regular computed tomography (CT) (14–16). Many investigations have studied the potential of radiomics to predict the risk for metastatic spread, progression-free survival, overall survival or biological phenotypes (17–19) showing encouraging results. In the context of radiotherapy, several studies indicated an association between CT heterogeneity and local tumor control (20–24). However, there is so far scarce data on radiomics as a method to individually tailor dose distribution to specific sub-volumes within a tumor (25–27).

Here, we investigated CT radiomics for identification of radioresistant sub-volumes of the primary tumor leading to persistence or recurrence after curative radiochemotherapy. Reliable detection of these resistant sub-volumes may allow for a tailored treatment by increasing radiation dose to these parts.

METHODS

Study Population

We retrospectively analyzed patients with primary locally advanced HNSCC (cT3/4 or cN+) treated at our institution between June 2004 and October 2015, who experienced a local in-field tumor recurrence. Local in-field tumor recurrence was defined as a recurrence occurring within the high-dose target volume (excluding lymph nodes), and had to be confirmed by biopsy. Only patients that received a definitive high dose radiation treatment with an equivalent total dose of 68 - 70 Gy in 2-2.11 Gy fractions and a concomitant systemic therapy with either cisplatin and/or cetuximab during the course of radiotherapy were included. A further requirement was a contrast-enhanced CT (CE-CT) imaging before treatment and a FDG-PET/CT at the time of local recurrence. Contrast protocol was not standardized, however patients with contrast visible only in thyroid were not included. Follow up of patients was done according to the institutional routine. The first FDG-PET/CT for treatment response was usually done three months after completion of treatment. A subset of patients included in this study and now reviewed retrospectively has been treated within a prospective trial back then (28).

Data analysis was approved by the Swissethics and was carried out in accordance with Swissethics guidelines and regulations. Patients gave informed general consent.

We identified in total 66 patients fulfilling the above mentioned inclusion criteria. However, twenty-one cases had to be excluded, as a reliable image registration between follow-up ^{18}F-fluorodeoxyglucose (^{18}F-FDG) PET and CE pretreatment planning CT was not possible. Five additional patients were excluded due to the small size of recurrence (volume < 27 voxels), which precludes a reasonable radiomics analysis. Thus, in total 40 patients were available and were split into a training cohort, consisting of

retrospectively collected data (n = 28), and a validation cohort, comprising patients from a prospective phase II clinical study (n = 12) (28). In the majority of cases, the primary tumor was located in the oropharynx, for further details see **Table 1**.

Contrast-Enhanced Planning CT Image Acquisition

Iodine contrast was injected intravenously prior to CT imaging. The contrast protocol varied over the years of data acquisition. Images were acquired on five different scanners (Siemens Somatom Definition AS n = 20, Siemens Volume Zoom n = 13, Siemens Somatom Plus n = 5, GE Healthcare, Discovery STE n = 1 and GE Healthcare, Discovery LS n = 1) with 120 or 140 kV tube voltage, 2 – 3.27mm slice thickness and <1mm in-plane voxel size. The tube current varied between 120-450 mAs, and angular exposure adaption was applied in 7 scans (17%). Images were reconstructed using filtered back projection and soft kernel (B30).

Definition of Primary Tumor and Recurrence Volume

The primary tumor was contoured based on the CE planning CT assisted by the co-registered pretreatment diagnostic FDG-PET/CT and MRI scans if available. However, in contrary to the FDG-PET, which was available in all the patients, a pretreatment MRI was only present in 5 cases. The recurrence region was contoured as high FDG uptake in the follow-up FDG-PET/CT at the time of the detection of the recurrence (**Supplementary Figure S1**). High FDG uptake region was considered as 40% SUVmax threshold-based sub-volume in the recurrence PET-CT. Details of the PET scanning protocol can be found in the supplement (**Supplementary Table S1**). The contours were then transferred to the initial planning CT by rigid registration with focus on the tumor region (Eclipse v.15, Varian Medical Systems, USA). Two structures were created for further analysis: the overlap of tumor recurrence and primary tumor (GTV_{rec}) and the primary tumor contour minus the recurrence ($GTV_{control}$). Prior to radiomics analysis both GTV_{rec} and $GTV_{control}$ were postprocessed, by removal of contours on slices affected by metal artifacts.

TABLE 1 | Patient characteristics of all included patients.

	Training Cohort (n=28)	Validation Cohort (n=12)
Tumor location		
Oropharynx	19 (68%)	6 (50%)
Hypopharynx	5 (18%)	4 (33%)
Larynx	3 (11%)	0 (0%)
Oral Cavity	1 (3%)	2 (17%)
Time to recurrence (median [range] months)	7 [2 - 59]	8 [4 -24]
T stage		
1	1 (3%)	0 (0%)
2	5 (18%)	0 (0%)
3	7 (25%)	3 (25%)
4	15 (54%)	9 (75%)
N stage		
0	5 (18%)	2 (17%)
1	4 (14%)	1 (8%)
2a	0 (0%)	1 (8%)
2b	7 (25%)	2 (17%)
2c	12 (43%)	6 (50%)
3	0 (0%)	0 (0%)
HPV status		
Positive	3 (11%)	3 (25%)
Negative	12 (43%)	8 (67%)
Unknown	13 (46%)	1 (8%)

Radiomics Analysis for Differentiation of Controlled to Non-Controlled Tumor Sub-Volumes and Detection of Radioresistant Volumes

Radiomics analysis was performed using Z-Rad software implementation (Python v 2.7). This implementation was benchmarked in the Image Biomarker Standardization Initiative (29, 30). Images were resampled to cubic voxels (3.3 mm, largest voxel dimension in the dataset) using trilinear interpolation. Hounsfield unit range of -20 to 180 HU was set to limit the analysis to soft tissue. Due to the small size of analyzed volumes, the feature set was limited to intensity (n = 25) and texture features (n = 136). The full list of analyzed features is presented in the **Supplementary Table S2**. Volumes with less than 27 voxels were excluded from analysis.

In total, three different radiomics models were built (**Figure 1**): The "bi-regional radiomics" analysis aims to differentiate tumor regions, which stay controlled after treatment ($GTV_{control}$) and the non-controlled region, which resulted in a recurrence detected on the FDG-PET CT (GTV_{rec}). For detection of radioresistant volumes within the primary tumor without prior knowledge on their location and prior to any treatment two "local radiomics models" were built. Here, the predefined region of interest (ROI), corresponding to the primary GTV and from now on named GTV, is divided into smaller sub-regions, which are later used as a mask for feature extraction. As a consequence, instead of one vector of features describing the GTV, local radiomics returns a set of radiomics-based parametric maps. These maps visualize for example changes in heterogeneity or contrast across the GTV. The local features are calculated based only on the voxels within the GTV, whereas the neighboring voxels (e.g. healthy tissue, manually excluded artifacts) are set to 'not a number' (discarded). We have implemented two approaches for the definition of sub-regions (**Figure 1**):

- Division of the GTV into a fixed number of sub-regions with the center of the grid attached to the center of the GTV (in this study 8 sub-regions, the number was chosen to provide sufficient information in 3D and to ensure large enough neighborhood to define texture). The same number of sub-regions was analyzed per patient and their volume depended on the GTV size.
- Division of the GTV into sub-regions with fixed size. In the second approach, the placement of the grid is optimized to cover a GTV volume as large as possible. The full coverage is rarely feasible, because in order to avoid meaningless features calculated only on a few voxels, we have set a threshold T = 25 of minimum number of ROI voxels (voxels with value different than 'not a number') in the sub-region. In this study, we decided for 5x5x5 voxels grid with 5 voxel shift. There was no overlap between sub-regions to ensure

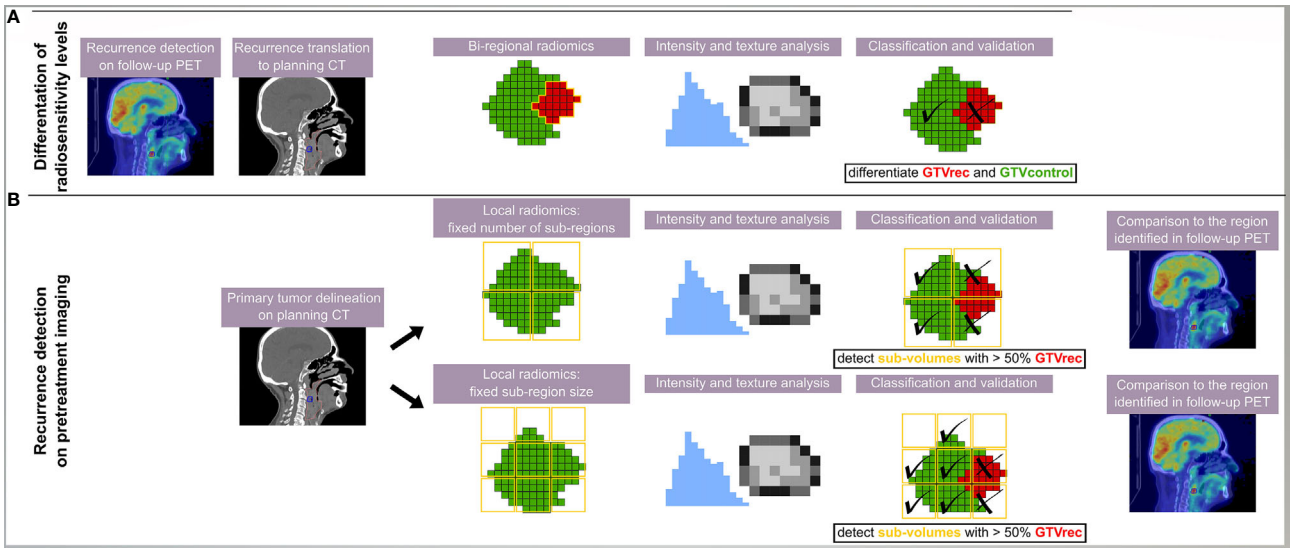


FIGURE 1 | Scheme of the analysis giving an overview on all three radiomics models. The recurrence region was identified on PET/CT imaging and rigidly transferred to the contrast-enhanced planning CT. Different models were trained using different methods and aiming at different purposes. **(A)** In the bi-regional radiomics, features were extracted from GTV_{rec} and $GTV_{control}$ and only a differentiation between recurrent and controlled sub-volumes was performed; **(B)** in two local radiomics models, a detection task was performed and thus sub-volumes were defined without any prior information on the location of recurrence. In the classification task, a sub-volume was classified as recurrence (X) if more than 50% of the voxels belonged to the original contour of the recurrence (red).

independent description of the sub-regions. The 5 voxels grid size was a trade-off between ensuring a large enough neighborhood to define texture and a high spatial resolution of the model. Other sizes of the grid were not tested. In contrary to the first approach different numbers of sub-regions were analyzed per patient but their size was constant.

Statistical Analysis

The classification and detection tasks were performed using logistic regression. In the bi-regional model, the two labels were assigned to $GTV_{control}$ and GTV_{rec} volumes. For local radiomics models, the recurrent label was assigned to regions with more than 50% of voxels overlapping with GTV_{rec}. The 50% threshold was chosen arbitrarily, other thresholds were not tested. The remaining sub-regions were labeled as control.

The following feature reduction and binary classification procedure was used. Features with high and moderate correlation to number of voxels in the local region were excluded ($r > 0.5$). Principal component analysis combined with univariate logistic regression was used for dimensionality reduction (20). The number of retained principal components was computed using the Horn method. Radiomic features were grouped based on their Spearman correlation to principal components. Per group, only the feature with the largest area under operator receiver characteristic curve (AUC) was chosen, given that p-value < 0.05. The final model was built in the multivariable logistic regression with backward selection of variables based on the Akaike Information Criterion. The final model was tested in the separate set as described in the sections below. The confidence intervals were computed with 2000 stratified bootstrap replicates.

For the bi-regional radiomics model and the first local radiomics model based on 8 sub-volumes, training was performed on the training cohort and the results validated in the validation cohort as previously specified in the `study population` section. The 75th percentile threshold of predictions in the training cohort was set as classification threshold to optimize the sensitivity and specificity of the model.

The second local radiomics model with a fixed 5x5x5 grid size, was trained and tested in leave-one-out cross-validation (LOOCV) on the patient level. The full split into training and validation cohort was not feasible due to the presence of metal artifacts in some images. These artifacts, especially when present in the middle of the GTV, strongly influence optimal placement of the grid and thus for this experiment only patients with no visible artifacts in the GTV were selected. Data from 23 patients with no visible CT artifacts were analyzed. The final model was trained on 12 patients (= 48%), who had both classes of sub-volumes (at least one sub-region with 50% contribution of recurrent voxels), which was a requirement of LOOCV.

RESULTS

Differentiation of Controlled to Non-Controlled Tumor Regions

On average, the volume of GTV_{rec} was 20% (range: 2%-71%) of the initial primary tumor volume. A logistic regression model with backward selection of variables showed a good discrimination between GTV_{rec} and $GTV_{control}$, in both the training and the validation cohort, see **Figure 2A** (training AUC = 0.79 (95%CI: 0.66 – 0.91); validation AUC = 0.88 (95%

CI: 0.72 – 1.00). This bi-regional radiomics model consisted of two features: "GLRLM gray level non-uniformity" and "merged GLCM sum entropy" - indicating that sub-volumes showing higher density heterogeneity have a higher propensity of recurrence. In the validation cohort, the model achieved high sensitivity (= 0.75) and specificity (= 0.83).

Detection of Radioresistant Sub-Volumes – Local Radiomics Model

The first model with division of the GTV into 8 equi-volume regions comprised three features: "GLCM cluster shade", "GLDZM gray level variance" and "histogram median". This local radiomics model showed slightly inferior performance (AUC = 0.70; 95%CI: 0.53–0.86) in the validation cohort compared to the bi-regional radiomics model (**Figure 2B**). In the validation cohort, in all patients in whom recurrence sub-volumes were detected, at least one of the sub-volumes was correctly identified. The median size of detected recurrence was 43% of the entire recurrence volume. This result is linked to the threshold for recurrence sub-volume definition (50%) and thus can be further improved in the future studies.

For the second definition with division of the GTV into sub-volumes of an equal size (5x5x5 voxels) the rigid grid placement was chosen on an individual patient basis in order to maximize the coverage of studied tumor. 12 patients (= 48%) had both classes of sub-volumes (recurrence and tumor control, based on a 50% criterion). The average AUC in leave-one-out cross-validation was 0.68. In three cases the AUC was below 0.5, indicating worse than random prediction. This was tracked back to a small number of analyzed sub-volumes (e.g. only two sub-volumes) or a high recurrence involvement (more than 25%) in most of the sub-volumes, meaning that in this tumor no true controls with no contribution of radioresistant clones were present. Similar features were selected in different cross validation folds: GLSZM zone size entropy (n=12), GLCM joint maximum (n=9) and histogram range (n=8).

Detailed summary of the all models together with numbers of sub-volumes used for model training and validation is presented in **Table 2**. Number of sub-volumes in the local radiomics analysis per patient is summarized in the **Supplementary Figures S2** and **S3**.

DISCUSSION

HNSCCs, exhibiting a high heterogeneity in CT images, were previously shown to respond worse to radiochemotherapy (20–24). Several prognostic radiomic signatures have been proposed recently, but studies on predictive signatures with potential impact on treatment optimization are scarce (31). In this study, we were able to differentiate intratumoral levels of radiosensitivity by means of CT radiomics. Further, we proposed an algorithm for pretreatment detection of radioresistant regions.

Identification of treatment resistant tumor sub-volumes by means of medical imaging has been previously investigated, mainly by detection on hypoxic imaging but also different imaging modalities. Tumor hypoxia is a known adverse prognostic factor for local control after radiotherapy of HNSCC (32). However, only two studies tried to correlate spatial location of tumor recurrence and initial hypoxic region, and only one of them showed overlap of those regions (33, 34). A correlation of high FDG uptake parts of a tumor to regions of more resistant tumor sub-volumes leading to recurrences has been shown (35, 36). Accordingly, FDG-based dose escalation strategies have already been exploited and proven to be feasible (37). However, also other imaging modalities like diffusion-weighted MRI or dynamic contrast-enhanced MRI, were shown to be able to identify risk factors for worse outcomes in head and neck cancer (38). Comparison studies have shown that the volumes defined by

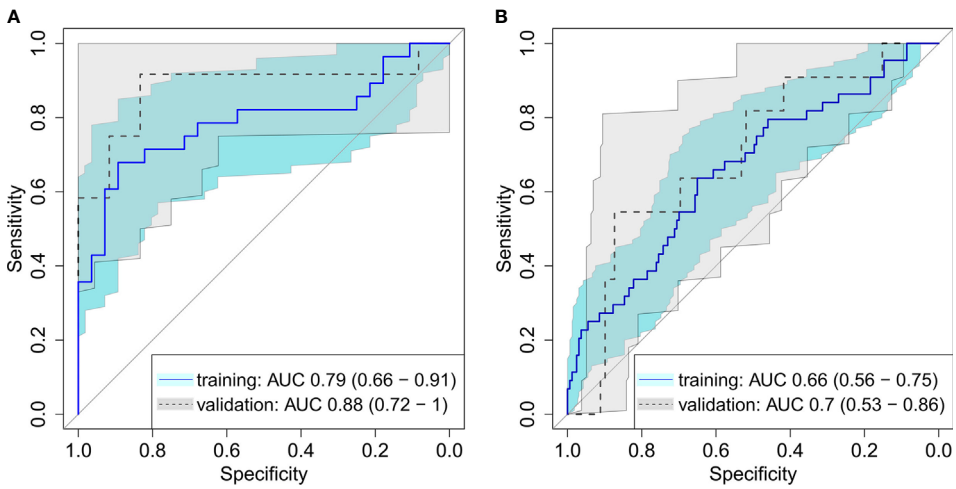


FIGURE 2 | Receiver operating characteristic for **(A)** differentiation between recurrent and controlled sub-volumes in bi-regional radiomics model showing a good discrimination between the radioresistance levels, in both the training and the validation cohort **(B)** detection of recurrent sub-volumes with local radiomics showing a good performance of the model with ability to detect radioresistant sub-volumes of the tumor on pretreatment CT images.

TABLE 2 | Details of the final radiomics models.

Model	Model endpoint	AUC training	AUC validation	Selected features	Model coefficients	No of analyzed sub-volumes/ No of recurrent sub-volumes	
						training	validation
Bi-regional radiomics	Sub-volumes differentiation	0.79 (0.66-0.91)	0.88 (0.72 – 1.00)	GLRLM gray level non-uniformity merged GLCM sum entropy intercept	141.43 4.56 -32.64	56/28	24/12
Local radiomics: fixed number of sub-volumes	Recurrence detection	0.66 (0.56 – 0.75)	0.70 (0.53 – 0.86)	GLCM cluster shade GLDZM gray level variance histogram median intercept	0.0015 0.019 -0.11 -0.24	222/48	91/11
Local radiomics: fixed size of sub-volumes	Recurrence detection	—	0.68 (AUC < 0.5 in n=3 cases)	GLSZM zone size entropy*, GLCM joint maximum* histogram range*	— — —	114 (105 -122)#/ 41 (33 - 42) #	10 (2 - 19)#/2 (1-10)#

*Area under receiver operating characteristic curve (AUC) and 95% confidence intervals. The second local radiomics model was tested in the leave-one-out cross-validation, thus no results for the training cohort are shown and the validation AUC is the average over the folds. *Most frequently chosen features over different folds. #Median and range of the number of sub-volumes analyzed over different folds.*

FDG-PET versus DWI – MRI do not overlap completely and identify distinct volumes within the primary tumor (39, 40).

Recently, Beaumont et al. (25) correlated location of tumor recurrence to pretreatment local texture features in FDG PET imaging in a small cohort of 15 patients. The performance of their model was comparable to ours (median AUC of 0.71). However, in contrast to our study, their results were only tested in leave-one-out cross-validation, owing to a small number of patients. In addition, the study did not clearly define if CT information was used together with PET information for definition of the primary tumor contour. Delineation based solely on PET images may lead to neglecting tumor regions with low FDG uptake (41). Our local radiomics results indicate that sub-volumes with higher CT heterogeneity are more radioresistant, which is in agreement with previous studies showing that higher tumor-wide CT heterogeneity is linked with reduced local control rate (20–24). Local radiomics in PET/CT has also been investigated in the context of outcome prediction in nasopharyngeal cancer, showing a higher prognostic value than with models based on entire tumor volume analysis (26). However, this study did not use local radiomics for sub-regional detection of treatment resistance sub-volumes.

With the introduction of intensity modulated radiation therapy (IMRT) the concept of delivery higher doses to head and neck tumors has regained interest. IMRT allows not only for better sparing of OARs, but also enables delivery of simultaneously higher doses to selected areas (42), so called dose painting. The identification of radioresistant tumor sub-volumes that require higher radiation doses, is an unmet clinical need. The ESCALOX study currently investigates a dose escalation up to 80.5 Gy using a simultaneous integrated boost (SIB) to the whole primary tumor and large involved lymph nodes (43), whereas another group performed a planning study with dose escalation only to hypoxic areas within the tumor as defined by ^{18}F-Fluoroazomycin arabinosid (FAZA) PET/CT and found it to be feasible while respecting the maximum OAR constraints (44). Resistance of head and neck cancer cells to radiotherapy is not conditioned by one single biological feature but rather driven by several different mechanisms (39). Consequently, a method capturing all these distinct features is desirable. The underlying biology for the distinct radiomics signature is not simple nor exactly known for the individual case. But, a correlation of radiomic signature with underlying genomic alterations and biological phenotype has been shown (45). In our study, the selection of the radioresistant sub-volumes is meant to be done by radiomics analysis, which does not account for any particular biological background/histopathological difference. However, a strong relationship between medical images, or more precisely, the extracted, quantitative imaging features, and the underlying tumor phenotype and biology was shown previously (46, 47). Thus, in contrast to the above mentioned study, which selected the radioresistant parts solely based on hypoxic regions, our radiomics approach covers a broader spectrum of underlying biological phenotypes/alterations. Still, the histopathological factors behind the selected sub-volumes in our study are unknown, and thus definition of an adequate dose boost is not straightforward and should be a subject of further research.

This study was performed on relatively small patient cohort. However, for two out of three proposed models (bi-regional and local radiomics with fixed number of sub-volumes), a successful model validation in *a priori* defined cohort was performed. Remarkably, the performance of the models was higher in the validation cohort, in comparison with the training cohort. The validation cohort is a prospective cohort with standardized imaging and treatment protocol – in comparison to the training cohort, which was retrospectively collected. Therefore, higher model performance in the validation might be a result of premature training or better data quality. In the head and neck region, HPV status is known to influence CT values distribution within primary tumor (48). Unfortunately, HPV status was unknown for a large proportion of patients (35%). Consequently, an analysis on HPV status as an effector is not possible due to the high number of missing values in the individual patient cohorts. The analyzed cohort is also heterogeneous in terms of tumor subsites, and theirs impact on the performance of the model should be further investigated. The presence of artifacts did not allow for a full

validation of the fixed grid size local model. With the introduction of iterative metal artifact reduction reconstruction algorithms in CT few years ago, this limitation of the model is addressed for future studies (49). Availability of larger data collections may also permit for testing different settings in the local radiomics analysis, such as variable size of the grid or smaller value of the translation vector, allowing for an overlap between sub-volumes. Today, it is not clear which methodology provides optimal results. In our study, we assumed that all of the analyzed tumor sub-volumes are independent and thus, no overlap was allowed. Additionally, classification was performed using logistic regression. In the future, segmentation algorithms may be tested to improve the predictive power of local radiomics.

Additional validation of the proposed approach may be derived from surgical cohorts, where pretreatment local radiomics maps may be correlated with full-mount tumor histopathology. Alternatively, if contrast-enhanced CT is present at the time of recurrence similarity of the radiomics features in the corresponding areas in the two, sequential investigations could be measured to evaluate if the signature remains stable over time.

This study has an inherent selection bias, since only patients with observed recurrences were analyzed. In the real pretreatment classification, a standard radiomics analysis can be used prior to local radiomics to preselect patients with high risk of tumor recurrence, as shown by other studies (20–24).

In conclusion, this is the first study indicating that tumor radioresistance can be localized on pretreatment CT images with validation of the radiomics model in an independent cohort. This is a proof of concept study, indicating that local CT radiomics can be used as predictive biomarker in radiotherapy and potential target for dose intensification.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee canton Zurich. The patients/participants provided general consent to participate in this study.

AUTHOR CONTRIBUTIONS

MB – performed data analysis, supervised implementation of local radiomics, wrote part of the manuscript and reviewed the manuscript. MP - contributed to study design, performed image registration and contouring, wrote part of the manuscript and reviewed the manuscript. OR - provided expertise, contributed to study design, reviewed contours and registration, and reviewed the manuscript. TF, HS, EH-S, LR and LB - performed image registration and contouring, reviewed the manuscript. MS – performed standard radiomics analysis, reviewed the manuscript, AA – implemented local radiomics, reviewed the manuscript. MH, MG and ST-L - provided expertise, contributed to study design and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This project was supported by the Swiss National Science Foundation Sinergia grant (310030_173303). The clinical study used as one of the cohorts was supported by a research grant from Merck (Schweiz) AG. The funder bodies were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.664304/full#supplementary-material

REFERENCES

1. Gatta G, Botta L, Sánchez MJ, Anderson LA, Pierannunzio D, Licitra L, et al. Prognoses and Improvement for Head and Neck Cancers Diagnosed in Europe in Early 2000s: The EUROCARE-5 Population-Based Study. *Eur J Cancer* (2015) 51(15):2130–43. doi: 10.1016/j.ejca.2015.07.043
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA: Cancer J Clin* (2019) 69(1):7–34. doi: 10.3322/caac.21551
3. Pignon J-P, Maître Al, Maillard E, Bourhis J. Meta-Analysis of Chemotherapy in Head and Neck Cancer (MACH-NC): An Update on 93 Randomised Trials and 17,346 Patients. *Radiother Oncol* (2009) 92(1):4–14. doi: 10.1016/j.radonc.2009.04.014
4. Ho AS, Kraus DH, Ganly I, Lee NY, Shah JP, Morris LGT. Decision Making in the Management of Recurrent Head and Neck Cancer. *Head & Neck* (2014) 36: (1):144–51. doi: 10.1002/hed.23227
5. Leclerc M, Maingon P, Hamoir M, Dalban C, Calais G, Nuyts S, et al. A dose Escalation Study With Intensity Modulated Radiation Therapy (IMRT) in T2N0, T2N1, T3N0 Squamous Cell Carcinomas (SCC) of the Oropharynx, Larynx nd Hypopharynx Using a Simultaneous Integrated Boost (SIB) Approach. *Radiother Oncol* (2013) 106(3):333–40. doi: 10.1016/j.radonc.2013.03.002
6. Ferreira BC, Sá-Couto P, Khouri L, Lopes M. Biological Dose-escalated Definitive Radiation Therapy in Head and Neck Cancer. *British J Rad* (2017) 90(1072):20160477. doi: 10.1259/bjr.20160477
7. Lauve A, Morris M, Schmidt-Ullrich R, Wu Q, Mohan R, Abayomi O, et al. Simultaneous Integrated Boost Intensity-Modulated Radiotherapy For Locally Advanced Head-and-Neck Squamous Cell Carcinomas: II—Clinical Results. *Int J Radiat OncologyBiologyPhysics* (2004) 60(2):374–87. doi: 10.1016/j.ijrobp.2004.03.010
8. Bourhis J, Overgaard J, Audry H, Ang KK, Saunders M, Bernier J, et al. Hyperfractionated or Accelerated Radiotherapy in Head And Neck Cancer: A Meta-Analysis. *Lancet (London England)* (2006) 368(9538):843–54. doi: 10.1016/s0140-6736(06)69121-6
9. Sher DJ, Adelstein DJ, Bajaj GK, Brizel DM, Cohen EEW, Halthore A, et al. Radiation Therapy for Oropharyngeal Squamous Cell Carcinoma: Executive Summary of an ASTRO Evidence-Based Clinical Practice Guideline. *Pract Radiat Oncol* (2017) 7(4):246–53. doi: 10.1016/j.prro.2017.02.002
10. Mroz EA, Tward AD, Pickering CR, Myers JN, Ferris RL, Rocco JW. High Intratumor Genetic Heterogeneity Is Related to Worse Outcome in Patients With Head and Neck Squamous Cell Carcinoma. *Cancer* (2013) 119 (16):3034–42. doi: 10.1002/cncr.28150

11. Thorwarth D. Radiotherapy Treatment Planning Based on Functional PET/CT Imaging Data. *Nuclear Med Rev* (2012) 15(C):43–7. doi: 10.5603/NMR.2012.0003
12. Skjøtskift T, Evensen ME, Furre T, Moan JM, Amdal CD, Bogsrud TV, et al. Dose Painting For Re-Irradiation of Head And Neck Cancer. *Acta Oncologica* (2018) 57(12):1693–9. doi: 10.1080/0284186X.2018.1512753
13. Welz S, Monnich D, Pfannenberg C, Nikolaou K, Reimold M, La Fougere C, et al. Prognostic Value Of Dynamic Hypoxia PET in Head and Neck Cancer: Results From a Planned Interim Analysis of a Randomized Phase II Hypoxia-Image Guided Dose Escalation Trial. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2017) 124(3):526–32. doi: 10.1016/j.radonc.2017.04.004
14. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: Extracting More Information from Medical Images Using Advanced Feature Analysis. *Eur J Cancer (Oxford Engl 1990)* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
15. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding Tumour Phenotype By Noninvasive Imaging Using A Quantitative Radiomics Approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5006
16. Leijenaar RT, Bogowicz M, Jochems A, Hoebers FJ, Wesseling FW, Huang SH, et al. Development and Validation of a Radiomic Signature to Predict HPV (P16) Status From Standard CT Imaging: A Multicenter Study. *British J Rad* (2018) 91: (1086):20170498. doi: 10.1259/bjr.20170498
17. Giraud P, Giraud P, Gasnier A, El Ayachy R, Kreps S, Foy J-P, et al. Radiomics and Machine Learning for Radiotherapy in Head and Neck Cancers. *Front Oncol* (2019) 9(174). doi: 10.3389/fonc.2019.00174
18. Bogowicz M, Vuong D, Huellner MW, Pavic M, Andratschke N, Gabrys HS, et al. CT Radiomics and PET Radiomics: Ready for Clinical Implementation? *Q J Nucl Med Mol* (2019) 63(4):355–70. doi: 10.23736/s1824-4785.19.03192-3
19. Tanadini-Lang S, Balermpas P, Guckenberger M, Pavic M, Riesterer O, Vuong D, et al. Radiomic Biomarkers For Head And Neck Squamous Cell Carcinoma. *Strahlenther Onkol* (2020) 196(10):868–78. doi: 10.1007/s00066-020-01638-4
20. Bogowicz M, Riesterer O, Stark LS, Studer G, Unkelbach J, Guckenberger M, et al. Comparison of PET and CT Radiomics For Prediction of Local Tumor Control in Head and Neck Squamous Cell Carcinoma. *Acta Oncologica* (2017) 56(11):1531–6. doi: 10.1080/0284186X.2017.1346382
21. Anderson MD. Cancer Center Neck Quantitative Imaging Working Group. Investigation of Radiomic Signatures for Local Recurrence Using Primary Tumor Texture Analysis in Oropharyngeal Head And Neck Cancer Patients. *Sci Rep* (2018) 8(1):1524. doi: 10.1038/s41598-017-14687-0
22. Vallières M, Kay-Rivest E, Perrin LJ, Liem X, Furstoss C, Aerts HJ, et al. Radiomics Strategies For Risk Assessment of Tumour Failure in Head-and-Neck Cancer. *Sci Rep* (2017) 7(1):10117. doi: 10.1038/s41598-017-10371-5
23. Leger S, Zwanenburg A, Pilz K, Zschaeck S, Zöphel K, Kotzerke J, et al. CT Imaging During Treatment Improves Radiomic Models For Patients With Locally Advanced Head And Neck Cancer. *Radiother Onco* (2019) 130:10–7. doi: 10.1016/j.radonc.2018.07.020
24. Bogowicz M, Tanadini-Lang S, Guckenberger M, Riesterer O. Combined CT Radiomics of Primary Tumor and Metastatic Lymph Nodes Improves Prediction of Loco-Regional Control in Head and Neck Cancer. *Sci Rep* (2019) 9(1):15198. doi: 10.1038/s41598-019-51599-7
25. Beaumont J, Acosta O, Devillers A, Palard-Novello X, Chajon E, de Crevoisier R, et al. Voxel-based Identification of Local Recurrence Sub-Regions From Pre-treatment PET/CT For Locally Advanced Head And Neck Cancers. *EJNMMI Res* (2019) 9(1):90. doi: 10.1186/s13550-019-0556-z
26. Xu H, Lv W, Feng H, Du D, Yuan Q, Wang Q, et al. Subregional Radiomics Analysis of PET/CT Imaging With Intratumor Partitioning: Application to Prognosis for Nasopharyngeal Carcinoma. *Mol Imaging Biol* (2019) 22 (5):1414–26. doi: 10.1007/s11307-019-01439-x
27. McGarry SD, Bukowy JD, Iczkowski KA, Unteriner JG, Duvnjak P, Lowman AK, et al. Gleason Probability Maps: A Radiomics Tool for Mapping Prostate Cancer Likelihood in MRI Space. *Tomography* (2019) 5(1):127. doi: 10.18383/j.tom.2018.00033
28. Riesterer O, Pruschy M, Bender S, Sharma A, Bogowicz M, Tanadini-Lang S, et al. Consolidation Cetuximab After Concurrent Triplet Radiochemotherapy+ Cetuximab in Patients With Advanced Head and Neck Cancer: A Randomized Phase II Study. *Radiother Oncol* (2020) 150:62–9. doi: 10.1016/j.radonc.2020.06.011
29. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJ, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics For High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295:(2):328–38. doi: 10.1148/radiol.2020191145
30. Radiomics team USZ. *(2020) Z-Rad Documentation*. Available at: https://medical-physics-usz.github.io/ (Accessed March 26, 2020).
31. Ou D, Blanchard P, Rosellini S, Levy A, Nguyen F, Leijenaar RT, et al. Predictive and Prognostic Value Of CT Based Radiomics Signature in Locally Advanced Head And Neck Cancers Patients Treated With Concurrent Chemoradiotherapy Or Bioradiotherapy And Its Added Value To Human Papillomavirus Status. *Oral Onco* (2017) 71:150–5. doi: 10.1016/j.oraloncology.2017.06.015
32. Vaupel P. Tumor Microenvironmental Physiology and its Implications For Radiation Oncology. *Semin Radiat Oncol* (2004) 14(3):198–206. doi: 10.1016/j.semradonc.2004.04.008
33. Zschaeck S, Haase R, Abolmaali N, Perrin R, Stützer K, Appold S, et al. Spatial Distribution Of FMISO In Head And Neck Squamous Cell Carcinomas During Radio-Chemotherapy And Its Correlation To Pattern Of Failure. *Acta Oncologica* (2015) 54: (9):1355–63. doi: 10.3109/0284186X.2015.1074720
34. Boeke S, Thorwarth D, Mönnich D, Pfannenberg C, Reischl G, La Fougère C, et al. Geometric Analysis Of Loco-Regional Recurrences in Relation to Pre-Treatment Hypoxia in Patients With Head and Neck Cancer. *Acta Oncologica* (2017) 56: (11):1571–6. doi: 10.1080/0284186X.2017.1372626
35. Due AK, Vogelius IR, Aznar MC, Bentzen SM, Berthelsen AK, Korreman SS, et al. Recurrences After Intensity Modulated Radiotherapy for Head and Neck Squamous Cell Carcinoma More Likely to Originate From Regions With High Baseline [18F]-FDG Uptake. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2014) 111(3):360–5. doi: 10.1016/j.radonc.2014.06.001
36. Soto DE, Kessler ML, Piert M, Eisbruch A. Correlation Between Pretreatment FDG-PET Biological Target Volume and Anatomical Location of Failure After Radiation Therapy for Head and Neck Cancers. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2008) 89(1):13–8. doi: 10.1016/j.radonc.2008.05.021
37. Madani I, Duprez F, Boterberg T, Van de Wiele C, Bonte K, Deron P, et al. Maximum Tolerated Dose in a Phase I Trial on Adaptive Dose Painting by Numbers For Head and Neck Cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2011) 101(3):351–5. doi: 10.1016/j.radonc.2011.06.020
38. Ng SH, Liao CT, Lin CY, Chan SC, Lin YC, Yen TC, et al. Dynamic Contrast-enhanced MRI, Diffusion-Weighted MRI and (18)F-FDG PET/CT for the Prediction of Survival in Oropharyngeal or Hypopharyngeal Squamous Cell Carcinoma Treated With Chemoradiation. *Eur Radiol* (2016) 26(11):4162–72. doi: 10.1007/s00330-016-4276-8
39. Perri F, Pacelli R, Della Vittoria Scarpati G, Cella L, Giuliano M, Caponigro F, et al. Radioresistance in Head And Neck Squamous Cell Carcinoma: Biological Bases And Therapeutic Implications. *Head Neck* (2015) 37 (5):763–70. doi: 10.1002/hed.23837
40. Houweling AC, Wolf AL, Vogel WV, Hamming-Vrieze O, van Vliet-Vroegindeweij C, van de Kamer JB, et al. FDG-PET And Diffusion-Weighted MRI In Head-and-Neck Cancer Patients: Implications For Dose Painting. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2013) 106(2):250–4. doi: 10.1016/j.radonc.2013.01.003
41. Purohit BS, Ailianou A, Dulguerov N, Becker CD, Ratib O, Becker M. FDG-PET/CT Pitfalls In Oncological Head And Neck Imaging. *Insights Into Imaging* (2014) 5(5):585–602. doi: 10.1007/s13244-014-0349-x
42. Bhide SA, Newbold KL, Harrington KJ, Nutting CM. Clinical Evaluation of Intensity-Modulated Radiotherapy for Head and Neck Cancers. *Br J Radiol* (2012) 85(1013):487–94. doi: 10.1259/bjr/85942136
43. Pigorsch SU, Wilkens JJ, Kampfer S, Kehl V, Hapfelmeier A, Schläger C, et al. Do Selective Radiation Dose Escalation And Tumour Hypoxia Status Impact The Loco-Regional Tumour Control After Radio-Chemotherapy Of Head & Neck Tumours? The ESCALOX Protocol. *Radiat Oncol* (2017) 12(1):45. doi: 10.1186/s13014-017-0776-1
44. Servagi-Vernat S, Differding S, Sterpin E, Hanin F-X, Labar D, Bol A, et al. Hypoxia-Guided Adaptive Radiation Dose Escalation In Head And Neck Carcinoma: A Planning Study. *Acta Oncol* (2015) 54(7):1008–16. doi: 10.3109/0284186X.2014.990109
45. Zhu Y, Mohamed ASR, Lai SY, Yang S, Kanwar A, Wei L, et al. Imaging-Genomic Study of Head and Neck Squamous Cell Carcinoma: Associations

Between Radiomic Phenotypes and Genomic Mechanisms Via Integration of The Cancer Genome Atlas and The Cancer Imaging Archive. *JCO Clin Cancer Inf* (2019) 3:1–9. doi: 10.1200/cci.18.00073

46. Colen R, Foster I, Gatenby R, Giger ME, Gillies R, Gutman D, et al. NCI Workshop Report: Clinical and Computational Requirements for Correlating Imaging Phenotypes with Genomics Signatures. *Trans Oncol* (2014) 7(5):556–69. doi: 10.1016/j.tranon.2014.07.007
47. Zwirner K, Hilke FJ, Demidov G, Socarras Fernandez J, Ossowski S, Gani C, et al. Radiogenomics in Head And Neck Cancer: Correlation Of Radiomic Heterogeneity And Somatic Mutations in TP53, FAT1 and KMT2D. *Strahlenther Und Onkol Organ Der Deutschen Rontgengesellschaft [et al]* (2019) 195(9):771–9. doi: 10.1007/s00066-019-01478-x
48. Bogowicz M, Jochems A, Deist TM, Tanadini-Lang S, Huang SH, Chan B, et al. Privacy-preserving Distributed Learning Of Radiomics To Predict Overall Survival And HPV Status In Head And Neck Cancer. *Sci Rep* (2020) 10(1):1–10. doi: 10.1038/s41598-020-61297-4
49. Andersson KM, Dahlgren CV, Reizenstein J, Cao Y, Ahnesjö A, Thunberg PJMP. Evaluation of Two Commercial CT Metal Artifact Reduction Algorithms For Use In Proton Radiotherapy Treatment Planning In The Head And Neck Area. *Medical Physics* (2018) 45(10):4329–44. doi: 10.1002/mp.13115

Conflict of Interest: ST-L reports grants from Swiss National Science Foundation, during the conduct of the study, and grants from Siemens Healthcare, outside the submitted work. MG reports grants from Swiss National Science Foundation, during the conduct of the study, grants from Siemens Healthcare, and grants from Artificial Intelligence in oncological Imaging Network, University of Zurich, outside the submitted work. MH reports grants and personal fees from GE Healthcare, grants from Alfred and Annemarie von Sick legacy, and grants from Artificial Intelligence in oncological Imaging Network, University of Zurich, outside the submitted work.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

MINI REVIEW
published: 28 May 2021
doi: 10.3389/fonc.2021.664688

Current and Future Role of Medical Imaging in Guiding the Management of Patients With Relapsed and Refractory Non-Hodgkin Lymphoma Treated With CAR T-Cell Therapy

Laetitia Vercellino[1], Dorine de Jong[2], Roberta di Blasi[3], Salim Kanoun[4], Ran Reshef[5], Lawrence H. Schwartz[6] and Laurent Dercle[6]*

[1] Nuclear Medicine Department Saint Louis Hospital, Assistance Publique Hôpitaux de Paris, Paris, France, [2] Center for Cell Engineering, Memorial Sloan Kettering Cancer Center, New York, NY, United States, [3] Onco-Hematology Department Saint Louis Hospital, Assistance Publique Hôpitaux de Paris, Paris, France, [4] Cancer Research Center of Toulouse (CRCT), Team 9, INSERM UMR 1037, Toulouse, France, [5] Blood and Marrow Transplantation and Cell Therapy Program, Division of Hematology/Oncology and Columbia Center for Translational Immunology, Columbia University Irving Medical Center, New York City, NY, United States, [6] Department of Radiology, New York Presbyterian, Columbia University Irving Medical Center, New York City, NY, United States

OPEN ACCESS

Edited by:
Pilar López-Larrubia, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Ekta Dhamija, All India Institute of Medical Sciences, India
Inna Serganova, Cornell University, United States

***Correspondence:**
Laetitia Vercellino
laetitia.vercellino@aphp.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *05 February 2021*
Accepted: *05 May 2021*
Published: *28 May 2021*

Citation:
Vercellino L, de Jong D, di Blasi R, Kanoun S, Reshef R, Schwartz LH and Dercle L (2021) Current and Future Role of Medical Imaging in Guiding the Management of Patients With Relapsed and Refractory Non-Hodgkin Lymphoma Treated With CAR T-Cell Therapy. Front. Oncol. 11:664688. doi: 10.3389/fonc.2021.664688

Chimeric antigen receptor (CAR) T-cells are a novel immunotherapy available for patients with refractory/relapsed non-Hodgkin lymphoma. In this indication, clinical trials have demonstrated that CAR T-cells achieve high rates of response, complete response, and long-term response (up to 80%, 60%, and 40%, respectively). Nonetheless, the majority of patients ultimately relapsed. This review provides an overview about the current and future role of medical imaging in guiding the management of non-Hodgkin lymphoma patients treated with CAR T-cells. It discusses the value of predictive and prognostic biomarkers to better stratify the risk of relapse, and provide a patient-tailored therapeutic strategy. At baseline, high tumor volume (assessed on CT-scan or on [18F]-FDG PET/CT) is a prognostic factor associated with treatment failure. Response assessment has not been studied extensively yet. Available data suggests that current response assessment developed on CT-scan or on [18F]-FDG PET/CT for cytotoxic systemic therapies remains relevant to estimate lymphoma response to CAR T-cell therapy. Nonetheless, atypical patterns of response and progression have been observed and should be further analyzed. The potential advantages as well as limitations of artificial intelligence and radiomics as tools providing high throughput quantitative imaging features is described.

Keywords: lymphoma, CAR T-cell, immunotherapy, FDG PET/CT, CT scan, prognostic biomarker

INTRODUCTION

In 2017, reprogramming T lymphocytes to carry chimeric antigen receptor (CAR) targeting CD19 antigen became a novel immunotherapy commercially available for patients with refractory/relapsed B cell malignancies. Despite the unprecedented therapeutic responses achieved by CD19-CAR T-cells, the number of patients experiencing relapse stresses the need for reliable

biomarkers to closely monitor clinical response and implement early consolidation strategies. Medical imaging such as [18F]-FDG PET/CT is already used for diagnosis and evaluation of hematologic malignancies and the clinical significance of several PET parameters such as Total Metabolic Tumor Volume, standardized uptake value and Total Lesion Glycolysis consumption has been extensively demonstrated. Hence, this is an area of ongoing investigation in the context of CAR T-cell therapy. This review aims to summarize recent clinical data and to emphasize the importance of further investigation of medical imaging biomarkers for CAR T-cells to optimize and personalize medical care: risk stratification, prediction of response, response assessment, and early detection of relapse.

CAR T-CELL MANUFACTURING AND TREATMENT

CAR T-cell manufacturing begins with T-cell collection from patients or donors by aphaeresis. These cells are then genetically reprogrammed (e.g., using viral vectors) to express receptors for specific tumor antigens. CD19 CAR T-cell therapy uses a single chain variable fragment (scFv) derived from the variable heavy and variable light chains of an antibody against epitopes of the CD19 antigen. In second generation CARs, the scFv is connected through a transmembrane domain to a costimulatory domain (such as CD28 or 4-1BB) further linked to the CD3ζ intracellular signaling domain of the T-cell receptor (1, 2). Before CAR T-cell infusion, patients are often given a bridging therapy to control their disease during the CAR T-cell production which may take two to six weeks. Patients then receive a lymphodepleting treatment several days before infusion to create a favorable environment for CAR T-cells by removing unmodified T-cells and immunosuppressive regulatory T-cell and increase homeostatic cytokine levels (3). Finally, cells are expanded and infused back into patients to achieve tumor cell recognition and killing. The antitumor response driven by CAR T-cells is HLA independent and relies on antigen-receptor binding and on the co-stimulatory signals that enhance T-cell proliferation and/or persistence. After antigen recognition, CAR T-cells eliminate cancer cells through death receptors-, cytokines- or granzyme/perforin-induced killing.

CURRENT CLINICAL TRIAL LANDSCAPE

As CAR-T-cells are still entering routine clinical practice, most of the current knowledge about imaging and outcomes has been gained from review of clinical trials. The most common cancer subtypes studied are acute lymphoblastic leukemia and non-Hodgkin lymphoma and the majority of studies used an autologous cell source (4). While the treatment of relapsed/refractory leukemia and lymphoma is increasing in clinical practice, and the first CAR T product for multiple myeloma was just recently FDA-approved (5, 6), the applications of this therapy in solid cancers remain at a nascent stage and need further investigations (7). Of note, medical imaging guides the management of lymphoma patients while it has a more limited impact for leukemia. Therefore, the focus of this review is on the contribution of medical imaging in lymphoma patients treated with CAR T-cells.

RESPONSE RATES

CD19 CAR T-cell therapy has a high overall response rate with complete responses in up to 90% in adult and pediatric acute lymphoblastic leukemia motivating the initiation of hundreds of CAR T-cell clinical trials worldwide and the search for more efficient designs and new antigens (8, 9).

In clinical trials of aggressive lymphoma patients, complete response rates ranged from 40% to 59% (10). In indolent lymphoma, complete response rates were even higher (11, 12). Several factors, however, can limit long-term efficacy of CAR T-cells and lead to disease relapse (13). First, the delay to start CAR T-cell therapy may allow disease progression, with tumor volume increase, which stresses the importance of rapid and reliable manufacturing and bridging therapy. CAR T-cell function may also be decreased by the poor quality and low number of cells obtained from certain patients. In addition, treatment efficacy depends on the expression of the targeted antigen by the tumor cells. Therefore, tumor heterogeneity, mutations, down-regulation or loss of tumor antigen can decrease the recognition of tumors by CAR T-cells and subsequent therapeutic response. Moreover, CAR T-cells' viability and efficacy can be impaired by suboptimal stimulation leading to T cells exhaustion and relapse (14).

In addition to poor effector to target ratio in presence of a high tumor burden (15), antigen positive relapse can also occur as a result of tumor cells resistance to CAR T-cells (14) or immunosuppressive tumor microenvironment (16) inducing CAR T-cell dysfunction.

BASELINE BIOMARKERS PREDICTING RESPONSE AND OUTCOME

Biomarkers that predict short survival are critical for close monitoring during bridging therapy given that the manufacturing time may allow disease progression. Biomarkers that predict durable response at baseline may help identify patients more likely to benefit from this strategy. Biomarkers that indicate treatment failure at the 1-month milestone will identify patients who might benefit from early intervention such as therapies that reinvigorate CAR T-cells or a second infusion of CAR T-cells.

Prognostic Value of Response Under Bridging Chemotherapy

In 72 patients with relapsed/refractory diffuse large B-cell lymphoma who received CAR T-cells, Tordo et al. measured the kinetics of tumor bulk during bridging therapy determined

by the evolution of Total Metabolic Tumor Volume, Total Lesion Glycolysis or SUVmax. They demonstrated that biomarkers derived from the analysis of the kinetics of theses parameters during bridging therapy before lymphodepletion are better predictors of progression-free survival than baseline biomarkers (17). Thus, patients with satisfactory disease control before lymphodepletion had an overall longer progression-free survival.

Tumor Volume on CT-Scan

On CT-scan, tumor volume is typically estimated using the sum of the product of perpendicular diameters of measurable target tumor lesions. In the first series reporting the efficacy of CAR T-cell in lymphoma, there was only a non-significant trend for the predictive and prognostic impact of tumor bulk assessed by morphological imaging. Schuster et al. found that the median sum of the product of perpendicular diameters was 20 cm2 (range 3-100) in responding patients and 30 cm2 (range 3-157) in non-responding patients (18). A recently published report on the TRANSCEND NHL01 trial also found a trend towards worse outcomes in patients with greater sum of the product of perpendicular diameters, with an objective response rate of 76.8% in patients with the sum of the product of perpendicular diameters <50 cm2 and 61.4% in patients with ≥50 cm2 (19). Likewise, Neelapu et al. found that patients with bulky disease (>10 cm) had an objective response rate of 71% (95CI: 0.44-0.90) compared with 85% (95CI: 0.75-0.91) for patients without bulky disease (20). More recently, an analysis in the same cohort found an association between tumor burden evaluated by the sum of the products of diameters of target lesions and durable response, however this parameter had limited sensitivity and specificity, which may be due to the fact that it does not take into account total tumor burden (15). Expansion of CAR T-cells in the blood was also predictive of response, along with markers of inflammation such as IL-6 and CRP. Interestingly, in this study the best predictor of durable response was the peak CAR T-cell levels in the blood normalized to pretreatment tumor burden. Durable responders had a higher peak CAR T-cell to tumor burden ratio than non-responders or responding patients relapsing within one year.

Tumor Volume on [18F]-FDG PET-Scan

[18F]-FDG PET is a routine, standard of care imaging study that estimates tumor glucose consumption. [18F]-FDG PET is preferred for the staging and restaging of FDG-avid lymphomas because it outperforms CT scans in these diseases (21). Additionally, the overall metabolic tumor volume or Total Metabolic Tumor Volume (i.e, metabolically active tumor volume with significantly increased glucose metabolism) assessed on [18F]-FDG PET has good prognostic value. For instance, higher tumor volume in aggressive lymphomas before initiating first-line chemotherapy predicts shorter progression-free and overall survival. Therefore, tumor volume along with other parameters such as International Prognostic Index, Eastern Cooperative Oncology Group performance status and cell of origin could improve patient risk stratification (22, 23). Several groups have explored the predictive/prognostic value of metabolic tumor volume in patients undergoing CAR T-cell therapy and preliminary results suggest that Total Metabolic Tumor Volume is a relevant imaging biomarker.

In a study done on a small cohort (n=19) of patients with non-Hodgkin lymphomas, with a best overall response rate of 79%, the median Total Metabolic Tumor Volume was 72 cm^3. Lower tumor volume was observed in responding (58.1 cm^3) than in non-responding patients (110.8 cm^3), though this did not reach statistical significance (24). Likewise, overall survival was not significantly different in patients above and below the median (8.6 months vs 11.5 months). The absence of prognostic value might be due to the small size of this cohort. Nonetheless, the authors found that patients with more severe cytokine release syndrome (grade 3-4), had significantly higher Total Metabolic Tumor Volume than patients with no or mild cytokine release syndrome.

Dean et al. observed a stronger correlation between Total Metabolic Tumor Volume and outcome in a larger cohort of 96 patients with diffuse large B-cell lymphoma (25). In a sub-group of 48 patients the median Total Metabolic Tumor Volume (determined by a manual method) was 147 mL. Lower tumor volume was associated with prolonged overall and progression-free survival. This was validated in a second sub-group (n=48), where median Total Metabolic Tumor Volume was lower (72.8 mL), and in the entire study population. Lower tumor volume was also associated with higher overall and complete response rates. In a subgroup of 72 patients with "true baseline" PET (no bridging therapy, or PET performed after bridging chemotherapy) the same results were observed. In addition, high tumor volume was associated with more grade 3-4 cytokine release syndrome but not with neurotoxicity.

In another study (n=116 patients with aggressive non-Hodgkin lymphoma), extension of lymphoma measured by more than two involved extranodal sites both at times of enrollment (decision of CAR T-cells and before bridging therapy if applicable) and treatment, and high Total Metabolic Tumor Volume (superior to 80 mL) at the time of treatment were predictive of progression-free survival, overall survival, and early progression (occurring during the first month) after CAR T-cell treatment in patients with R/R diffuse large B-cell lymphoma. Of note, elevated CRP at time of CAR T-cell infusion was also associated with a worse outcome (but with a low odds ratio). Combining the number of extranodal sites>2 and high tumor volume (>80 mL) allowed to establish 3 prognostic groups with 0, 1 or 2 adverse parameters, more distinctly than the revised International Prognostic Index (26).

ON-TREATMENT IMAGING BIOMARKERS: MEASURING RESPONSE

Learning Curve

Since CAR-T-cell therapy is relatively novel, there is a critical need to evaluate the reproducibility in assessing response since

even expert radiologists will have to familiarize with the patterns of response to CAR T-cells. The optimal time point for follow-up and therapeutic evaluation is not known and on initial studies was determined empirically. Second, pseudoprogression may occur, as observed in other immunotherapies, but there are no definitive criteria to define it. Finally, immune response could generate atypical uptake linked to an inflammatory process, such as observed with immunotherapy (colitis, thyroiditis…) (27–29).

Typical Patterns of Response and Progression on CT-Scan

In 101 patients with relapsed/refractory aggressive B-cell non-Hodgkin lymphoma enrolled in the ZUMA-1 study, axicabtagene ciloleucel (axi-cel) had an overall response rate of 83% and complete response rate of 58%. The median duration of response was 11 months. Ongoing long-term responses were seen in 39% of patients after a median follow up of 27.1 months (30). In a recent update, the three-year overall survival was 47% (31). Eleven out of 33 patients with partial responses at 1 month, and 11 of 24 patients with stable disease at 1 month, subsequently attained a complete response without any additional therapy (30). In these cases of responses improving over time, complete resolution of FDG-avid lesions after CAR T-cell therapy may take up to 9-12 months and anecdotally even longer. Additionally, a complete response and partial response at the 3-month milestone were associated with similar progression-free survival, further highlighting the complexity in using early imaging findings for prognostic purposes.

Real-world evidence using commercial CAR T-cells, including patients with comorbidities found similar results with overall response rate of 82% and complete response rate of 64%, with estimated 12-months PFS of 47% and OS of 68% (32). Among the patients with a partial response at 1 month 32% achieved complete response at 3 months and only 1 out of 14 patients with stable disease achieved a complete response at 3 months.

Typical Patterns of Response and Progression on [18F]-FDG PET

Shah et al. reported a case series of 7 patients with aggressive and indolent non-Hodgkin lymphoma treated with CAR T-cells, evaluated with [18F]-FDG PET (33). Three of these patients (all with follicular lymphoma) had a complete metabolic response at 1 month and remained disease-free at 2 years. Two patients with a partial metabolic response experienced later progression, at 3- and 6-months post-infusion respectively, and two patients had progressive disease as soon as 1 month post infusion. At this time-point, no Cytokine Release Syndrome-related metabolic activity impaired FDG PET interpretation.

A recent retrospective report on 10 patients with aggressive lymphoma, underlined the importance of early evaluation of therapeutic efficacy, with all patients with a partial or metabolic response at 3 months having shown metabolic response at 1 month, and only one patient with unfavorable outcome experiencing early metabolic response (34).

Examples of therapeutic response assessment with FDG PET are presented in **Figures 1** and **2**.

Atypical Patterns of Response and Progression

Wang et al. reported 3 cases of pseudoprogression that may cause local complications, due to compression of adjacent organs for example. Compared with other immunotherapies, pseudoprogression was very early, occurring as soon as 4 to 5 days after CAR T-cell infusion (24). More studies will be needed to better understand and describe the different patterns of response after CAR T-cell therapy.

TOXICITY

CAR T-cell is associated with a wide range of toxicities. Across studies, all patients experience at least one adverse event. On the most serious side of the spectrum are high grades Cytokine Release Syndrome (CRS) and Immune effector cell-associated Neurotoxicities Syndrome (ICANS). CRS is a systemic inflammatory response characterized among others by fever, hypotension, hypoxia and potential multiple organ failure whereas ICANS is characterized by various neurologic symptoms ranging from confusion and tremor to aphasia, dysgraphia, seizures or coma. These generally occur in the first weeks after CAR T-cell infusion as a results of high levels of cytokines not only produced by CAR T-cells but also by their activation of myeloid cells (3, 35). Pyrexia, fatigue, cytopenias and infections are also frequent after CAR T-cell therapy.

In the ZUMA-1 study, axi-cel infusion was associated with 11% incidence of grade 3 or higher CRS and 32% of ICANS grade 3 or more, with similar results in the real-world data described by Nastoupil et al. (30, 32, 36). In the JULIET study, CRS occurred in 58% patients, with grade 3 or higher CRS in 22%, while neurologic events were observed in 21% of patients, and grade 3 or higher in 12% (37). As mentioned above, several analyses showed that higher baseline tumor burden was associated with severe CRS (24, 25).

For neurotoxicity, along with abnormalities on electroencephalogram and transcranial Doppler ultrasound, dedicated brain FDG PET can contribute to diagnosis (the main observed abnormalities being cortical hypometabolism) and follow-up (38). Structural abnormalities are usually absent, but CT and MRI may identify concomitant events such as ischemic stroke or subarachnoid hemorrhage (38). However, there is a lack of prospectively collected data on the subject. Additionally, a recent study suggested that a higher FDG avidity of lymphoma, evaluated by SUVmax, was associated with more neurotoxicity (34). If confirmed in larger studies, this could also be of importance for patient management after CAR T-cell infusion. The mainstay of treatment of these toxicities is steroids, vasopressors and tocilizumab treatments.

CHALLENGES AND PERSPECTIVES

There are several challenges to improve implementation of CAR-T-cells in routine clinical practice and the outcome of patients

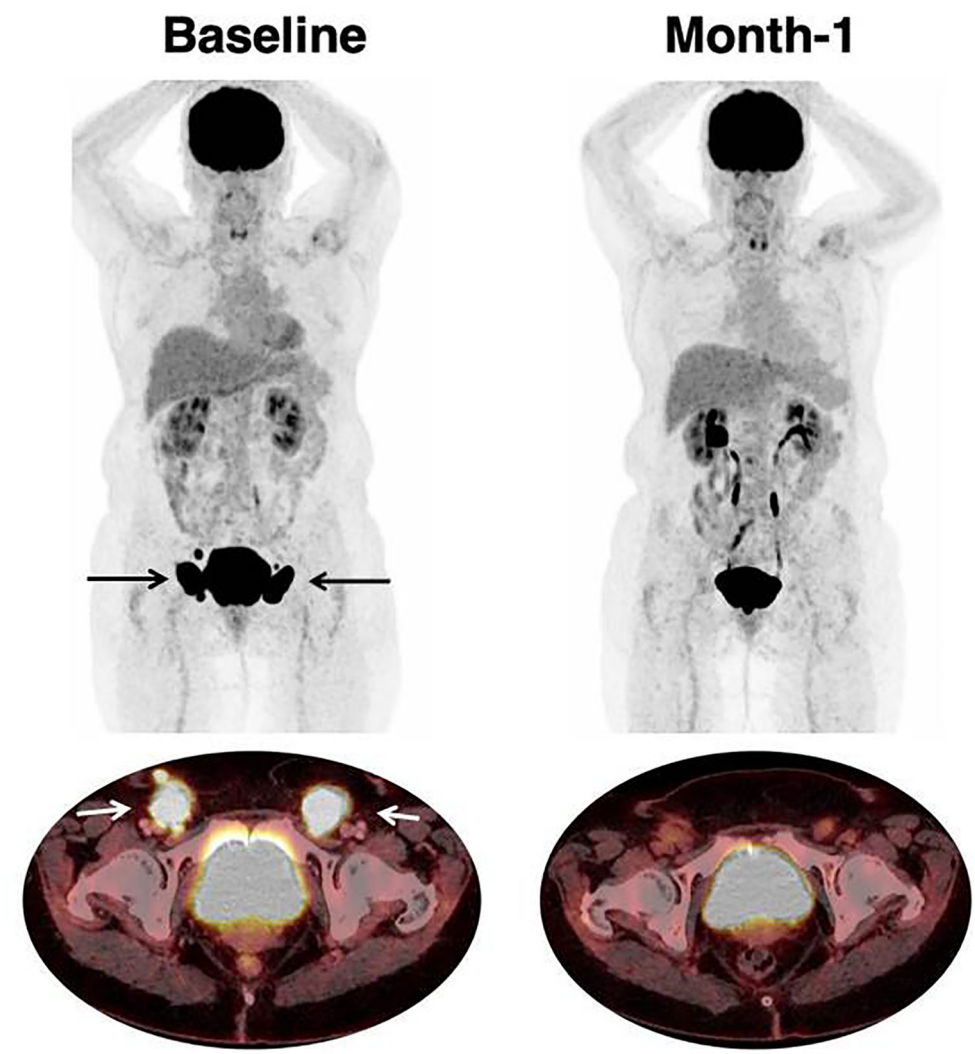


FIGURE 1 | Response to CAR T-cell therapy. 66 year-old patient with past medical history of follicular lymphoma. The patient relapsed with DLBCL, treated with two lines of prior chemotherapy. Baseline imaging showed a low tumor volume (TMTV was 47 mm^3) which is typically associated with favorable outcome and response to CAR T-cell therapy. Inguinal lymphadenopathies are indicated with black (on Maximum Intensity Projection) and white arrows (on axial fusion image). Follow-up imaging showed a partial response on CT-scan with residual disease. [18F]-FDG PET reclassified this patient as a complete metabolic response which persisted at month-6.

receiving this innovative treatment. These challenges include technical challenges (CAR-T-cell development, manufacturing), standardization of clinical trial results to facilitate the comparison (protocols, pre-conditioning of patients, CAR-T-cell formulation, quality and persistence), and identifying robust tools to optimize treatment decision (4, 7). Among these tools, imaging techniques may play a critical role. The role and use of medical imaging techniques remain to be defined but results presented above suggest that imaging will be a pivotal tool to guide treatment decisions.

Beyond Total Metabolic Tumor Volume, recent data suggest that lesion dissemination assessed on PET/CT by means of the largest distance between two lesions (normalized with the body surface area), contributes to assess the spread of the disease, and has a prognostic value, independently of Total Metabolic Tumor Volume in a cohort of first-line chemotherapy diffuse large B-cell lymphoma (39). Radiomics could further contribute to extract clinically meaningful data from medical images. Recent findings from a wide range of solid tumor types suggest that a signature combining a limited subset of pretreatment (40) or on-treatment (41, 42) imaging biomarkers are able to help to identify patients who might benefit from early intervention. Among these quantitative imaging biomarkers, several have been shown to predict responses to immunotherapies with immune checkpoint blockers such as increased tumor volume, increased tumor glucose metabolism, tumor organotropism in visceral tissues, and lower skeletal muscle index. These are all associated with unfavorable outcomes (43–47).

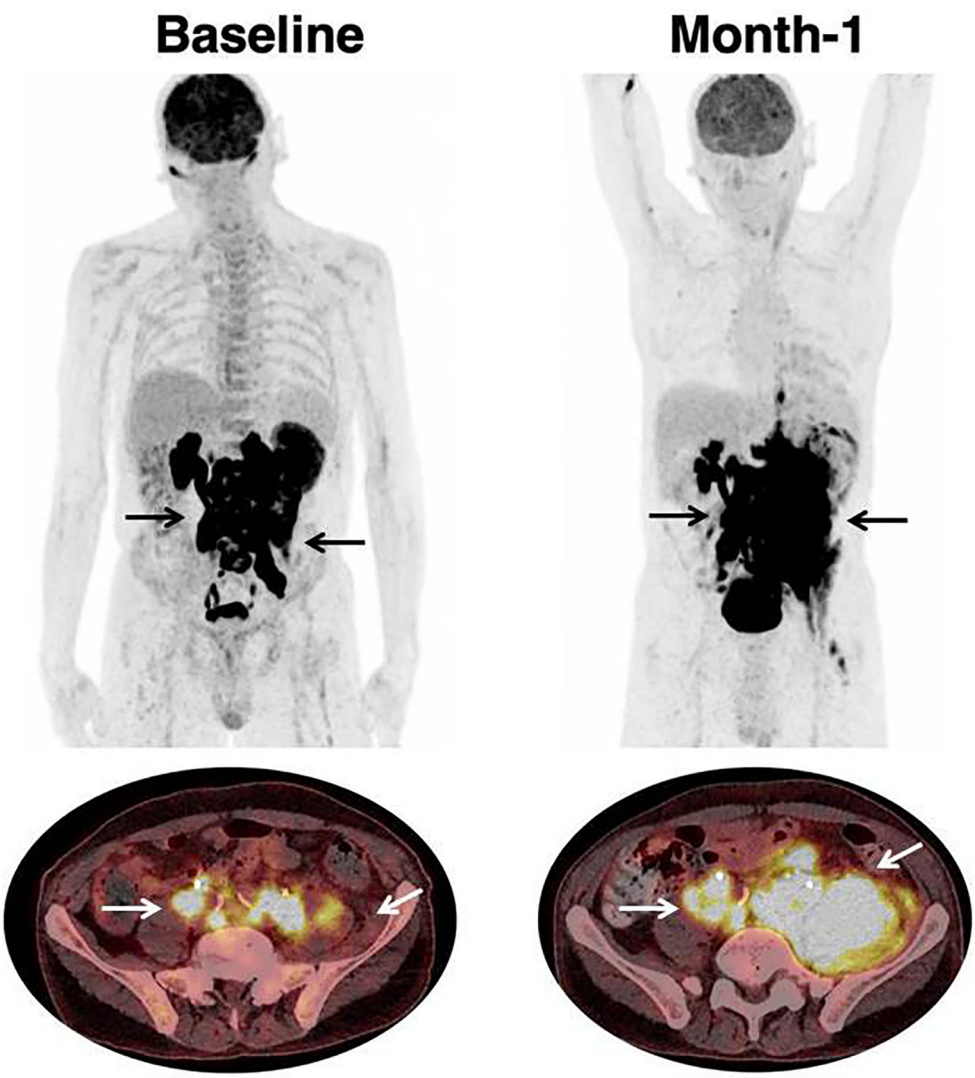


FIGURE 2 | Progression in a patient treated with CAR T-cell therapy. 68 year-old patient with past medical history of DLBCL diagnosed one year prior to treatment initiation. Patient had Stage IV disease, with rearrangement of the MYC and BCL6 genes, and treated with two prior lines of chemotherapy. Black (on Maximum Intensity Projection) and white arrows (on axial fusion images) show infradiaphragmatic lymphadenopathies, with muscular infiltration. Baseline imaging showed high tumor volume which is typically associated with unfavorable outcome and lower response rate to CAR T-cell therapy. Follow-up imaging showed a progression on CT-scan as well as on [18F]-FDG PET. At month-1, there were new lesions as well as an increase in tumor volume. The prognosis was poor; hence salvage treatment and later best supportive care were initiated. Patient died at month-2.

Hence, it is very likely that similar technologies will be applicable to CAR T-cells and that computational models will be applied to data from CT or PET/CT scans to predict outcome, while accounting for technical variability between machines and centers. These tools eliminate the bias of investigator assessment and multicenter variability, allowing their implementation in large multisite trials. Artificial intelligence could be used to combine previously cited biomarkers to build robust prognostic/predictive models. One challenge is that building robust models using artificial intelligence requires creating large datasets, hence the need to aggregate data from multiple institutions to avoid overfitting (48, 49). Eventually, deep learning could contribute to determine radiomics signature correlated with survival.

CONCLUSION

The role of medical imaging, and PET/CT in particular, in lymphoma patients treated with CAR T-cells is twofold. First, the pre-infusion Total Metabolic Tumor Volume seems promising for its prognostic value, and should probably be associated with biological parameters, such as CRP at time of lymphodepletion (26, 50). Some data also suggest that high

tumor volume could be correlated with more severe cytokine release syndrome, with possible direct impact on patient management and monitoring. Second, the evaluation of response with CAR T-cell is an ongoing challenge, with more data needed, especially on the possibility of pseudoprogression, slow or late responses as well as the timing of relapses. Beyond FDG, a better knowledge and understanding of imaging data could contribute to detect and treat toxicities timely (51) and further tailor the therapeutic strategy, with the use of next-generation CAR T-cells, combination therapeutics especially in patients with high tumor burden and potentially rapid implementation of salvage therapies in case of relapse such as new CAR T-cell infusion (targeting the same or other antigens), immunomodulatory agents or radiation therapy (52).

AUTHOR CONTRIBUTIONS

LV, DJ, and LD designed the mini review, drafted the work, and approved the content. RB, SK, RR, and LS revised it critically for important intellectual content and approved the content. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Sadelain M. CAR Therapy: The CD19 Paradigm. *J Clin Invest* (2015) 125:3392–400. doi: 10.1172/JCI80010
2. Sadelain M. Cd19 Car T Cells. *Cell* (2017) 171:1471. doi: 10.1016/j.cell.2017.12.002
3. Yakoub-Agha I, Chabannon C, Bader P, Basak GW, Bonig H, Ciceri F, et al. Management of Adults and Children Undergoing Chimeric Antigen Receptor T-Cell Therapy: Best Practice Recommendations of the European Society for Blood and Marrow Transplantation (EBMT) and the Joint Accreditation Committee of ISCT and EBMT (Jacie). *Haematologica* (2020) 105:297–316. doi: 10.3324/haematol.2019.229781
4. Pettitt D, Arshad Z, Smith J, Stanic T, Holländer G, Brindley D. Car-T Cells: A Systematic Review and Mixed Methods Analysis of the Clinical Trial Landscape. *Mol Ther* (2018) 26:342–53. doi: 10.1016/j.ymthe.2017.10.019
5. Xiang X, He Q, Ou Y, Wang W, Wu Y. Efficacy and Safety of CAR-Modified T Cell Therapy in Patients With Relapsed or Refractory Multiple Myeloma: A Meta-Analysis of Prospective Clinical Trials. *Front Pharmacol* (2020) 11:544754. doi: 10.3389/fphar.2020.544754
6. Mullard A. FDA Approves First BCMA-Targeted CAR-T Cell Therapy. *Nat Rev Drug Discovery* (2021) 20:332. doi: 10.1038/d41573-021-00063-1
7. Holzinger A, Barden M, Abken H. The Growing World of CAR T Cell Trials: A Systematic Review. *Cancer Immunol Immunother* (2016) 65:1433–50. doi: 10.1007/s00262-016-1895-5
8. Hartmann J, Schüßler-Lenz M, Bondanza A, Buchholz CJ. Clinical Development of CAR T Cells-Challenges and Opportunities in Translating Innovative Treatment Concepts. *EMBO Mol Med* (2017) 9:1183–97. doi: 10.15252/emmm.201607485
9. Liu B, Song Y, Liu D. Clinical Trials of CAR-T Cells in China. *J Hematol Oncol* (2017) 10:166. doi: 10.1186/s13045-017-0535-7
10. Chavez JC, Bachmeier C, Kharfan-Dabaja MA. Car T-Cell Therapy for B-Cell Lymphomas: Clinical Trial Results of Available Products. *Ther Adv Hematol* (2019) 10:2040620719841581. doi: 10.1177/2040620719841581
11. Jacobson C. *Primary Analysis of Zuma-5: A Phase 2 Study of Axicabtagene Ciloleucel (Axi-Cel) in Patients With Relapsed/Refractory (R/R) Indolent non-Hodgkin Lymphoma (Inhl)*. ASH. Available at: https://ash.confex.com/ash/2020/webprogram/Paper136834.html (Accessed January 6, 2021).
12. Fowler N. *Efficacy and Safety of Tisagenlecleucel in Adult Patients With Relapsed/Refractory Follicular Lymphoma: Interim Analysis of the Phase 2 Elara Trial*. ASH. Available at: https://ash.confex.com/ash/2020/webprogram/Paper138983.html (Accessed January 6, 2021).
13. Namuduri M, Brentjens RJ. Enhancing CAR T Cell Efficacy: The Next Step Toward a Clinical Revolution? *Expert Rev Hematol* (2020) 13:533–43. doi: 10.1080/17474086.2020.1753501
14. Singh N, Lee YG, Shestova O, Ravikumar P, Hayer KE, Hong SJ, et al. Impaired Death Receptor Signaling in Leukemia Causes Antigen-Independent Resistance by Inducing Car T-Cell Dysfunction. *Cancer Discovery* (2020) 10:552–67. doi: 10.1158/2159-8290.CD-19-0813
15. Locke FL, Rossi JM, Neelapu SS, Jacobson CA, Miklos DB, Ghobadi A, et al. Tumor Burden, Inflammation, and Product Attributes Determine Outcomes of Axicabtagene Ciloleucel in Large B-Cell Lymphoma. *Blood Adv* (2020) 4:4898–911. doi: 10.1182/bloodadvances.2020002394
16. Epperly R, Gottschalk S, Velasquez MP. A Bump in the Road: How the Hostile Aml Microenvironment Affects Car T Cell Therapy. *Front Oncol* (2020) 10:262. doi: 10.3389/fonc.2020.00262
17. Tordo J, Sesques P, Ferrant E, Safar V, Wallet F, Karlin L, et al. *Prognostic Impact of 18F-FDG PET-CT in Patients With Aggressive B Cell Lymphoma Treated With anti-CD19 Car T Cells* EHA (2020).
18. Schuster SJ, Svoboda J, Chong EA, Nasta SD, Mato AR, Anak Ö, et al. Chimeric Antigen Receptor T Cells in Refractory B-Cell Lymphomas. *N Engl J Med* (2017) 377:2545–54. doi: 10.1056/NEJMoa1708566
19. Abramson JS, Palomba ML, Gordon LI, Lunning MA, Wang M, Arnason J, et al. Lisocabtagene Maraleucel for Patients With Relapsed or Refractory Large B-Cell Lymphomas (TRANSCEND NHL 001): A Multicentre Seamless Design Study. *Lancet* (2020) 396:839–52. doi: 10.1016/S0140-6736(20)31366-0
20. Neelapu SS, Locke FL, Bartlett NL, Lekakis LJ, Miklos DB, Jacobson CA, et al. Axicabtagene Ciloleucel Car T-Cell Therapy in Refractory Large B-Cell Lymphoma. *N Engl J Med* (2017) 377:2531–44. doi: 10.1056/NEJMoa1707447
21. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Mueller SP, et al. Role of Imaging in the Staging and Response Assessment of Lymphoma: Consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol* (2014) 32:3048–58. doi: 10.1200/JCO.2013.53.5229
22. Cottereau A-S, Lanic H, Mareschal S, Meignan M, Vera P, Tilly H, et al. Molecular Profile and FDG-PET/CT Total Metabolic Tumor Volume Improve Risk Classification At Diagnosis for Patients With Diffuse Large B-Cell Lymphoma. *Clin Cancer Res* (2016) 22:3801–9. doi: 10.1158/1078-0432.CCR-15-2825
23. Toledano MN, Desbordes P, Banjar A, Gardin I, Vera P, Ruminy P, et al. Combination of Baseline FDG PET/CT Total Metabolic Tumour Volume and Gene Expression Profile Have a Robust Predictive Value in Patients With Diffuse Large B-Cell Lymphoma. *Eur J Nucl Med Mol Imaging* (2018) 45:680–8. doi: 10.1007/s00259-017-3907-x
24. Wang J, Galal A. Can Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Predict Chimeric Antigen Receptor T Cell Adverse Effects? *Biol Blood Marrow Transplant* (2019) 25:e187–8. doi: 10.1016/j.bbmt.2019.05.003
25. Dean EA, Mhaskar RS, Lu H, Mousa MS, Krivenko GS, Lazaryan A, et al. High Metabolic Tumor Volume Is Associated With Decreased Efficacy of Axicabtagene Ciloleucel in Large B-Cell Lymphoma. *Blood Adv* (2020) 4:3268–76. doi: 10.1182/bloodadvances.2020001900
26. Vercellino L, Di Blasi R, Kanoun S, Tessoulin B, Rossi C, D'Aveni-Piney M, et al. Predictive Factors of Early Progression After CAR T-Cell Therapy in Relapsed/Refractory Diffuse Large B-Cell Lymphoma. *Blood Adv* (2020) 4:5607–15. doi: 10.1182/bloodadvances.2020003001
27. Mekki A, Dercle L, Lichtenstein P, Nasser G, Marabelle A, Champiat S, et al. Machine Learning Defined Diagnostic Criteria for Differentiating Pituitary Metastasis From Autoimmune Hypophysitis in Patients Undergoing Immune Checkpoint Blockade Therapy. *Eur J Cancer* (2019) 119:44–56. doi: 10.1016/j.ejca.2019.06.020

28. Mekki A, Dercle L, Lichtenstein P, Marabelle A, Michot J-M, Lambotte O, et al. Detection of Immune-Related Adverse Events by Medical Imaging in Patients Treated With Anti-Programmed Cell Death 1. *Eur J Cancer* (2018) 96:91–104. doi: 10.1016/j.ejca.2018.03.006
29. Dercle L, Mokrane F-Z, Schiano de Colella JM, Stamatoullas A, Morschhauser F, Brice P, et al. Unconventional Immune-Related Phenomena Observed Using 18F-FDG PET/CT in Hodgkin Lymphoma Treated With Anti PD-1 Monoclonal Antibodies. *Eur J Nucl Med Mol Imaging* (2019) 46:1391–2. doi: 10.1007/s00259-019-04310-x
30. Locke FL, Ghobadi A, Jacobson CA, Miklos DB, Lekakis LJ, Oluwole OO, et al. Long-Term Safety and Activity of Axicabtagene Ciloleucel in Refractory Large B-Cell Lymphoma (ZUMA-1): A Single-Arm, Multicentre, Phase 1-2 Trial. *Lancet Oncol* (2019) 20:31–42. doi: 10.1016/S1470-2045(18)30864-7
31. Jacobson C. *Long-Term Survival and Gradual Recovery of B Cells in Patients With Refractory Large B Cell Lymphoma Treated With Axicabtagene Ciloleucel (Axi-Cel)*. ASH. Available at: https://ash.confex.com/ash/2020/webprogram/Paper134362.html (Accessed January 6, 2021).
32. Nastoupil LJ, Jain MD, Feng L, Spiegel JY, Ghobadi A, Lin Y, et al. Standard-of-Care Axicabtagene Ciloleucel for Relapsed or Refractory Large B-Cell Lymphoma: Results From the US Lymphoma Car T Consortium. *J Clin Oncol* (2020) 38:3119–28. doi: 10.1200/JCO.19.02104
33. Shah NN, Nagle SJ, Torigian DA, Farwell MD, Hwang W-T, Frey N, et al. Early Positron Emission Tomography/Computed Tomography as a Predictor of Response After CTL019 Chimeric Antigen Receptor -T-Cell Therapy in B-Cell Non-Hodgkin Lymphomas. *Cytotherapy* (2018) 20:1415–8. doi: 10.1016/j.jcyt.2018.10.003
34. Derlin T, Schultze-Florey C, Werner RA, Möhn N, Skripuletz T, David S, et al. 18f-Fdg PET/CT of Off-Target Lymphoid Organs in CD19-targeting Chimeric Antigen Receptor T-Cell Therapy for Relapsed or Refractory Diffuse Large B-Cell Lymphoma. *Ann Nucl Med* (2021) 35:132–8. doi: 10.1007/s12149-020-01544-w
35. Giavridis T, van der Stegen SJC, Eyquem J, Hamieh M, Piersigilli A, Sadelain M. Car T Cell-Induced Cytokine Release Syndrome is Mediated by Macrophages and Abated by IL-1 Blockade. *Nat Med* (2018) 24:731–8. doi: 10.1038/s41591-018-0041-7
36. Jacobson CA, Hunter B, Armand P, Kamihara Y, Ritz J, Rodig SJ, et al. Axicabtagene Ciloleucel in the Real World: Outcomes and Predictors of Response, Resistance and Toxicity. *Blood* (2018) 132:92–2. doi: 10.1182/blood-2018-99-117199
37. Schuster SJ, Bishop MR, Tam CS, Waller EK, Borchmann P, McGuirk JP, et al. Tisagenlecleucel in Adult Relapsed or Refractory Diffuse Large B-Cell Lymphoma. *N Engl J Med* (2019) 380:45–56. doi: 10.1056/NEJMoa1804980
38. Rubin DB, Danish HH, Ali AB, Li K, LaRose S, Monk AD, et al. Neurological Toxicities Associated With Chimeric Antigen Receptor T-Cell Therapy. *Brain* (2019) 142:1334–48. doi: 10.1093/brain/awz053
39. Cottereau AS, Meignan M, Nioche C, Capobianco N, Clerc J, Chartier L, et al. Risk Stratification in Diffuse Large B Cell Lymphoma Using Lesion Dissemination and Metabolic Tumor Burden Calculated From Baseline PET/CT. *Ann Oncol* (2021) 32:404–411. doi: 10.1016/j.annonc.2020.11.019
40. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A Radiomics Approach to Assess Tumour-Infiltrating CD8 Cells and Response to anti-PD-1 or anti-PD-L1 Immunotherapy: An Imaging Biomarker, Retrospective Multicohort Study. *Lancet Oncol* (2018) 19:1180–91. doi: 10.1016/S1470-2045(18)30413-3
41. Dercle L, Lu L, Schwartz LH, Qian M, Tejpar S, Eggleton P, et al. Radiomics Response Signature for Identification of Metastatic Colorectal Cancer Sensitive to Therapies Targeting EGFR Pathway. *J Natl Cancer Inst* (2020) 112:902–12. doi: 10.1093/jnci/djaa017
42. Dercle L, Fronheiser M, Lu L, Du S, Hayes W, Leung DK, et al. Identification of Non-Small Cell Lung Cancer Sensitive to Systemic Cancer Therapies Using Radiomics. *Clin Cancer Res* (2020) 26:2151–62. doi: 10.1158/1078-0432.CCR-19-2942
43. Seban R-D, Nemer JS, Marabelle A, Yeh R, Deutsch E, Ammari S, et al. Prognostic and Theranostic 18F-FDG PET Biomarkers for Anti-PD1 Immunotherapy in Metastatic Melanoma: Association With Outcome and Transcriptomics. *Eur J Nucl Med Mol Imaging* (2019) 46:2298–310. doi: 10.1007/s00259-019-04411-7
44. Seban R-D, Mezquita L, Berenbaum A, Dercle L, Botticella A, Le Pechoux C, et al. Baseline Metabolic Tumor Burden on FDG PET/CT Scans Predicts Outcome in Advanced NSCLC Patients Treated With Immune Checkpoint Inhibitors. *Eur J Nucl Med Mol Imaging* (2020) 47:1147–57. doi: 10.1007/s00259-019-04615-x
45. Dercle L, Seban R-D, Lazarovici J, Schwartz LH, Houot R, Ammari S, et al. 18f-Fdg PET and CT Scans Detect New Imaging Patterns of Response and Progression in Patients With Hodgkin Lymphoma Treated by Anti-Programmed Death 1 Immune Checkpoint Inhibitor. *J Nucl Med* (2018) 59:15–24. doi: 10.2967/jnumed.117.193011
46. Seban R-D, Moya-Plana A, Antonios L, Yeh R, Marabelle A, Deutsch E, et al. Prognostic 18F-FDG PET Biomarkers in Metastatic Mucosal and Cutaneous Melanoma Treated With Immune Checkpoint Inhibitors Targeting PD-1 and CTLA-4. *Eur J Nucl Med Mol Imaging* (2020) 47:2301–12. doi: 10.1007/s00259-020-04757-3
47. Dercle L, Ammari S, Champiat S, Massard C, Ferté C, Taihi L, et al. Rapid and Objective CT Scan Prognostic Scoring Identifies Metastatic Patients With Long-Term Clinical Benefit on anti-PD-1/-L1 Therapy. *Eur J Cancer* (2016) 65:33–42. doi: 10.1016/j.ejca.2016.05.031
48. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and Challenges for the Implementation of Computational Medical Imaging (Radiomics) in Oncology. *Ann Oncol* (2017) 28:1191–206. doi: 10.1093/annonc/mdx034
49. Reuzé S, Schernberg A, Orlhac F, Sun R, Chargari C, Dercle L, et al. Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls, and Challenges. *Int J Radiat Oncol Biol Phys* (2018) 102:1117–42. doi: 10.1016/j.ijrobp.2018.05.022
50. Sesques P, Ferrant E, Safar V, Wallet F, Tordo J, Dhomps A, et al. Commercial anti-CD19 Car T Cell Therapy for Patients With Relapsed/Refractory Aggressive B Cell Lymphoma in a European Center. *Am J Hematol* (2020) 95:1324–33. doi: 10.1002/ajh.25951
51. Krebs S, Ponomarev V, Slovin S, Schöder H. Imaging of CAR T-Cells in Cancer Patients: Paving the Way to Treatment Monitoring and Outcome Prediction. *J Nucl Med* (2019) 60:879–81. doi: 10.2967/jnumed.119.227561
52. Imber BS, Sadelain M, DeSelm C, Batlevi C, Brentjens RJ, Dahi PB, et al. Early Experience Using Salvage Radiotherapy for Relapsed/Refractory non-Hodgkin Lymphomas After CD19 Chimeric Antigen Receptor (CAR) T Cell Therapy. *Br J Haematol* (2020) 190:45–51. doi: 10.1111/bjh.16541

Conflict of Interest: LS is a paid advisory board member for Roche and Novartis, and reports receiving commercial research grants from Merck and Boehringer Ingelheim. RR has consulting or advisory role: Atara Biotherapeutics, Novartis, Magenta Therapeutics, Bristol-Myers Squibb, Gilead Sciences, and received research Funding from Atara Biotherapeutics, Incyte, Pharmacyclics, Shire, Immatics, Takeda, Gilead Sciences, Precision Biosciences, Astellas Pharma. RB received honoraria from Novartis and Gilead.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

MINI REVIEW
published: 28 May 2021
doi: 10.3389/fonc.2021.560585

Optimizing the Management of Cancer Patients Treated With Systemic Therapies During the COVID-19 Pandemic: The New Role of PCR and CT Scan

Alessandro A. Viansone[1,2*], *Samy Ammari*[3], *Laurent Dercle*[4] *and Monica Arnedos*[1]

[1] Breast Unit, Department of Medicine, Gustave Roussy Cancer Campus, Villejuif, France, [2] Breast Unit–Oncology Unit, Department of Medicine, University Hospital of Parma, Parma, Italy, [3] Radiology Department, Gustave Roussy Cancer Campus, Villejuif, France, [4] Radiology Department, Columbia University Medical Center New York Presbyterian Hospital, New York, NY, United States

OPEN ACCESS

Edited by:
Samata Kakkad,
Merck, United States

Reviewed by:
Jacobo Rogado,
Hospital Universitario Infanta Leonor, Spain
Monika Engelhardt,
University of Freiburg, Germany

***Correspondence:**
Alessandro A. Viansone
alessandro.viansone@gustaveroussy.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 09 May 2020
Accepted: 12 March 2021
Published: 28 May 2021

Citation:
Viansone AA, Ammari S, Dercle L and Arnedos M (2021) Optimizing the Management of Cancer Patients Treated With Systemic Therapies During the COVID-19 Pandemic: The New Role of PCR and CT Scan. Front. Oncol. 11:560585. doi: 10.3389/fonc.2021.560585

In late 2019 and early 2020, the world witnessed the outbreak of the SARS-CoV-2 (also referred as COVID-19) in Wuhan, China. Its rapid expansion worldwide and its contagiousness rate have forced the activation of several measures to contain the pandemic, mostly through confinement and identification of infected patients and potential contacts by testing.

Keywords: COVID-19, PCR, CT-scan, oncology, cancer

INTRODUCTION

It is well established that patients with cancer are more susceptible to infections since they tend to be older, have multiple comorbidities, and because of the immunosuppressive state caused by anticancer treatments (1), so they could be potentially at particular risk from COVID-19 infection (2).

In a Chinese analysis (3) Professor He and colleagues found that patients with cancer presented an important risk of secondary events due to COVID-19 if we compare it with a population without cancer: 39% in the oncologic population *versus* 8%, hazard ratio of 3.56. Zhang and colleagues in a retrospective analysis identified 28 patients positive for COVID and with a cancer among 1,276 patients admitted in hospital. This prevalence (2.2%) is 1.7 (95% confidence interval, 1.2, 2.4) times higher than the Chinese population of the same age (4). The probability of dying for COVID with a cancer was of 28.6% (5). Zhang found that the administration of anticancer therapies was an independent predictor of death. They also described the high proportion of patients who acquired the infection in hospital, when they came for cancer treatment.

Although these data included a very small and heterogeneous sample of patients with cancer, reports from Italy confirm the potential higher risk of COVID-19 infection in patients with cancer, indicating that the 20% of patients dead for COVID-19 infection also had a concomitant diagnosis of cancer (6). More recently, clinical data on COVID-19 cases from two hospitals in New York City (7) observed that 23 out of the 393 (5.9%) reported cases were patients with cancer, and 10 of them required invasive mechanical ventilation (representing 7.7% of the total of patients requiring

invasive mechanical ventilation). In another institution from New York City, 28% of COVID-19 + cancer patients died from COVID-19 with a case fatality rate of 37% for hematologic malignancies and 25% for solid malignancies. This study reported that in this population, older age, multiple comorbidities, need for ICU support, and elevated levels of D-dimers, LDH, and lactate predicted poorer outcome (7).

Therefore, healthcare professionals were rapidly faced with the challenge of profoundly re-organizing healthcare systems at unprecedented pace during the COVID-19 crisis to balance the competing risks of death from cancer *vs.* death or serious complications from infection.

The Chinese series suggested that postponing adjuvant chemotherapy or elective surgery for less aggressive cancers should be considered. Furthermore, more intensive care should be done for the patients with cancer who are infected (3). These measures have been later taken by healthcare organizations like ASCO and ESMO who have deepened these recommendations extending it to postponing clinics or balancing the cost/benefit ratio according to outcome, with prioritization of adjuvant therapies (8). Similarly, the benefits and risks of palliative therapies and the options of temporary stopping the therapy and switch to oral drugs, if available, during the pandemic needed to be considered. Recommendations for management of cancer patients in clinical trials have also been made available by different regulatory agencies (9). However, as this outbreak prolongs in time and with the unknown risk-benefit balance between undertreating patients with cancer (with resultant increase in cancer-related morbidity and mortality) and preventing the expected higher morbidity and mortality from COVID-19, initiation of systemic cancer therapy seems unavoidable.

The scarce of systematic information on prevalence and incidence in overall population toughen the possibility of real comparisons with patients with cancer although contagiousness rates seem rather high. For example, estimates on infection rates vary from 0.76% for residents of Iceland to 36% for residents of Boston. This likely overstates the overall population mean, which some observers have suggested is around 40% (10).

Systematic reports about the prevalence of cancer in patients with COVID-19 and the real incidence of COVID-19 infection among cancer patients are starting to be made public. A review and meta-analysis reported by Desai et al. (11) found 11 studies of patients with COVID-19 with the data of cancer prevalence: 2% (95% CI 2.0 to 3.0%: I^2 = 83.2%) in patients treated for COVID-19. A similar meta-analysis by Emami et al. (12) reported a prevalence of malignancy of 0.92% (95% CI, 0.56–1.34%). In a recent dedicated session to cancer patients and COVID-19 infection during the last AACR annual symposium, several cancer institutions and hospitals across the world presented updates on their management and outcome of cancer patients with COVID-19 infection. These series reported variability within different countries in terms of incidence and prevalence. For example, in our own series, of Gustave Roussy Hospital, including more than 1,300 tested cancer patients, COVID-19 positivity was observed in around 12%, whereas retrospective series including 1,524 cancer patients from Wuhan reported a COVID-19 positivity rate of only 0.79% (13).

The standard method to diagnose infection by COVID-19 is through identification by RT-PCR SARS-CoV-2 testing. Since SARS-CoV-2 is usually transmitted by the upper respiratory tract, mostly swab samples are taken from the nasopharynx. It has also been accepted to perform the swabs directly to the oropharynx, although some studies suggest that the nasopharynx would be somewhat more sensitive than the oropharynx (14). Positivity from nasopharynx swabs is variable and ranges from 53.6 to 73.3% depending on the series (15). There are several factors that might affect the performance (sensitivity and specificity) of the test like the quality of the sample, the sampling technique, transportation process, or limited gene detection. In fact, it has been shown that high viral loads soon after symptom onset, which then gradually decreased towards the detection limit at about day 21, with no obvious difference in viral loads across sex, age groups, and disease severity (16). In fact, similar viral loads have been documented in the upper respiratory tract of both symptomatic and asymptomatic cases (17) and in the pre-symptomatic phase (18).

Several strategies are available in order to increase sensitivity and specificity of this testing. In a report of 67 patients with confirmed COVID-19 infection, the duration of positive test in nasopharyngeal swabs has a median of 12 days (range, 3–38), in sputum of 19 days (range, 5–37), and in stools of 18 days (range, 7–26). SARS-CoV-2 RNA was detectable for a duration of 30 days (19). After a negative test of nasopharyngeal swabs among 46 patients, 28 (60.9%) and 14 (30.4%) patients were still positive in sputum and stools.

Another approach consists in the realization of chest CT scans in patients suspected or tested for COVID-19 in addition to RT-PCR for nasal or oropharynx swabs. Consistently between patients' series, the main symptoms associated to SARS-CoV-2 infection in addition to fever is the presence of pulmonary symptoms ranging from dry cough to pneumonia up to acute distress respiratory syndrome leading to death (3). Therefore initially it was suggested that lung cancer patients or patients who had suffered previous lung surgery would be at higher risk of lung complications from SARS-CoV-2 so more intensive follow-up and chest CT scrutiny should be required for these patients. Initial data from Wuhan series (20) indeed reported a higher risk of COVID-19 infection among their cancer patients (7 out of 28; 25%). A recent international series compilation of 200 patients with thoracic tumors affected with COVID-19 (21) reported an extremely high death rate of 34.6%, mostly due to acute respiratory distress syndrome and multi-organ failure. These findings have not been replicated by other series like Wuhan reports (22) where mortality was not affected by type of cancer (any cancer *vs.* lung cancer; HR = 0.727; p = 0.589). What it has been noteworthy is that chest CT demonstrates typical radiographic features in almost all COVID-19 patients irrespective of presence of cancer or type of cancer. These include ground-glass opacities, multifocal patchy consolidation, and/or interstitial changes with a peripheral distribution (23). Those typical pulmonary abnormalities were also observed in

patients with negative RT-PCR results but clinical symptoms in small-scale studies (24).

A report of the correlation of chest CT and RT-PCR testing in 1,014 cases from China found a RT-PCR positivity in 59% (n = 601) of the patients and a chest CT positivity of 88% (n = 888). The negative RT-PCR results is correlated to a 75% (n = 308) with positive chest CT findings. The combination of RT-PCR and chest CT gave respectively a sensitivity, specificity, and accuracy of 97% (n = 580), 25% (n = 105), and 68% (n = 685), a positive predictive value of 65% (n = 580), and a negative predictive value of 83% (n = 105) (25). Another work analyzed 1,099 hospitalized patients with a positive test for COVID-19: 86% (n = 840) of patients had CT imaging with finding of ground-glass opacity, local patchy shadowing, or interstitial pneumopathies; 17.9% (n = 157) of patients has no radiographic abnormality (26).

Despite the standardized use of CT scans in addition to (or instead of) RT-PCR for the diagnoses of a suspected COVID-19 infection might still be controversial (27). First, the findings on chest imaging in COVID-19 are indeed not specific, and overlap with other infections, including influenza, H1N1, SARS, and MERS. Second, there are issues related to cleaning imaging equipment to control the spread of infection in health care facilities where CT scans are frequently used. For instance, portable radiography units are less expensive, can be cleaned easier, and could be an alternative. Chest CT scan might provide prognostic information as, some published data reported that the presence of a patchy consolidation by lung CT scan at patient admission was associated to possibility of a severe event in a multivariate analysis in COVID-19+ patients (HR = 5.438; CI 1.498–19.748; p = 0.010). In addition, it might prove useful for the management of patients with COVID-19 infection, especially in highly symptomatic cases. Indeed, as additional knowledge of this infection becomes available, several reports have shown that COVID-19 infection might be associated to an inflammatory syndrome evidenced by high levels of inflammatory markers and increased risk of thromboembolism associated to this infection (28). This might be important in patients with cancer with already increased phenomena of coagulopathy and thrombosis. Hence, the interest of associating chest imaging in cancer patients to identify underlying pulmonary embolism, which might contribute to worsen respiratory symptoms and require specific additional treatment (29). Beyond the role of CT scan without intravenous contrast agent injection for the diagnostic workup, prognostic evaluation, and follow-up of COVID-19 infection, selected patients may benefit from contrast enhanced CT pulmonary angiography to diagnose potentially life-threatening pulmonary embolism and start appropriate therapies. A study reported a high frequency of either pulmonary embolism in critically ill ICU patients with COVID-19 complications [7.1, 20.6 (30), and 49% (31)].

Rogado et al. detected 45/1,069 COVID-19 diagnoses in cancer patients *vs* 42,450/6,662,000 in total population (p < 0.00001) in a Spanish hospital. Mortality rate: 19/45 cancer patients *vs* 5,586/42,450 (p = 0.0001). Mortality was associated with older median age, adjusted by staging, and histology (74 *vs* 63.5 years old, OR 1.06, p = 0.03). Patients who combined hydroxychloroquine and azithromycin presented 3/18 deaths, regardless of age, staging, histology, cancer treatment, and comorbidities (OR 0.02, p = 0.03) (32).

In a Spanish series Mestre-Gomez et al. retrospectively reviewed 452 electronic medical records of patients admitted to Internal Medicine Department of a secondary hospital in Madrid during COVID-19 pandemic outbreak. Ninety-one patients had a Computed Tomography pulmonary angiography (CTPA). The cumulative incidence of PE was assessed with a clinical, analytical, and radiological characterization compared in patients with and without PE. The incidence of PE was 6.4%. They evaluated the D-dimer peak and they found a significant elevation in PE *vs* non-PE patients (14,480 *vs* 7,230 mcg/dl, p = 0.03). In multivariate analysis that plasma D-dimer peak was confirmed as an independent predictor of PE with a best cut off point of >5,000 μg/dl (33).

In a hematologic series at the University Hospital Freiburg, Shoumariyeh et al. analyzed a retrospective cohort of 39 patients with hematological and solid cancers hospitalized for COVID-19. With univariate and multivariate Cox analysis they found that the presence of a malignancy was not significantly associated with survival or severe events. Instead the high IL-6 levels at COVID-19 diagnosis (HR = 6.95, P = .0121) and age ≥65 years (HR = 6.22, P = .0156) were related to mortality. Another find of Shoumariyeh et al. was about patients with a hematological malignancy that showed a longer duration of clinical improvement and longer hospitalization compared to patients with a solid cancer (34).

In our institution, we have therefore taken the decision to perform RT-PCR SARS-CoV-2 testing by nasopharyngeal swab in a specifically dedicated area to our patients at day −2 or −1 before administration of chemotherapy or immunotherapy treatment, if feasible. The reason is to discover the disease in its pre-symptomatic phase to prevent initiation of a potentially immunosuppressive treatment to diminish chances of major complications and to be able to intensively follow up these patients. In addition, identifying those pre-symptomatic and asymptomatic patients will help rapidly isolate these COVID-19+ patients to prevent further spreading of the disease to health personnel and to other patients. For those patients with respiratory symptoms a chest CT scan is also performed to increase sensitivity to the diagnoses of the disease. In addition, CT pulmonary angiogram is performed for those cases with acute inflammatory syndrome to rule out pulmonary embolism and to initiate intensive anticoagulation therapy. For patients confirmed COVID-19+, systemic treatments are delayed and patient is surveyed or treated accordingly. All COVID-19+ cases will be re-tested at around 15 days from the initial testing. For those patients asymptomatic throughout the infection with a negative test at 15 days, systemic treatment can be initiated. In case of a symptomatic course of the infection, systemic anticancer treatment can be started around 15 days after the end of symptoms. If the second PCR test still is positive, patients need to be rested at 7–15 days and systemic treatment must be delayed if possible. For those cases still COVID-19+ after at least 15 days from the diagnostic PCR,

chemotherapy can be initiated if the patient has no symptoms for at least 7 days for those cases where benefit/risk is in favor for cancer therapy. Soon, as validated RT-PCR SARS-CoV-2 testing becomes available in local laboratories, testing will be performed at the same time as the standard pre-chemotherapy blood samples before every cycle of treatment, with results being directly reported to our institution.

In the near future, the availability of laboratory IgM or IgG validated testing to evaluate the previous SARS-CoV-2 exposure will be helpful in order to fully picture the real incidence of this disease in overall population and more particularly in cancer patients. This will become crucial if further studies observe the acquisition of an adaptive immunity capable of preventing re-infection to SARS-CoV-2 infection, or at least to severe forms to allow treating cancer patients without additional risk from COVID-19.

AUTHOR CONTRIBUTIONS

AV and MA: oncologist point of view and writing and bibliographic search. SA and LD: radiologist point of view and writing and figure creation. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Kamboj M, Sepkowitz KA. Nosocomial Infections in Patients With Cancer. *Lancet Oncol* (2009) 10:589–59. doi: 10.1016/S1470-2045(09)70069-5
2. Sica A, Massarotti M. Myeloid Suppressor Cells in Cancer and Autoimmunity. *J?Autoimmun* (2017) 85:117–25. doi: 10.1016/j.jaut.2017.07.010
3. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer Patients in SARS-CoV-2 Infection: A Nationwide Analysis in China. *Lancet Oncol* (2020) 21:335–7. doi: 10.1016/S1470-2045(20)30096-6
4. Ferlay J, Ervik M, Lam F. Global Cancer Observatory: Cancer Today. *Int Agency Res Cancer* (2018) 2:435–54. doi: 10.1002/ijc.33232
5. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in CHINA: Summary of a rEport of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* (2020) 2:321–42. doi: 10.1001/jama.2020.2648
6. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA* (2020) 3:654–6. doi: 10.1001/jama.2020.4683
7. Mehta V, Goel S, Kabarriti R, Cole D, Goldfinger M, Acuna-Villaorduna A, et al. Case Fatality Rate of Cancer Patients with COVID-19 in a New York Hospital System. *Cancer Discov* (2020) 10:935–41. doi: 10.1158/2159-8290.CD-20-0516
8. Hanna TP, Gerald GA, Christopher CM. Cancer, COVID-19 and the Precautionary Principle: Prioritizing Treatment During a Global Pandemic. *Nat Rev* (2020) 17:268–70. doi: 10.1038/s41571-020-0362-6
9. Bruno de Paula H, Araújo I, Bandeira L, Barreto NMPB, Doherty GJ. Recommendations From National Regulatory Agencies for Ongoing Cancer Trials During the COVID-19 Pandemic. *Lancet Oncol* (2020) 8):455–7. doi: 10.1016/S1470-2045(20)30226-6
10. Oran P, Topol J. Getting a Handle on Asymptomatic SARS-CoV-2 Infection. *Scripps Res* (2020) 5:678–9. doi: 10.7326/M20-3012
11. Desai R. COVID-19 and Cancer: Lessons from a Pooled Meta-Analysis. *JCO* (2020) 8:557–9. doi: 10.1200/GO.20.00097
12. Song Y, Xia J, Guo Q, Song T, He J, Peiris M. Prevalence of Underlying Diseases in Hospitalized Patients with COVID-19: A Systematic Review and Meta-Analysis. *Arch Acad Emerg Med* (2020) 8:765–70. doi: 10.1016/j.ajem.2020.10.022
13. Yu J, Ouyang W, Chua MLK, Xie C. SARS-CoV-2 Transmission in Patients with Cancer at a Tertiary Care Hospital in Wuhan, China. *JAMA Oncol* (2020) 25:34–8. doi: 10.1001/jamaoncol.2020.0980
14. Zou L, Ruan F, Huang M, Liang L, Kang M. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *N Engl J Med* (2020) 382:1177–9. doi: 10.1056/NEJMc2001737
15. Yang Y, Yang M, Shen C. Evaluating the Accuracy of Different Respiratory Specimens in the Laboratory Diagnosis and Monitoring the Viral Shedding of 2019-nCoV Infections. *Nat* (2020) 2:212–5. doi: 10.1101/2020.02.11.20021493
16. He X, Lau E, Wu P. Temporal Dynamics in Viral Shedding and Transmissibility of COVID-19. *Nat Med* (2020) 4:665–7. doi: 10.1038/s41591-020-0869-5
17. Hoehl S, Rabenau H, Berger A. Evidence of SARS-CoV-2 Infection in Returning Travelers From Wuhan, China. *N Engl J Med* (2020) 4:1278–80. doi: 10.1056/NEJMc2001899
18. Bhangu A, Lawani I, Ng-Kamstra JS, Wang Y, Chan A. Global Guidance for Surgical Care During the COVID-19 Pandemic. *Br J Surg* (2020) 15:898–1. doi: 10.1002/bjs.11646
19. Horn L, Whisenant JG, Torri V, Huang L-C, Trama A, Paz-Ares LG, et al. Thoracic Cancers International COVID-19 Collaboration (TERAVOLT): Impact of Type of Cancer Therapy and COVID Therapy on Survival. *J Clin Oncol* (2020) 38:(18 Suppl). doi: 10.1200/JCO.2020.38.18_suppl.LBA111
20. Zhang L, Zhu F, Xie L, Wang C, Wang J, Chen R, et al. Clinical Characteristics of COVID-19-Infected Cancer Patients: A Retrospective Case Study in Three Hospitals Within Wuhan, China. *Ann Oncol* (2020) 3:363–70. doi: 10.1016/j.annonc.2020.03.296
21. Mitchell RL, Blakely LJ, Schleicher SM, Poole SL, Dickson NR, Patton J, et al. Maintaining Treatment Volumes During the COVID-19 Pandemic. *J Clin Oncol* (2020) 38(suppl 29; abstr 103). doi: 10.1200/JCO.2020.38.29_suppl.103
22. Xie X, Zhong Z, Zhao W. Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology* (2020) 6:65–74. doi: 10.1148/radiol.2020200343
23. Chung M, Bernheim A, Mei X. CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV). *Radiology* (2020) 7:660–7. doi: 10.1148/radiol.2020200230
24. Huang P, Liu T, Huan L. Use of Chest CT in Combination with Negative RT-PCR Assay for the 2019 Novel Coronavirus but High Clinical Suspicion. *Radiology* (2020) 8:445–553. doi: 10.1148/radiol.2020200330
25. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* (2020) 8:32–40. doi: 10.1148/radiol.2020200642
26. Guan W, Zheng N, Wen L. Clinical Characteristics of Coronavirus Disease 2019 in China. *NEJM* (2020) 4:67–74. doi: 10.1056/NEJMoa2002032
27. Hope A. A role for CT in COVID-19? What Data Really Tell us so Far. *Lancet* (2020) 395:665–78. doi: 10.1016/S0140-6736(20)30728-5
28. Klok FA, Kruip M, vanderMeer N. Incidence of Thrombotic Complications in Critically ill ICU Patients with COVID-19. *Thromb Res* (2020) 4:323–45. doi: 10.1016/j.thromres.2020.04.013
29. Zhang Y, Xiao M, Zhang S, Xia P. Coagulopathy and Antiphospholipid Antibodies in Patients with COVID-19. *N Engl J Med* (2020) 8:567–34. doi: 10.1056/NEJMc2007575
30. Poissy J, Goutay J, Caplan M, Parmentier E, Duburcq T, Lassalle F, et al. Pulmonary Embolism in COVID-19 Patients: Awareness of an Increased Prevalence. *CIRCULATION* (2020) 142:184–6. doi: 10.1161/CIRCULATIONAHA.120.047430
31. Klok FA, Kruipb MJHA, van der Meercd NJM, Arbouse MS, Gommersf D, Kantg KM, et al. Confirmation of the High Cumulative Incidence of Thrombotic Complications in Critically ill ICU Patients with COVID-19: An Updated Analysis. *Thromb Res.* (2020) 191:148–50. doi: 10.1016/j.thromres.2020.04.041
32. Rogado J, Obispo B, Pangua C, Serrano-Montero G, Martín Marino A, Pérez-Pérez M, et al. COVID-19 Transmission, Outcome and Associated Risk Factors in Cancer Patients at the First Month of the Pandemic in a Spanish

Hospital in Madrid. *Clin Trans Oncol* (2020) 22:2364–8. doi: 10.1007/s12094-020-02381-z

33. Mestre-Gómez B, Lorente-Ramos RM, Rogado J, Franco-Moreno A, Obispo B, Salazar-Chiriboga D, et al. Incidence of Pulmonary Embolism in Non-Critically Ill COVID-19 Patients. Predicting Factors for a Challenging Diagnosis. *J Thromb Thrombolysis* (2020) 1–7. doi: 10.1007/s11239-020-02190-9
34. Shoumariyeh K, Biavasco F, Ihorst G, Rieg S, Nieters A, Kern WV, et al. COVID-19 in Patients with Hematological and Solid Cancers at a Comprehensive Cancer Center in Germany. *Cancer Med* (2020) 22):8412–22. doi: 10.1002/cam4.3460

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 June 2021
doi: 10.3389/fonc.2021.659313

Markers of Prognosis for Early Stage Cervical Cancer Patients (Stage IB1, IB2) Undergoing Surgical Treatment

Chen Xu[1,2†], Tie Ma[3†], Hongzan Sun[1,2], Xiaohan Li[3] and Song Gao[4*]*

[1] Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China, [2] Liaoning Provincial Key Laboratory of Medical Imaging, Shengjing Hospital of China Medical University, Shenyang, China, [3] Department of Pathology, Shengjing Hospital of China Medical University, Shenyang, China, [4] Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang, China

Background: For individuals with cervical cancer, large tumor volume, lymph node metastasis, distant metastasis, and parauterine infiltration are usually associated with a poor prognosis. Individuals with stage 1B1 and 1B2 cervical cancer usually do not have these unfavorable prognostic factors. Once the disease progresses, the prognosis becomes extremely poor. Therefore, investigating the prognostic markers of these cervical cancer patients is necessary for treatment.

Methods: This retrospective study included 95 cervical cancer patients treated with surgery. The patients were divided into progressor and non-progressor groups according to postoperative follow-up results. T-test (or Mann–Whitney U test), chi-squared test (or Fisher's exact test) and receiver operating characteristic (ROC) curves were used to evaluate imaging, hematology, and clinicopathological index differences between the two groups. Cox analysis was performed to select the independent markers of progression-free survival (PFS) when developing the nomogram. Validation of the nomogram was performed with 1000 bootstrapped samples. The performance of the nomogram was validated with ROC curves, generated calibration curves, and Kaplan-Meier and decision curve analysis (DCA).

Results: Cervical stromal invasion depth, lymphovascular space invasion (LVSI), human papilloma virus (HPV-16), Glut1, D-dimer, SUVmax and SUVpeak showed significant differences between the two groups. Multivariate Cox proportional hazard model showed SUVpeak ($p = 0.012$), and HPV-16 ($p = 0.007$) were independent risk factors and were used to develop the nomogram for predicting PFS. The ROC curves, Kaplan-Meier method, calibration curves and DCA indicated satisfactory accuracy, agreement, and clinical usefulness, respectively.

Conclusions: SUVpeak level (≥7.63 g/cm3) and HPV-16 negative status before surgery were associated with worse PFS for patients with cervical cancer. Based on this result, we constructed the nomogram and showed satisfactory performance. Clinically, individualized clinical decision-making can be performed on patients based on this result.

Keywords: cervical cancer, positron-emission tomography, computed tomography, human papilloma virus, prognosis

OPEN ACCESS

Edited by:
Romain-David Seban, Institut Curie, France

Reviewed by:
Jitti Hanprasertpong, Prince of Songkla University, Thailand
Valerio Gallotta, Catholic University of the Sacred Heart, Italy

****Correspondence:***
Xiaohan Li
2412171534@qq.com
Song Gao
gaos@sj-hospital.org

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *27 January 2021*
Accepted: *17 May 2021*
Published: *02 June 2021*

Citation:
Xu C, Ma T, Sun H, Li X and Gao S (2021) Markers of Prognosis for Early Stage Cervical Cancer Patients (Stage IB1, IB2) Undergoing Surgical Treatment. Front. Oncol. 11:659313. doi: 10.3389/fonc.2021.659313

INTRODUCTION

Cervical cancer has the second highest incidence of female malignant tumors (1). The International Federation of Gynecology and Obstetrics (FIGO) cancer staging system is used in the formulation of treatment and prognosis plans for patients with cervical cancer. In the recently updated 2018 FIGO staging, it is noteworthy that, assessment of the abdominopelvic retroperitoneal lymph nodes was included in the FIGO system (2–4). Regardless of parametrial infiltration and tumor size, the presence of nodal metastases now indicates stage IIIC. Radical hysterectomy with lymphadenectomy will be effective for those cervical cancer patients (stage 1B1, 1B2) with the following characteristics: invasive carcinoma confined to the uterine cervix, no more than 5 mm invasion, tumor size < 4 cm in its greatest dimension and no lymph node metastasis. Fortunately, these patients often do not need postoperative adjuvant radiotherapy, since, of course, radiotherapy has serious side effects (5). However, although the prognosis of cervical cancer patients with these features is excellent, if they relapse, then their prognosis is very poor (6). Therefore, identifying those cervical cancer patients who are prone to recurrence after surgery is of great significance, since they will require postoperative adjuvant treatment such as radiotherapy and chemotherapy, as well as closer follow-up, to reduce the chance of recurrence and death.

The following three methods were used in predicting the prognosis of cervical cancer: 1) evaluation of pathological characteristics obtained from postoperative pathological specimens, including aberrant molecular signaling pathway proteins; 2) preoperative imaging; and 3) hematological examination.

It is well known that pathological features such as age, tumor stage and grade, cervical stromal invasion depth, lymphovascular space invasion (LVSI); and preoperative high-risk human papillomavirus (HR-HPV) statue were important factors in the prognosis of cervical cancer. Different molecular factors involving loss of tumor suppressor genes and aberrant molecular signaling pathways, such as TP53-induced glycolysis and apoptosis regulator (TIGAR), cytokine involved primarily in angiogenesis (Vegf-A), mammalian facilitative glucose transporter family (Glut-1), epithelial mesenchymal transition related protein (E-cadherin), immune-linked factor by triggering pro-inflammatory immune-associated reactions (Cox-2), the extracellular matrix molecule (Tenascin-C), have recently been identified in the pathogenesis of cervical cancer (7–11). The impact of these proteins on cervical cancer prognosis needs further investigation.

^{18}F-FDG PET/CT, a functional imaging technique, provides quantified metabolic information (the mean and maximum standardized uptake values (SUV_{mean} and SUV_{max}), metabolic tumor volume (MTV) and total lesion glycolysis (TLG))and has a well-established role in the management of patients with cervical cancer (12–14). Many studies have shown that hematological parameters such as hemoglobin (Hb), coagulation indexes [D-dimer, fibrinogen (Fg)], tumor marker for squamous cell carcinoma [squamous cell carcinoma antigen (SCCA)], and novel systemic inflammation response index (SIRI) are potential prognostic marker of malignant tumors (15–19).

In this article, our aim was to find independent prediction parameters from these parameters and establish and validate a nomogram to predict the progression-free survival (PFS) of patients with cervical cancer.

MATERIALS AND METHODS

Population Characteristics

Patients deemed to have a high suspicion of cervical cancer with strong evidence from PET/CT (PET/CT examinations were performed within one week before treatment) were retrospectively enrolled from January 1st, 2013, to December 31th, 2015 with the following criteria: (1) The patient underwent surgery (open surgery) and was confirmed as cervical squamous cell carcinoma by postoperative pathological results; (2) According to FIGO2018 pathological staging, the patient was confirmed to be stage IB1 or IB2; (3) Patients who had not undergone other treatments without surgery, and (4) Patients with complete follow-up records.

The Institutional Review Boards at the local institutions endorsed the study. We implemented it in accordance with the ethical standards of the 1964 Helsinki Declaration and subsequent amendments, and gave up informed consent of all participants.

Surgical Protocol

Numerous studies have compared the effect of the Minimally Invasive and Abdominal Radical Hysterectomy for cervical cancer (20–22). The conclusion shows that for early cervical cancer the minimally invasive surgery (laparoscopic or robot-assisted radical hysterectomy) has a higher recurrence rate than open surgery. In recent years, patients of our hospital have also accepted open abdominal surgery as the main surgical method. Therefore, the patients included all received the open abdominal Type-C radical hysterectomy. The incision was located on the lower abdomen. The resection scope included the uterus, parauterine, upper vagina, and partial tissues around the vagina and pelvic lymph nodes. A sufficient length of adjacent connective tissues, including the front vesicocervix ligament (anterior and posterior lobes), the lateral main ligament, the posterior uterosacral ligament and rectovaginal ligament, also should be removed. The ovary was selectively removed according to the patient's clinical conditions and willingness. The resection of lymph nodes involved the obturator lymph nodes and the internal, common, and external iliac vessel lymph nodes. Para-aortic nodes should be removed only in the following cases: (1) The preoperative PET/CT showed that there were suspicious metastatic lymph nodes in the para-aortic area [SUVmax of lymph nodes is higher than the background metabolism level (23, 24)]; and (2) the possibility of metastasis in the pelvic or para-aortic lymph nodes was suspected intra-operatively.

Tumor Recurrence Prediction and Patient Follow-Up

PFS was defined as the time interval between the date of surgery and the date at which the first recurrence of disease was confirmed. Starting from the completion of the surgery, follow-up examinations (chest CT, abdominal CT and pelvic MR) were performed approximately every 3 months for the first 2 years, and then every 6 months or 1 year for the next 3 years. Recurrent disease was defined as local recurrence, metastasis of pelvic or para-aortic lymph nodes, and metastases in distant organs following surgery. For patients with abnormal follow-up examinations results, PET/CT or biopsy was recommended, and the results used as the standard for recurrence. The end of the follow-up time is December 31th, 2020.

Pathologic Diagnosis

A pathologist with more than 10 years of experience evaluated sections using HE staining and immunohistochemistry (IHC) of CD31 and D2-40. The following results based on postoperative pathology report were obtained: tumor differentiation grade, stage, lymph node metastasis, LVSI, cervical stromal invasion depth and histologic tumor type.

Tumor sections were obtained from our hospital pathology department and we purchased antibodies from Abcam (Shanghai, China). In addition, Cox-2, E-cadherin, Tenascin-C, Glut-1, Tigar, and Vegf-A protein expression were studied based on IHC, which was performed using the Leica BOND-MAX system (Leica Biosystems, Shanghai, China). Thorough mixing was done with the following polyclonal antibodies: Cox-2 at 1:100 dilution; E-cad at 1:700 dilution; Tenascin-C at 1:100 dilution; Glut-1 at 1:500 dilution; Tigar at 1:300 dilution; and Vegf-A at 1:50 dilution. Then the tissue section was sealed after the Imaging Mass Cytometry (IMC) was finished, the section was placed under the lens of a scanner (Pannoramic MIDI, 3D Histech), and moved gradually, imaging while moving and then scanning and imaging all the tissue information on the tissue section to form a file. After image scanning was completed, the DensitoQuant software application in the QuantCenter (QuantCenter is a piece of analysis software for Pannoramic viewer) automatically recognized and set all dark browns on the tissue section as strongly positive, brown-yellow as moderately positive, light yellow as weakly positive, and blue cell nuclei as negative. Furthermore, for each tissue, strongly positive, moderately positive, weakly positive and negative areas, and the percentage of positive areas, were identified. A negative expression was defined by a < 10% positive expression.

HPV Examination

For HPV detection we used an AB I2Prism7000 PCR detector, a Hybrid-Max medical nucleic acid molecular hybridizer (AB, USA), and a PCR kit (Qiagen, USA). HPV detection used an AB I2Prism7000 PCR detector, a Hybrid-Max medical nucleic acid molecular hybridizer (AB, USA), and a PCR kit (Qiagen, USA). Fifteen types of the HR-HPV types (HPV16, 18, 31, 33, 34, 39, 45, 51, 52, 53, 56, 58, 59, 66, and 68) were detected and diagnosed by a pathologist with more than 10 years of diagnostic experience.

Hematology Diagnostic Criteria

Included in the novel systemic inflammation index were the following: NLR, defined as neutrophil-to-lymphocyte ratio; LMR, defined as mononuclear cell-to-lymphocyte ratio; and SII, defined as ratio of product of platelets and neutrophils to lymphocytes (25, 26). Evaluation was undertaken of the relationship between the novel systemic inflammation index and prognosis based on the ROC curve and calculation of the optimal cut-off value for meaningful indexes. Anemia was defined as a hemoglobin count below 11.0 g/dL. The optimal D-dimer, Fg and SCCA cut-off range for assessing prognosis was 0-0.5 mg/L, 2-4 g/L, and 0-2.5 ng/ml. These cut-off ranges were selected as they have been recognized as standard pathological definitions.

PET/CT Scanning and Image Acquisition

All patients were fasted for more than 6 hours before the 18F-FDG PET/CT examination, blood glucose levels were controlled below 7 mmol/L, and 18F-FDG was injected with the patient in a quiet state from 3.70 to 5.55 MBq/kg. After 45–60 minutes, 18F-FDG PET/CT was performed. The patient was placed in a GE Discovery Elite scanner (GE Healthcare, USA) with the scan ranging from the skull top to the middle of the thigh (120s/bed). The thickness of the CT scan layer was 3.75 mm, the tube voltage was 120–140 keV, and the tube current was 80mA. PET images were reconstructed using an Iterative adaptive algorithm.

Image Analysis

Transmitting all ^{18}F-FDG PET/CT images to GE AW4.6 workstation (GE Healthcare, USA), 2 radiologists with more than 5 years of radiodiagnostic experience used PET volume computer-assisted reading (PET VCAR) software to perform imaging analysis. After outline of a lesion region of interest (ROI) at the cross-sectional level of the largest area of the tumor, the fused-PET software automatically calculated $SUV_{mean,}$ SUV_{max}, MTV and TLG of the entire tumor. MTV was calculated by a 40% SUV_{max} threshold (27, 28).

Statistical Analysis

The patients were divided into a positive and a negative group according to prognosis. Differences in patient characteristics between the groups were compared with the t-test (Mann-Whitney U test, if not normal distribution) or chi-squared (Fisher's exact test, if not the number of assumptions necessary). For quantitative variables, receiver operating characteristic (ROC) curves were generated to assess the area under the ROC curve (AUC) for differentiating prognosis. We used the cut-off threshold values for differentiating these meaningful quantitative variables. The Kaplan-Meier method was used to build survival functions and the log-rank test was used to compare survivals for all variables. Prognostic variables were evaluated by univariate and multivariate Cox proportional hazard model (variables with $p<0.1$ in the univariate Cox proportional hazard model were used in the multivariate analysis). In multiple testing, a correction was performed, based on the Benjamini-Hochberg procedure. Finally, these final prognostic variables were incorporated to construct the

nomogram. The nomogram adopted the 1-, 2-, and 5-year PFS as primary endpoints. Validation of the nomogram was performed with 1000 bootstrapped samples, that is, we adopted the method of bootstrapped samples to set up a control group, randomly selected 1000 times from the existing samples, arbitrarily once, and then a final time to get the original sample number (95). The predictive capability was evaluated by ROC analysis. Generated calibration curves were used to visualize the difference between the predicted and actual 1-, 2-, and 5-year PFS. Decision curve analysis (DCA) was introduced to evaluate clinical utility of the nomogram (29). Based on the median of the total scores, a risk stratification system was developed, and the cervical cancer patients were divided into two risk subgroups, including high-risk and low-risk groups, and the Kaplan-Meier method and log rank test were used to compare the differences between the two subgroups. The statistical analysis was conducted by MedCalc (version 15.2.2), SPSS (version 22.0), R (version 3.5.3), and $p<0.05$ indicated a significant statistical difference (if not specified).

RESULTS

Survival and Disease Control

The median follow-up was 62.8 months (range, 2-96). A total of 95 patients were enrolled: 24 (25.3%) patients manifested progressive disease and 71 (74.7%) patients presented no evidence of disease (**Figure 1**).

The Relationship Between Disease Progression and Clinicopathological Characteristics in Patients

The clinicopathological characteristics of the patients are summarized in **Table 1A**. Cervical stromal invasion depth (chi-squared test, p=0.046; log-rank test, p=0.032), LVSI (chi-squared test, p=0.032; log-rank test, p=0.008) and HPV-16 (chi-squared test, p=0.001; log-rank test, p<0.001) showed differences between patients with and without disease progression. The remaining indicators had no significant correlation.

The Relationship Between Disease Progression and Protein Expression in Cancer Tissue

Chi-squared test and Kaplan-Meier method showed that patients with Glut1-negative had a statistically significant better clinical outcome than patients with Glut-1-positive (chi-squared test, p=0.039; log-rank test, p=0.039) (**Table 1B**). The remaining protein indicators were not statistically significant.

The Relationship Between Disease Progression and Hematology-Related Parameters in Patients

Chi-squared test and Kaplan-Meier method showed only D-dimer was related to disease progression and PFS (**Table 1C**). The ROC curve showed that no novel systemic inflammation index was statistically significant for disease progression.

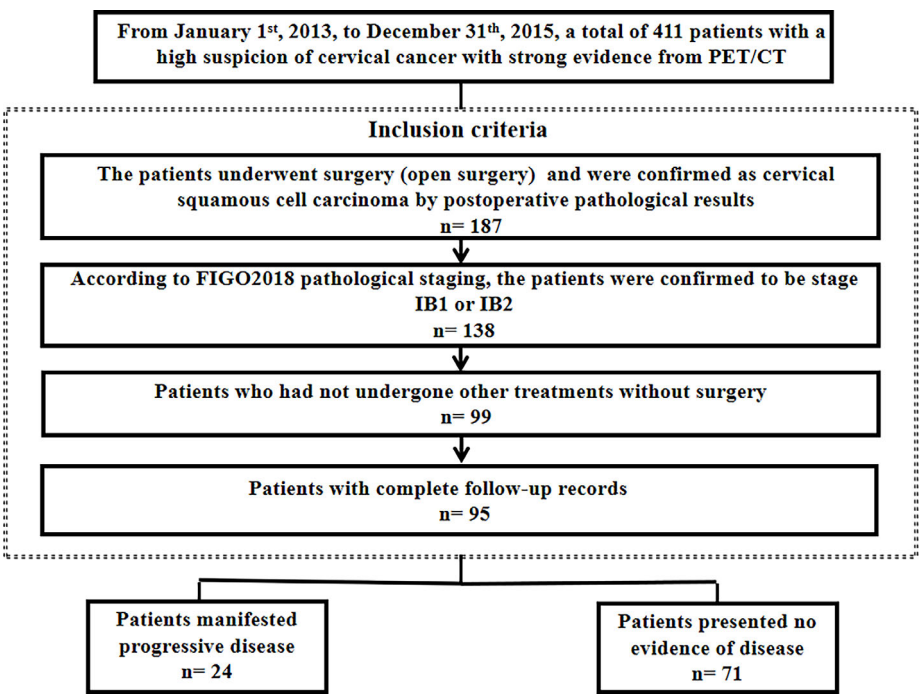


FIGURE 1 | Patient selection.

TABLE 1A | Patients' clinicopathological characteristics.

Feature	Total	Nun-progressor	Progressor	Value - p		
				chi-squared	log-rank	t-test/U-test
No. of patients	95	71	24			
Mean age (years)	49.40 ± 9.26	49.06 ± 9.14	50.42 ± 9.73	/	/	0.968
FIGO stage (2018):						
Ib1	11 (11.6%)	9	2	0.565	0.531	/
Ib2	84 (88.4%)	62	22			
Differentiation grade:						
Well-moderately differentiated	86 (90.5%)	64	22	0.825	0.748	/
Poorly differentiated	9 (9.5%)	7	2			
Cervical stromal invasion depth:						
< ½	36 (37.9%)	31	5	0.046	0.032	/
≥ ½	59 (62.1%)	40	19			
LVSI						
Negative	71 (74.7%)	57	14	0.032	0.008	/
Positive	24 (25.3%)	14	10			
HR-HPV:						
Negative	19 (20%)	7	12	0.194	0.176	/
Positive	76 (80%)	17	59			
HPV-16						
Negative	43 (45.3%)	25	18	0.001	<0.001	/
Positive	52 (54.7%)	46	6			
HPV-18						
Negative	88(45.3%)	67	21	0.266	0.226	/
Positive	7(45.3%)	4	3			

FIGO stage (2018): Postoperative pathological staging; LVSI, lymphovascular space invasion; HR-HPV, high-risk human papillomavirus; Normal distribution data, means ± standard deviations; t-test. Non-normal distribution data, medians and interquartile ranges; U-test. /, No statistics.

The Relationship Between Disease Progression and PET/CT Parameters in Patients

SUV_{max} (U test, p=0.014) and SUV_{peak} (U test, p=0.002) showed significant between-group differences (**Table 2**). ROC analysis showed that SUV_{max} (AUC 0.668, p=0.006), SUV_{peak} (AUC 0.716, p<0.001) had a positive effect on predicting disease progression (**Figure 2**). The optimal cut-off threshold values for SUV_{max}, SUV_{peak} were 9.12 g/cm^3 (sensitivity 91.67, specificity 39.44) and 7.63 g/cm^3 (sensitivity 91.67, specificity 47.89), respectively. Kaplan-Meier method showed the DFS rates of patients exhibiting high SUV_{max} and SUV_{peak} of the primary tumor were significantly lower than those of patients exhibiting low SUV_{max} and SUV_{peak} of the primary tumor (p=0.006 and p=0.001, respectively) (**Figure 2**).

Univariate and Multivariate Cox Proportional Hazard Model

Univariate Cox proportional hazard model showed that cervical stromal invasion depth (p=0.041), LVSI (p=0.018), HPV-16 (p=0.001), Glut1 (p=0.048), D-dimer (p=0.040), SUV_{max} (p=0.016) and SUV_{peak} (p=0.005) were associated with PFS. Multivariate Cox proportional hazard model showed that SUV_{peak} (p=0.012), and HPV-16 (p=0.007) were independent risk factors for PFS (**Table 3**).

Nomogram Development and Validation

SUV_{peak} and HPV-16 were incorporated to develop the nomogram for predicting 1-, 2-, and 5-year PFS. The nomogram showed that SUV_{peak} made the largest contribution to the prognosis, followed by HPV-16 which showed a certain amount of impact on the PFS (**Figure 3**). ROC analysis showed the AUCs at 1-, 2-, and 5-year PFS reached 0.828, 0.808 and 0.814 in the original model, and 0.814, 0.835 and 0.841 in the validation model, respectively (**Figures 4B, D**). Whether in the original model or in the validation model, the calibration curves demonstrated considerable agreement between the nomogram and predicted survival (**Figures 4A, C**). Clinical utility of the nomogram was evaluated by DCA. The nomogram showed enormous positive net benefits across wide ranges of mortality risk in both models, demonstrating its predominant clinical utility in predicting PFS (**Figure 5**). In particular, it has the greatest clinical utility for 5-year PFS in both models. In addition, we calculated the total scores for the original model to build a risk stratification system based on our nomogram, and then distinguished the patients according to the median quantile of total scores into high-risk subgroups and low-risk subgroups. The PFS in the two subgroups was exactly separated by this system (**Figures 6A, B**).

DISCUSSION

A poor prognosis for patients with cervical cancer is usually associated with a large tumor volume, lymph node metastasis, distant metastasis, and parauterine infiltration. In this study, we explored the prognostic indicators of patients with stage 1B1 and 1B2 cervical cancer without these factors, and established a prognostic nomogram and risk stratification system. Our study showed SUV_{max}, SUV_{peak}, Glut1, HPV-16, cervical stromal invasion depth, LVSI and D-dimer were associated with PFS. SUV_{peak} and HPV-16 were identified as factors independently impacting disease progression.

TABLE 1B | Patients' protein expression.

Feature	Total	Nun-progressor	Progressor	Value - p		
				chi-squared	log-rank	t-test/U-test
No. of patients	95	71	24			
Vegf-A:						
Negative	49 (51.6%)	37	12	0.858	0.837	/
Positive	46 (48.4%)	34	12			
Glut1:						
Negative	41 (43.2%)	35	6	0.039	0.039	/
Positive	54 (56.8%)	36	18			
E-cadherin:						
Negative	52 (54.7%)	40	12	0.590	0.487	/
Positive	43 (45.3%)	31	12			
Cox-2						
Negative	66 (69.5%)	48	18	0.496	0.547	/
Positive	29 (30.5%)	23	6			
Tenascin-C:						
Negative	50 (52.6%)	35	15	0.264	0.332	/
Positive	45 (47.4%)	36	9			
Tigar:						
Negative	30 (31.6%)	22	8	0.831	0.848	/
Positive	65 (68.4%)	49	16			

TABLE 1C | Patients' hematology-related parameters.

Feature	Total	Nun-progressor	Progressor	Value - p		
				chi-squared	log-rank	t-test/U-test
No. of patients	95	71	24			
Hb:						
≥ 110 g/dL	51 (53.7%)	41	10	0.172	0.180	/
<110 g/dL	44 (46.3%)	30	14			
D-Dimer:						
Negative	45 (47.4%)	38	7	0.039	0.033	/
Positive	50 (52.6%)	33	17			
Fg:						
Negative	57 (60.0%)	42	15	0.772	0.691	/
Positive	38 (40.0%)	29	9			
SCCA:						
Negative	58 (61.1%)	45	13	0.424	0.416	/
Positive	37 (38.9%)	26	11			
NLR:	1.92 (1.32, 3.00)	2.09 (1.38, 3.14)	1.54 (1.25, 2.46)	/	/	0.084
LMR:	0.20 (0.16, 0.27)	0.20 (0.17, 0.29)	0.18 (0.14, 0.27)	/	/	0.271
SII:	411.06 (298.67, 710.35)	416.20 (312.26, 710.35)	375.32 (293.99, 579.40)	/	/	0.333

Hb, Hemoglobin; Fg, fibrinogen, SCCA, squamous cell carcinoma antigen; NLR, neutrophil-to-lymphocyte ratio; LMR, mononuclear cell-to-lymphocyte ratio; SII, systemic immunity-inflammation index. Normal distribution data: means ± standard deviations; t-test. Non-normal distribution data: medians and interquartile ranges; U-test. /, No statistics.

TABLE 2 | Patients' PET/CT parameters.

	Nun-progressor	Progressor	Value - p
No. of patients	71	24	
SUV_{max} (g/cm^3)	10.87 (7.58, 13.66)	13.29 (10.57, 20.80)	0.014
SUV_{mean} (g/cm^3)	6.23 (4.17, 8.31)	7.78 (5.57, 10.15)	0.149
SUV_{peak} (g/cm^3)	8.27 (5.29, 11.47)	10.12 (8.84, 18.69)	0.002
MTV (cm^3)	10.20 ± 6.28	10.31 ± 6.50	0.876
TLG (g)	47.70 (32.72, 87.22)	54.10(38.30, 153.19)	0.014

Normal distribution data: means ± standard deviations; t-test.
Non-normal distribution data: medians and interquartile ranges; U-test.

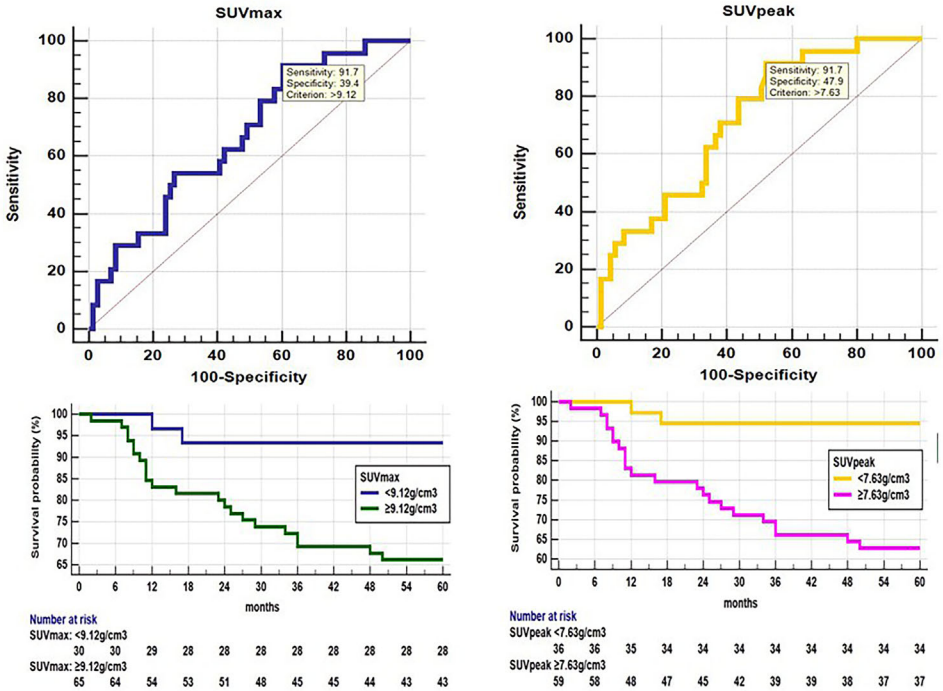


FIGURE 2 | ROC analysis shows that SUV_{max}(AUC 0.668, 95% confidence interval (CI) 0.564-0.761, p=0.006) and SUV_{peak} (AUC 0.716, 95% CI 0.615-0.804, p<0.001) had a positive effect on disease progression. The optimal cut-off threshold values for SUV_{max} and SUV_{peak} were 9.12 g/cm^3 (sensitivity 91.67, specificity 39.44), 7.63 g/cm^3 (sensitivity 91.67, specificity 47.89), respectively. Kaplan-Meier survival graph shows significantly different PFS between the groups categorized by SUVmax and SUVpeak above and below cut-off value (p<0.05, log-rank test).

TABLE 3 | Prognostic factors for DFS selected by Cox analysis.

Univariate Cox proportional hazard model: enter variable <0.1 **Method: Forward: LR** **Variable**	**Reference**	**Characteristic**	**P**	**HR**	**95% CI**
Cervical stromal invasion depth	<½	≥½	0.041	2.801	1.045-7.507
LVSI	negative	positive	0.018	2.661	1.181-5.998
HPV-16	positive	negative	0.001	4.651	1.844-11.736
Glut1	negative	positive	0.048	2.544	1.010-6.411
D-dimer	negative	positive	0.040	2.511	1.041- 6.056
SUV_{max}	<9.12 g/cm^3	≥9.12 g/cm^3	0.016	5.901	1.387-25.103
SUV_{peak}	<7.63 g/cm^3	≥7.63 g/cm^3	0.005	8.022	1.885-34.135
Multivariate Cox proportional hazard model: **Benjamini-Hochberg procedure was applied to the final analyses; enter variable <0.05** **Method: Forward: LR**					
Variable	**Reference**	**Characteristic**	**P**	**HR**	**95% CI**
SUV_{peak}	<7.63 g/cm^3	≥7.63 g/cm^3	0.012	8.342	1.956-35.574
HPV-16	positive	negative	0.007	4.834	1.909-12.239

LVSI, lymphovascular space invasion; HPV, human papillomavirus.

The basis of PET imaging is the "Warburg effect", that is, tumor cells mainly produce ATP through a higher rate of glycolysis compared with normal cells. SUV_{max}, a type of PET parameter, reflects the metabolism of the most active part of the lesion. SUV_{peak}, another type of SUV, refers to the average SUV within a spherical VOI positioned around the most active metabolism point (30). The glucose uptake of cells is mediated by Glut, and Glut-1 is an important isomer of glucose transporter. Univariate cox analysis shows that SUV_{max}, SUV_{peak} and Glut1 have a certain correlation with PFS, but multivariate analysis shows that only SUV_{peak} is considered to be an independent factor, among these three parameters. SUV_{max} is the most commonly used PET parameter in the clinic, but it is a single element measurement value and is susceptible to image resolution and noise. Compared with SUV_{max}, the value of SUV_{peak} is more stable and accurate because it is not easy affected by tracking bed position, scanning time and the size of the

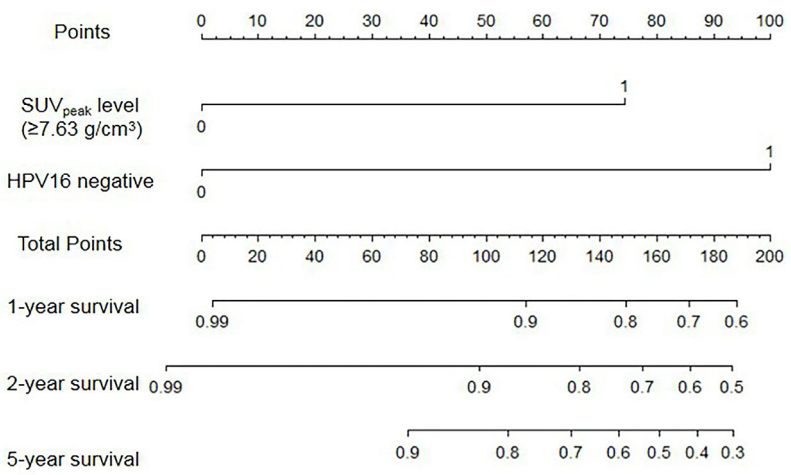


FIGURE 3 | Nomogram for predicting 1-, 2-, and 5-year PFS of cervical cancer.

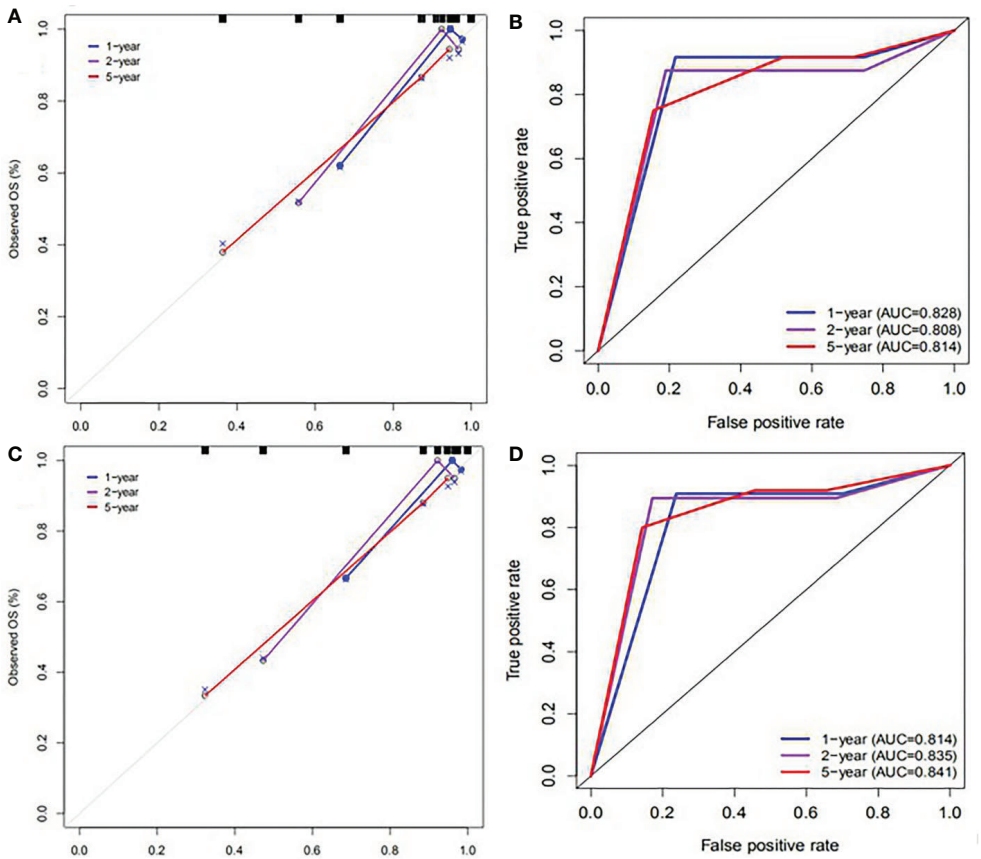


FIGURE 4 | Calibration plots and ROC curves for predicting PFS at 1-, 2-, and 5- year points. **(A)** The calibration plots for predicting PFS in the original model. **(B)** ROC curves of the nomogram for predicting PFS in the original model. **(C)** The calibration plots for predicting PFS in the validation model. **(D)** ROC curves of the nomogram for predicting PFS in the validation model.

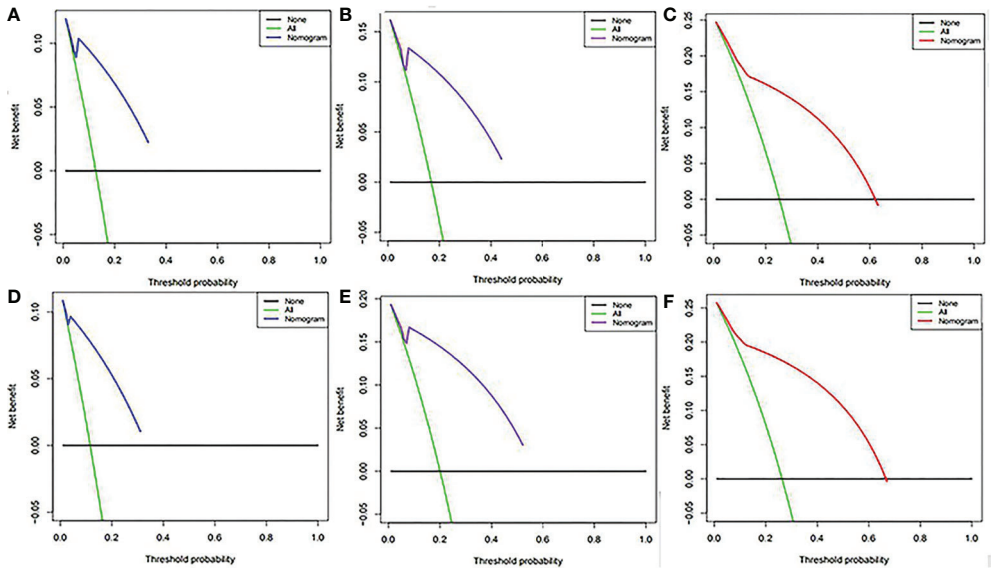


FIGURE 5 | Decision curve analysis of the nomogram for predicting PFS at 1-,**(A)** 2-,**(B)** and 5-year **(C)** points in the original model and PFS at 1-**(D)**, 2-**(E)** and 5-year **(F)** points in the validation model. The percentage of threshold probability was represented by the x-axis, whereas the net benefit was represented by the y-axis, calculated by adding the true positives and subtracting the false positives.

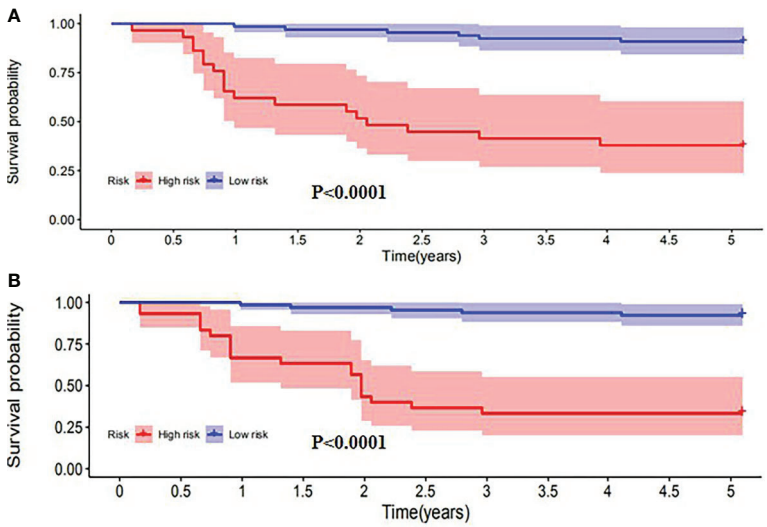


FIGURE 6 | Kaplan-Meier analysis of PFS for patients stratified by the risk stratification system in the original model **(A)** and validation model **(B)**.

lesion (31, 32). A study on the survival analysis of patients with cervical cancer showed that among the five classic PET parameters: SUV_{max}, SUV_{mean}, SUV_{peak}, MTV, TLG and six texture features, SUV_{peak} is the most accurate parameter in predicting the disease progression (33). Zhang Le et al. showed that SUV_{peak} had the highest correlation with the clinicopathological features of cervical cancer (34). These all support our results. Compared with other glucose metabolism parameters, SUV_{peak} is more closely related to the prognosis of patients undergoing surgery. It is worth noting that MTV, TLG and PFS are not related in our study. Yoo et al. conducted a prognostic study of cervical cancer at stage I–IV and showed that MTV and TLG are important PET parameters for predicting disease progression (35). The study of Maura et al. and Sangwon et al. had similar findings (36, 37). This may be because the largest diameter of the cancerous lesions in our study subjects was less than 4cm. However, MTV and TLG reflect the overall metabolic burden of the tumor. Therefore, the parameters MTV and TLG related to the tumor volume did not play a major role in our research.

HPV is a genus of papilloma vacuole virus A, belonging to the papovaviridae family. It is a spherical DNA virus that can cause the proliferation of squamous epithelium of human skin and mucosa. High-risk HPV can be detected in most cervical cancer specimens. HPV16 is a common high-risk type of HPV that is most prone to persistent infection and can be detected in about 40% to 60% of patients with cervical cancer (38). However, a study showed that HPV-negative status is associated with a poor prognosis in patients with cervical cancer, which may be related to WNT/β-catenin signaling and non-synonymous somatic mutations (39). Another Meta-Analysis showed the presence of HPV-16 positivity appears to have no significant association with prognosis cervical cancer in PFS. But, eliminating a study with a strong impact on the outcome from the analysis would lead to a conclusion of a worse prognosis of HPV-16 negative in cervical cancer (40). In addition, one study showed that cervical cancer patients infected with HPV16 had a better prognosis than those with any other HPV type (41). These results are roughly consistent with our study, in which HPV16 negativity was associated with worse PFS for those patients with cervical cancer.

Cervical stromal invasion depth is an essential index in the standard pathological report, and represents the invasion status of tumor cells. Lymphovascular space invasion (LVSI) is a common clinical pathological phenomenon of malignant tumors. For tumors to form distant metastases and spread, tumor cells must first enter the circulatory system and spread through blood or lymphatic vessels. Therefore, LVSI is associated with distant metastasis, from a histological perspective (42). D-dimer is a degradation product produced following fibrinolysis, which participates in the coagulation process. Its plasma level can be used as an evaluation index for blood hypercoagulability and as a measure of whether secondary fibrinolytic hyperfunction occurs (43). Our research showed that the above three indicators were related to prognosis, but they were not independent prognostic factors.

Our combined nomogram developed could effectively identify the patients with progressive disease. It is observed that SUV_{peak} is of the most value to prognosis, followed by HPV-16. We established a control model using 1000 bootstrapped samples at the same time. The nomogram of the original group and the control group both showed excellent performance, as was indicated by the ROC curves, generate calibration curves, DCA and Kaplan-Meier method. The ROC curves and Kaplan-Meier method showed that the model had good prediction accuracy, and the calibration curves demonstrated consistency. However, even if the model has high accuracy, patients may not necessarily benefit clinically, and there may be false positives and false negatives. Sometimes it is more beneficial to avoid false positives, and at other times more desirable to avoid false negatives. Since neither situation can be avoided, a method with the greatest net benefit is needed. This is what DCA does. The DCA of our model showed great clinical performance, especially for patients with 5 year-PFS. The capability of the combined nomogram for prediction of progressive disease may facilitate personalized treatment decisions.

This study has a number of limitations. On one hand, this is a retrospective study based on the latest FIGO2018 staging. Because the early follow-up was not optimal, the sample size of our study was not sufficiently large, especially for positive cases, thus further future confirmation of the results is needed. Furthermore, there is a lack of data about several important factors, including some functional sequences of magnetic resonance, proteomics and genomics data, and pathway proteins related to prognosis, such as Hive, Caix, etc. Finally, in our original data, compared with patients with squamous cell carcinoma, the proportion of patients with adenocarcinoma is particularly small. In order to prevent extreme value bias, we only selected patients with squamous cell carcinoma as the research object. In the future, additional types of pathology will be included in our research.

CONCLUSION

In conclusion, SUV_{peak} level (≥7.63 g/cm^3) and HPV16 negative were identified as independent factors and could be associated with poor prognosis for patients with cervical cancer (Stage IB1, IB2). Based on this result, we established a nomogram and risk stratification system, and achieved satisfactory performance and clinical utility. These findings could contribute to test-individualized neoadjuvant treatment. It is worth emphasizing that the stage of cervical cancer of patients selected was only Stage IB1 and IB2. That is, patients with cervical cancer of other stages were not within the scope of this study. At the same time, cervical stromal invasion depth, LVSI, and D-dimer were related to prognosis, but they were not associated with poor DFS in the multivariate analysis. Cautions would be needed when clinical decision-making was made for patients with these risk factors.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

CX and SG designed the research. CX, TM, XL and SG performed the research and analyzed results. CX wrote the paper. CX, XL, HS, TM and SG edited the manuscript and provided critical comments. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by LIAONING Science & Technology Project (2017225012), LIAONING Science Natural Science Foundation (2019-MS-373) and 345 Talent Project.

ACKNOWLEDGMENTS

We would like to thank our researchers for their hard work.

REFERENCES

1. Tsu V, Jerónimo J. Saving the World's Women From Cervical Cancer. *New Engl J Med* (2016) 374:2509–11. doi: 10.1056/NEJMp1604113
2. Signorelli M, Guerra L, Montanelli L, Crivellaro C, Buda A, Dell"Anna T, et al. Preoperative Staging of Cervical Cancer: Is 18-FDG-PET/CT Really Effective in Patients With Early Stage Disease? *Gynecol Oncol* (2011) 123:236–40. doi: 10.1016/j.ygyno.2011.07.096
3. Susanna L, Mostafa A. 2018 FIGO Staging System for Uterine Cervical Cancer: Enter Cross-sectional Imaging. *Radiology* (2019) 292:15–24. doi: 10.1148/radiol.2019190088
4. Koji M, Hiroko M, Rachel SM, Ikuo K, Mikio M. Validation of the 2018 FIGO Cervical Cancer Staging System. *Gynecol Oncol* (2019) 53:87–93. doi: 10.1016/j.ygyno.2018.10.026
5. Ainhoa A, Stanley IG, Ralph RW. Radiotherapy and Immunotherapy for Cancer: From "Systemic" to "Multisite". *Clin Cancer Res* (2020) 26:2777–82. doi: 10.2147/CMAR.S239624
6. Pfaendler KS, Tewari KS. Changing Paradigms in the Systemic Treatment of Advanced Cervical Cancer. *Am J Obstet Gynecol* (2016) 214:22–30. doi: 10.1016/j.ajog.2015.07.022
7. Fendt SM, Lunt SY. Dynamic ROS Regulation by TIGAR: Balancing Anti-cancer and Pro-metastasis Effects. *Cancer Cell* (2020) 37:141–2. doi: 10.1016/j.ccell.2020.01.009
8. Iwasaki K, Yabushita H, Ueno T, Wakatsuki A. Role of Hypoxia-Inducible Factor-1α, Carbonic anhydrase-IX, Glucose Transporter-1 and Vascular Endothelial Growth Factor Associated With Lymph Node Metastasis and Recurrence in Patients With Locally Advanced Cervical Cancer. *Oncol Lett* (2015) 10:1970–78. doi: 10.3892/ol.2015.3524
9. Peng J, Qi S, Wang P, Li W, Song L, Liu C, et al. Meta-Analysis of Downregulated E-cadherin as a Poor Prognostic Biomarker for Cervical Cancer. *Future Oncol* (2016) 12:715–26. doi: 10.2217/fon.15.332
10. Aylas M, Pérez -Regadera GJ. EP-2061: Over-expression of EGFR and/or Cox-2 in Locally Advanced Squamous Cervical Cancer (LASC). *Radiotherapy Oncol* (2016) 119:S972–3. doi: 10.1016/S0167-8140(16)33312-6
11. Huseyin B, Adem A. Value of Tenascin-C Content and Association With Clinicopathological Parameters in Uterine Cervical Lesions. *Int J Cancer* (2002) 100:719–22. doi: 10.1002/ijc.10546
12. Lucia F, Miranda O, Abgral R, Bourbonne V, Dissaux G, Pradier O, et al. Use of Baseline 18 F-FDG PET/CT to Identify Initial Sub-Volumes Associated With Local Failure After Concomitant Chemoradiotherapy in Locally Advanced Cervical Cancer. *Front Oncol* (2020) 10:678. doi: 10.3389/fonc.2020.00678
13. Giacomo ML, Antonella M, Giulio V, Giulia D, Nicoletta N, Eugenia MDC, et al. Prognostic Value of Posttreatment 18F-FDG PET/CT and Predictors of Metabolic Response to Therapy in Patients With Locally Advanced Cervical Cancer Treated With Concomitant Chemoradiation Therapy: An Analysis of Intensity- and Volume-Based PET Parameters. *Eur J Nucl Med Mol Imaging* (2018) 45:2139–46. doi: 10.1007/s00259-018-4077-1
14. Herrera FG, Prior JO. The Role of PET/CT in Cervical Cancer. *Front Oncol* (2013) 3:34. doi: 10.3389/fonc.2013.00034
15. Zsuzsanna N, Orsolya H, Julia K, Dorottya V, Larisza L, Bence K, et al. D-Dimer as a Potential Prognostic Marker. *Pathol Oncol Res* (2012) 18:669–74. doi: 10.1007/s12253-011-9493-5
16. Moreno-Acosta P, Vallard A, Carrillo S, Gamboa O, Romero-Rojas A, Molano M, et al. Biomarkers of Resistance to Radiation Therapy: A Prospective Study in Cervical Carcinoma. *Radiat Oncol* (2017) 12:120. doi: 10.1186/s13014-017-0856-2
17. Liu Z, Shi H. Prognostic Role of Squamous Cell Carcinoma Antigen in Cervical Cancer: A Meta-Analysis. *Dis Markers* (2019) 1:1–10. doi: 10.1155/2019/6710352
18. Sun D, An L, Lv G. Albumin-Fibrinogen Ratio and Fibrinogen-Prealbumin Ratio as Promising Prognostic Markers for Cancers: An Updated Meta-Analysis. *World J Surg Oncol* (2020) 18:9. doi: 10.1186/s12957-020-1786-2
19. Chao B, Ju X, Zhang L, Xu X, Zhao Y. A Novel Prognostic Marker Systemic Inflammation Response Index (SIRI) for Operable Cervical Cancer Patients. *Front Oncol* (2020) 10:766. doi: 10.3389/fonc.2020.00766
20. Ramirez PT, Frumovitz M, Pareja R, Lopez A, Vieira M, Ribeiro R, et al. Minimally Invasive Versus Abdominal Radical Hysterectomy for Cervical Cancer. *N Engl J Med* (2018) 379:1895–904. doi: 10.1056/NEJMoa1806395
21. Gallotta V, Conte C, Federico A, Vizzielli G, Gueli AS, Tortorella L, et al. Robotic Versus Laparoscopic Radical Hysterectomy in Early Cervical Cancer: A Case Matched Control Study. *Eur J Surg Oncol* (2018) 44:754–9. doi: 10.1016/j.ejso.2018.01.092
22. Pedone AL, Bizzarri N, Kucukmetin A, Turco LC, Gallotta V, Carbone V, et al. Investigating the Possible Impact of Peritoneal Tumor Exposure Amongst Women With Early Stage Cervical Cancer Treated With Minimally Invasive Approach. *Eur J Surg Oncol* (2021) 47(5):1090–7. doi: 10.1016/j.ejso.2020.09.038. S0748-7983(20)30829-5.
23. Xu C, Du S, Zhang S, Wang B, Dong C, Sun H. Value of Integrated PET-IVIM MR in Assessing Metastases in Hypermetabolic Pelvic Lymph Nodes in Cervical Cancer: A Multi-Parameter Study. *Eur Radiol* (2020) 30:2483–92. doi: 10.1007/s00330-019-06611-z
24. Xu C, Yu Y, Li X, Sun H. Value of Integrated PET-IVIM MRI in Predicting Lymphovascular Space Invasion in Cervical Cancer Without Lymphatic Metastasis. *Eur J Nucl Med Mol Imaging* (2021). doi: 10.1007/s00259-021-05208-3
25. Wang C, He W, Yuan Y, Zhang Y, Li K, Zou R, et al. Comparison of the Prognostic Value of Inflammation-Based Scores in Early Recurrent Hepatocellular Carcinoma After Hepatectomy. *Liver Int* (2020) 40:229–39. doi: 10.1111/liv.14281
26. Hu B, Yang XR, Xu Y, Sun Y, Sun C, Guo Y, et al. Systemic Immune-Inflammation Index Predicts Prognosis of Patients After Curative Resection for Hepatocellular Carcinoma. *Clin Cancer Res* (2014) 20:6212–22. doi: 10.1158/1078-0432.CCR-14-0442
27. Kidd EA, Thomas M, Siegel BA, Dehdashti F, Grigsby PW. Changes in Cervical Cancer FDG Uptake During Chemoradiation and Association With Response. *Int J Radiat Oncol Biol Phys* (2013) 85:116–22. doi: 10.1016/j.ijrobp.2012.02.056
28. Akkas BE, Demirel BB, Dizman A, Vural GU. Do Clinical Characteristics and Metabolic Markers Detected on Positron Emission Tomography/Computerized Tomography Associate With Persistent Disease in Patients With in-Operable Cervical Cancer? *Ann Nucl Med* (2013) 27:756–63. doi: 10.1007/s12149-013-0745-1
29. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26:565–74. doi: 10.1177/0272989X06295361
30. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. *J Nucl Med* (2009) 50:122S–50S. doi: 10.2967/jnumed.108.057307
31. Sher A, Lacoeuille F, Fosse P, Laurent V, Aurélie CV, Djamel D, et al. For Avid Glucose Tumors, the SUV_{peak} is the Most Reliable Parameter for((18)F)FDG-PET/CT Quantification, Regardless of Acquisition Time. *EJNMMI Res* (2016) 6:21. doi: 10.1186/s13550-016-0177-8
32. Akamatsu G, Ikari Y, Nishida H, Tomoyuki N, Akihito O, Akira M, et al. Influence of Statistical Fluctuation on Reproducibility and Accuracy of SUVmax and SUVpeak: A Phantom Study. *J Nucl Med Technol* (2015) 43:222–26. doi: 10.2967/jnmt.115.161745
33. Schernberg A, Reuze S, Orlhac F, Irène B, Laurent D, Roger S, et al. A Score Combining Baseline Neutrophilia and Primary Tumor SUV(peak) Measured From FDG PET is Associated With Outcome in Locally Advanced Cervical Cancer. *Eur J Nucl Med Mol Imaging* (2018) 45:187–95. doi: 10.1007/s00259-017-3824-z
34. Zhang L, Sun H, Du S, Xu W, Xin J, Guo Q. Evaluation of 18F-FDG PET/CT Parameters for Reflection of Aggressiveness and Prediction of Prognosis in Early-Stage Cervical Cancer. *Nucl Med Commun* (2018) 9:1045–52. doi: 10.1097/MNM.0000000000000909
35. Yoo J, Choi JY, Moon SH, Bae DS, Park SB, Choe YS, et al. Prognostic Significance of Volume-Based Metabolic Parameters in Uterine Cervical Cancer Determined Using 18F-Fluorodeoxyglucose Positron Emission Tomography. *Int J Gynecol Cancer* (2012) 22:1226–33. doi: 10.1097/IGC.0b013e318260a905
36. Miccò M, Vargas HA, Burger IA, Marisa AK, Debra AG, Kay JP, et al. Combined Pre-Treatment MRI and 18F-FDG PET/CT Parameters as Prognostic Biomarkers in Patients With Cervical Cancer. *Eur J Radiol* (2014) 83:1169–76. doi: 10.1016/j.ejrad.2014.03.024
37. Han S, Kim H, Kim YJ, Suh CH, Woo S. Prognostic Value of Volume-Based Metabolic Parameters of 18F-FDG PET/CT in Ovarian Cancer: A Systematic Review and Meta-Analysis. *Ann Nucl Med* (2018) 32:669–77. doi: 10.2214/AJR.18.19734

38. Muñoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, et al. Epidemiologic Classification of Human Papillomavirus Types Associated With Cervical Cancer. *N Engl J Med* (2003) 348:518–27. doi: 10.1056/NEJMoa021641
39. Banister CE, Liu C, Pirisi L, Creek KE, Buckhaults PJ. Identification and Characterization of HPV-independent Cervical Cancers. *Oncotarget* (2017) 8:13375–86. doi: 10.18632/oncotarget.14533
40. Chen X, Zhang P, CHen S, Zhu H, Wang K, Ye L, et al. Better or Worse? The Independent Prognostic Role of HPV-16 or HPV-18 Positivity in Patients With Cervical Cancer: A Meta-Analysis and Systematic Review. *Front Oncol* (2020) 10:1733. doi: 10.3389/fonc.2020.01733
41. Hang D, Jia M, Ma H, Zhou J, Feng X, Lyu Z, et al. Independent Prognostic Role Ofhuman Papillomavirus Genotype in Cervical Cancer. *BMC Infect Dis* (2017) 17:391. doi: 10.1186/s12879-017-2465-y
42. Tjalling B, Elke EMP, Carien LC, Ina MJS, Jan JJ, Jan WMM, et al. Substantial Lymph-Vascular Space Invasion (LVSI) is a Significant Risk Factor for Recurrence. *Eur J Cancer* (2015) 51:1742–50. doi: 10.1016/j.ejca.2015.05.015
43. John DO. D-Dimer: An Overview of Hemostasis and Fibrinolysis, Assays, and Clinical Applications. *Adv Clin Chem* (2015) 69:1–46. doi: 10.1016/bs.acc.2014.12.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 13 July 2021
doi: 10.3389/fonc.2021.638262

Development of a Machine Learning Classifier Based on Radiomic Features Extracted From Post-Contrast 3D T1-Weighted MR Images to Distinguish Glioblastoma From Solitary Brain Metastasis

Alix de Causans[1,2,3†], Alexandre Carré[4,5†], Alexandre Roux[2,3,6], Arnault Tauziède-Espariat[2,3,7], Samy Ammari[8,9], Edouard Dezamis[2,3,6], Frederic Dhermain[4,5], Sylvain Reuzé[4,5], Eric Deutsch[4,5], Catherine Oppenheim[1,2,3], Pascale Varlet[2], Johan Pallud[2,3,6], Myriam Edjlali[1,2,3‡] and Charlotte Robert[4,5‡]*

[1] Neuroradiology Department, Hôpital Sainte-Anne, GHU-Paris Psychiatrie et Neurosciences, Paris, France, [2] Université de Paris, Paris, France, [3] Inserm, UMR1266, IMA-Brain, Institut de Psychiatrie et Neurosciences, Paris, France, [4] Radiothérapie Moléculaire et Innovation Thérapeutique, INSERM UMR1030, Gustave Roussy Cancer Campus, Université Paris Saclay, Villejuif, France, [5] Département de Radiothérapie, Gustave Roussy, Université Paris Saclay, Villejuif, France, [6] Service de Neurochirurgie, GHU Paris – Psychiatrie et Neurosciences – Hôpital Sainte-Anne, Paris, France, [7] Service de Neuropathologie, GHU Paris – Psychiatrie et Neurosciences – Hôpital Sainte-Anne, Paris, France, [8] Département de Radiologie, Gustave Roussy, Université Paris Saclay, Villejuif, France, [9] BioMaps UMR1281, Université Paris-Saclay, CNRS, INSERM, CEA, Orsay, France

Objectives: To differentiate Glioblastomas (GBM) and Brain Metastases (BM) using a radiomic features-based Machine Learning (ML) classifier trained from post-contrast three-dimensional T1-weighted (post-contrast 3DT1) MR imaging, and compare its performance in medical diagnosis *versus* human experts, on a testing cohort.

Methods: We enrolled 143 patients (71 GBM and 72 BM) in a retrospective bicentric study from January 2010 to May 2019 to train the classifier. Post-contrast 3DT1 MR images were performed on a 3-Tesla MR unit and 100 radiomic features were extracted. Selection and optimization of the Machine Learning (ML) classifier was performed using a nested cross-validation. Sensitivity, specificity, balanced accuracy, and area under the receiver operating characteristic curve (AUC) were calculated as performance metrics. The model final performance was cross-validated, then evaluated on a test set of 37 patients, and compared to human blind reading using a McNemar's test.

Results: The ML classifier had a mean [95% confidence interval] sensitivity of 85% [77; 94], a specificity of 87% [78; 97], a balanced accuracy of 86% [80; 92], and an AUC of 92% [87; 97] with cross-validation. Sensitivity, specificity, balanced accuracy and AUC were equal to 75, 86, 80 and 85% on the test set. Sphericity 3D radiomic index highlighted the highest coefficient in the logistic regression model. There were no statistical significant

OPEN ACCESS

Edited by:
Sebastian Cerdan, Autonomous University of Madrid, Spain

Reviewed by:
Sung Soo Ahn, Yonsei University Health System, South Korea
Isaac Shiri, Geneva University Hospitals (HUG), Switzerland

***Correspondence:**
Charlotte Robert
ch.robert@gustaveroussy.fr

[†]These authors have contributed equally to this work and share first authorship

[‡]These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *05 December 2020*
Accepted: *17 June 2021*
Published: *13 July 2021*

Citation:
de Causans A, Carré A, Roux A, Tauziède-Espariat A, Ammari S, Dezamis E, Dhermain F, Reuzé S, Deutsch E, Oppenheim C, Varlet P, Pallud J, Edjlali M and Robert C (2021) Development of a Machine Learning Classifier Based on Radiomic Features Extracted From Post-Contrast 3D T1-Weighted MR Images to Distinguish Glioblastoma From Solitary Brain Metastasis. Front. Oncol. 11:638262. doi: 10.3389/fonc.2021.638262

differences observed between the performance of the classifier and the experts' blinded examination.

Conclusions: The proposed diagnostic support system based on radiomic features extracted from post-contrast 3DT1 MR images helps in differentiating solitary BM from GBM with high diagnosis performance and generalizability.

Keywords: radiomics, machine learning, glioblastoma, brain metastasis, diagnostic decision support system

INTRODUCTION

Brain Metastases (BM) and Glioblastomas (GBM) are the two most frequent intra-cranial brain tumors in adults (1–3). Currently, Magnetic Resonance Imaging (MRI) is the modality of choice for brain tumor characterization. Usually, BM present an encapsulated contrast enhancement, with regular and well-defined boundaries, whereas GBM have heterogeneous contrast enhancement with very irregular and fuzzy boundaries (4–6). Nonetheless, their morphological characteristics remain very similar on MRI as both are lesions with annular contrast enhancement, having a necrotic center and a peritumoral zone in T2-weighted and Fluid-Attenuated Inversion Recovery (FLAIR) sequences. Advanced neuroimaging techniques such as perfusion MRI and Magnetic Resonance Spectroscopy (MRS) provide additional information to distinguish between the two tumor types, based on differences in the peritumoral area (7–10). Although in the past decades, various studies (11–13) have evaluated the diagnostic performance of perfusion imaging and MRS, they have shown heterogeneous results in distinguishing these two tumor types, resulting in sensitivities and specificities ranging from 64 to 100% and 60 to 100% respectively. This high heterogeneity reflects the difficulty experienced in daily practice to differentiate the two brain tumors, even using advanced neuroimaging techniques, particularly in the case of differentiating a GBM from a solitary BM revealing an unknown primary cancer [5 to 12% of BM (14, 15)]. Even though the final diagnostic will be given by a histopathological examination and a biomolecular analysis of the tumor tissue relying on the 2016 WHO classification (16), the presurgical distinction between these two types of tumors is crucial for adapting treatment strategies: for metastases less than 3–4 cm, a bloc resection or stereotactic radiosurgery will be planned depending on the lesion location (17), while GBM (18) should be treated with maximal safe resection, and concurrent chemoradiotherapy. Radiomics (19–22) is a recent area of research based on the simple observation that the human eyes have limitations, even those trained for medical image interpretation. Radiomics consists of extracting large numbers of predefined quantitative features from medical images with the ultimate goal of identifying subgroups of biomarkers able to guide patient's care and has shown promise in brain cancer detection, diagnosis, molecular mutation characterization, prognosis and outcome prediction (23–29). In our study, we hypothesized that the morphological differences observed on post-contrast 3DT1 MR images would lead to differences in radiomic features between the two tumor types. The aim of this study was to therefore develop a radiomic features-based Machine Learning (ML) classifier, to evaluate its diagnostic performance on an unseen test set of patients, and to compare it to the diagnosis performance of neuroradiologists. A strong emphasis was placed on favoring explainable classifiers to ease translation into clinic.

MATERIALS AND METHODS

The steps of our study are summarized in **Figure 1**.

Patients

This retrospective bicentric study was approved by the local institutional review board (n° IRB00011687 College de neurochirurgie IRB #1: 2020/29). The two Radiology Departments that participated in the study had the same 3 Tesla MRI scanners (MR 750, Discovery; General Electric Healthcare), with the same imaging parameters implemented. Medical records of patients who had histologically proven BM or GBM between January 2010 and May 2019 were screened in the two centers to constitute the training set. Inclusion criteria for the training set were: 1) patients more than 18 years of age, 2) with histologically-confirmed diagnosis of BM or GBM, and 3) and with pre-operative MRI. Exclusion criteria for the training set were: 1) lesions less than 2 cm, 2) extra-axial locations, 3) history of treatment before the MRI examination, 4) absence of 3D T1-weighted Fast SPoiled Gradient Recalled sequence, 5) image acquisition performed on a different machine to the 3 Tesla GE Discovery MR scanner, and 6) 3D T1-weighted sequence acquired with non-conventional parameters or inadequate quality (see section *MRI data*). The minimal size of 2 cm was chosen as GBM are usually >2 cm at the diagnosis. We therefore wanted to exclude small BM from the analysis, to avoid a bias of size. For BM, we included patients with one or more brain lesions. However in cases of multiple lesions, only the largest was segmented for radiomic feature extraction.

Secondly, a test set was constituted after completion of the model development process in order to evaluate the final performance of the radiomic classifier on unseen lesions. As well, the test set included patients from both centers. Inclusion criteria for the test set were the same as for the training set. All patients included in the test set were required to have solitary lesions so that neuroradiologists were not influenced in their final diagnosis. Exclusion criteria of the study were therefore the same as those of the training set plus patients having multifocal or infra-tentorial lesions. All inclusion and exclusion criteria are summarized in the flow chart (**Figure 2**).

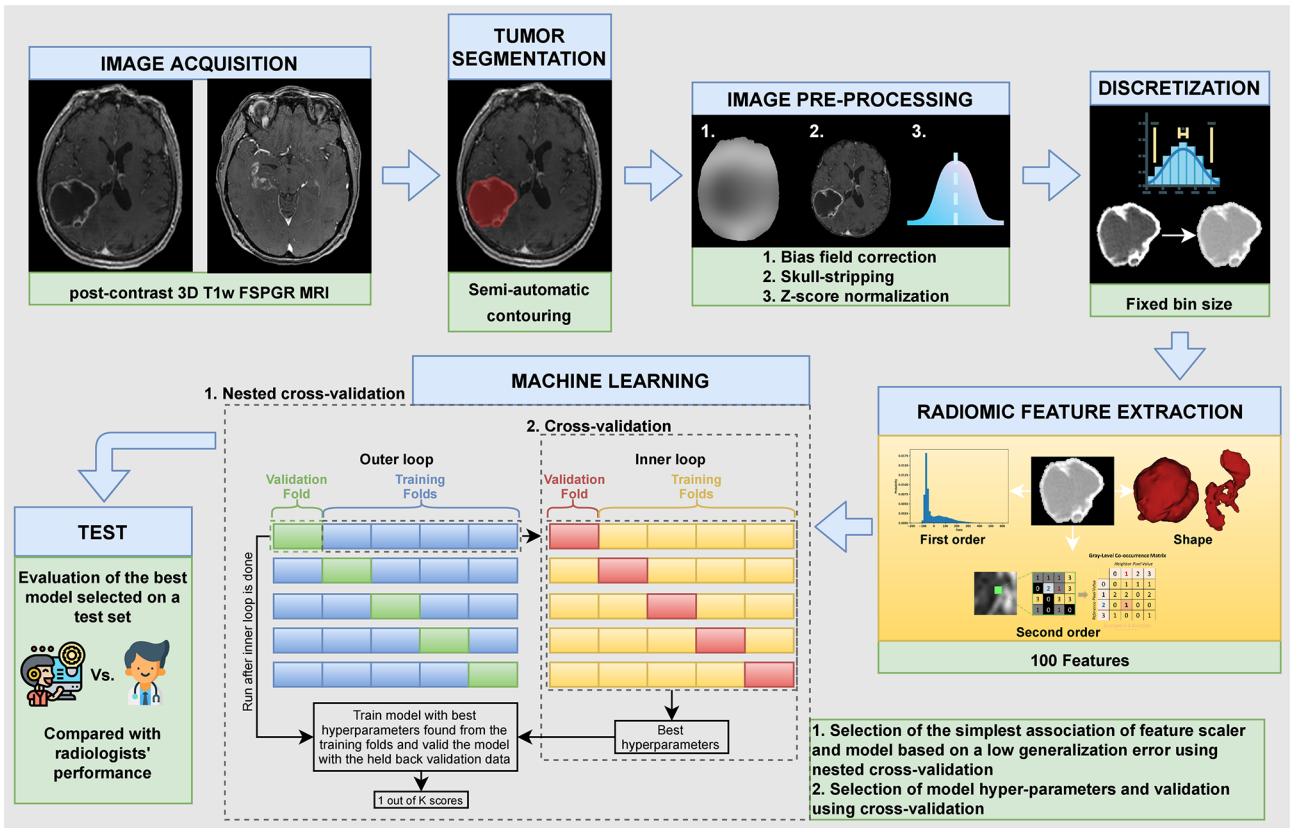


FIGURE 1 | Different steps of the study.

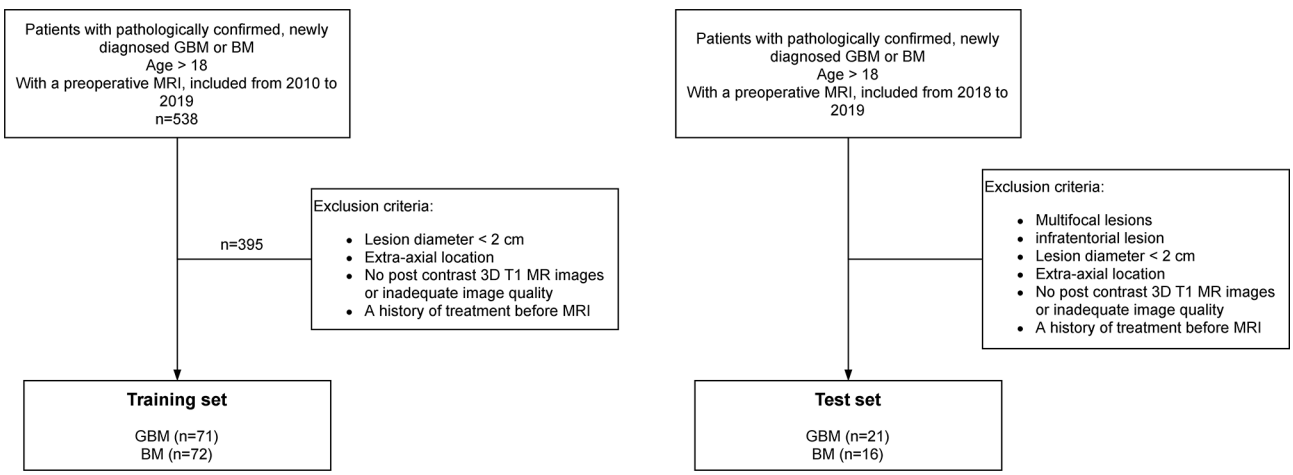


FIGURE 2 | Flow chart of patient inclusion.

MRI Data

MR acquisitions were performed on the same 3 Tesla MR scanner, even if at two clinical sites. MRI data included at least a post-contrast (gadoterate meglumine [Dotarem; Guerbet Laboratory]) three-dimensional T1-weighed Fast SPoiled Gradient Recalled (FSPGR) acquisition (post-contrast 3DT1), with the following parameters: repetition time: 10.2 ms; echo time: 3.4 ms; field of view: 22 cm; voxel size: 0.8 mm × 0.8 mm ×

1.2 mm. Patients were excluded from this study if other imaging protocols were followed. Post-contrast 3DT1 MR images were only used as inputs of the radiomic classifier. To compare the performance between the classifier and neuroradiologists, clinical conditions were mimicked, and all available sequences of the imaging exam were thus analyzed by the neuroradiologists, as routinely conducted in a clinical setting.

Image Analysis

Pre-Processing

MR image preprocessing included bias field correction using the N4ITK algorithm (30) from the Advanced Normalization Tools (ANTs) library (31), skull-stripping with the Brain Extraction Tool (BET) of the FSL software (FMRIB's Software Library) (32) and Z-score normalization with a scaling factor of 100. No spatial resampling was performed due to data homogeneity. As well, no noise filtering was applied.

Tumor Segmentation

A segmentation of the volume of interest, including the contrast-enhanced and necrotic regions, was performed semi-automatically using Olea Sphere© (Olea Medical, La Ciotat, France). These two sub-regions corresponded to Labels 4 and 1 of the BraTS 2012–2016 challenge (33). Within a region of interest defined by a trained radiologist (AdC, 5 years of experience), threshold-based gray level contouring and manual correction were used for the segmentation so that the volume of interest was carefully drawn along the tumor enhancement.

Feature Extraction

One hundred radiomic features were extracted from the 3D MR images using the Python library PyRadiomics 2.1.2 (34) in which the feature definitions are consistent with the Image Biomarker Standardization Initiative (IBSI) (35). The only exception is that PyRadiomics and IBSI use different definitions of the Kurtosis first-order feature, where Kurtosis is calculated using −3 and +3 in the IBSI and PyRadiomics referentials respectively. For first order features, an intensity shifting of 300 (equal to three standard deviations) was applied to ensure that the majority of the voxel intensities were positive before feature extraction. An absolute discretization with a fixed bin size equal to 37 was chosen (36, 37). This leads to a bin number of 32 considering the mean of the intensity intervals computed for all volumes of interest of patients of the training set (min intensity: 575, max intensity: 2069, mean intensity range: 1190). Six feature classes were considered: 18 first-order statistics, 14 shape-based features, 22 Gray Level Co-occurrence Matrix features (GLCM), 16 Gray Level Run Length Matrix features (GLRLM), 16 Gray Level Size Zone Matrix features (GLSZM), and 14 Gray Level Dependence Matrix features (GLDM).

Model Building

The establishment of the classification model was based on the scikit-learn library version 0.23.2 (38) and included two steps applied to the training set: (1) selection of the ML classifier and feature scaling method and 2) optimization of the hyper-parameters. In step 1), a nested cross-validation was used given the moderately-sized dataset and 144 ML models combining nine feature scaling methods (No Scaler, MaxAbsScaler, MinMaxScaler, Normalizer, PowerTransformer-Yeo–Johnson, QuantileTransformer-normal, QuantileTransformer-uniform, RobustScaler, StandardScaler) and 16 classifiers (AdaBoostClassifier, BaggingClassifier, BernoulliNB, DecisionTreeClassifier, Extra TreeClassifier, ExtraTreesClassifier, GaussianNB, Gradient BoostingClassifier, KNeighborsClassifier, LinearSVC, Logistic Regression, MLPClassifier, QuadraticDiscriminantAnalysis, RandomForestClassifier, RidgeClassifier, SGDClassifier) were compared. The nested cross-validation considered a stratified 5-fold cross-validation in the inner loop for hyper-parameter tuning (grid search strategy) and a stratified 5-fold cross-validation in the outer loop for the evaluation of the performance of the model. In step 2), the model showing the lowest generalization error, as assessed by the balanced accuracy, was kept and a ten-repeated 5-fold cross-validation was performed. In this second step, a grid search method was implemented to optimize the final set of hyper-parameters. Mean sensitivity, specificity, balanced accuracy, and area under the receiver operating characteristic curve (AUC) and their associated variances and 95% confidence intervals were calculated as performance metrics. Research spaces for hyper-parameter tuning with grid search during nested cross-validation and cross-validation are described in **Table S1**.

Evaluation on the Test Set and Comparison to Human Performance

The final model was fitted using the entire training set and its performance evaluated on the test set including 37 patients (21 GBM and 16 BM). Images of the test set were then blindly analyzed by five neuroradiologists (R1, R2, R3, R4, and R5). Two were neuroradiologists with more than 10 years of experience and three were radiology residents with about 6 months of training and practice in neuroradiology. The neuroradiologists had access to all MR sequences acquired in a routine MR imaging protocol, including 3D FLAIR, 2D T2, perfusion imaging, and pre- and post-contrast 3DT1 sequences.

Statistics

Sensitivity, specificity, balanced accuracy and AUC were used to assess the diagnosis performance of the radiomic model. We applied a McNemar's test and evaluated its p-value to assess if the differences were significant between the diagnostic performance of the radiomic classifier and the diagnostic performance of the readers. The threshold was set at 0.05.

RESULTS

Patients

267 GBM and 271 BM were pre-selected for the training set, and 71 GBM and 72 BM met the inclusion criteria respectively (**Figure 2**). Median [minimum value–maximum value] 2D maximal diameter was equal to 53.39 mm [24.11–88.12 mm] for GBM and 41.40 mm [20.77–77.92 mm] for BM. The test set included 37 patients (21 GBM and 16 BM). In this set, the

median 2D maximal diameter was equal to 54.93 mm [32.61–102.53 mm] and 33.85 mm [22.41–63.63 mm] for GBM and BM respectively. Patient characteristics and their repartition between Centers 1 and 2 are summarized in **Table 1**.

Selected Machine Learning Classifier

Table S2 summarizes the mean balanced accuracies and their associated standard deviations obtained for all tested combinations (scaling method + classifier). Combinations are ranked considering the lowest generalization error. The ML classifier providing the better performance using the nested cross-validation was the logistic regression combined to the power transform Yeo–Johnson scaling feature method which corresponds to a zero-mean, unit-variance normalization with a power transform applied feature wise to make distribution of each radiomic feature Gaussian-like. To limit overfitting, the classifier encompassed a ridge regression for regularization (l2 penalty assignment) with a C value equal to 0.7. The final logistic regression-based established signature was a combination of the 100 input radiomics features, in which the feature with the highest coefficient in the decision function was sphericity, with a coefficient of 1.48. All other features had absolute coefficient less than 0.96. The 20 predominant features had absolute coefficients superior to 0.38. Among these features, five were shape features, two were first-order metrics, and 13 were based on texture matrices, with 6 extracted from the GLCM matrix (**Figure 3**).

Diagnosis Performance of the Classifier With a Ten-Repeated 5-Fold Cross-Validation

The model differentiated BM from GBM on the validation sets with a mean sensitivity of 85% [95% CI = (77%; 94%)], a specificity of 87% [95% CI = (78%; 97%)], a balanced accuracy of 86% [95% CI = (80%; 92%)], and an AUC of 92% [95% CI = (87%; 97%)] (**Figure 4**).

Diagnosis Performance of the Radiomic Classifier on the Test Set

The classifier correctly identified 12/16 BM and 18/21 GBM. Corresponding sensitivity, specificity, balanced accuracy and AUC were respectively equal to 75, 86, 80, and 85% (**Figures 4** and **5**).

Performance of the Radiologists

The performances of the neuroradiologists are described in **Table 2**. Even though differences in diagnostic performance were not statistically significant, we can highlight the fact that two radiology residents (R3 and R4) had lower scores than the classifier (respective balanced accuracies of 72 and 72%) whereas the two neuroradiologists with 10 years of experience (R1 and R2) and one radiology resident (R5) had better scores than the classifier (respective balanced accuracies of 87, 94 and 88% *versus* balanced accuracy of 80% for the classifier).

DISCUSSION

We have developed a radiomic classifier to differentiate solitary BM and GBM based on post-contrast 3DT1 MR images with high diagnostic performances on the validation and test sets. There was no statistically significant difference between classifier predictions and human reading by five trained neuroradiologists (two neuroradiologists with 10 years of experience, and three radiology residents with about 6 months of training exclusively in neuroradiology in an expert center).

The radiomic classifier, a logistic regression combined to the power transform Yeo–Johnson scaling feature method, was chosen because of its high performance, simplicity, and because it allowed an interpretation of the underlying model. Indeed, the fact that the radiomic feature with the most important coefficient value in the classifier was a shape feature, i.e. sphericity, partly allows an explainability of our radiomic

TABLE 1 | Demographics and clinical characteristics at diagnosis of the patients included in the training set and in the test set.

Patients characteristics	Training set		Test set	
	BM (n = 72)	GBM (n = 71)	BM (n = 16)	GBM (n = 21)
Mean patient age—years	59.29	58.25	59.00	58.19
(standard deviation)	(13.29)	(14.59)	(10.9)	(14.5)
Proportion of female gender (%)	53	38	50	52
Proportion of male gender (%)	47	62	50	48
Largest diameter in mm	41.40	53.39	33.85	54.93
median [range]	[20.77–77.92]	[24.11–88.12]	[22.41–63.63]	[32.61–102.53]
Patients from Center 1	56 (77.8%)	69 (97.2%)	5 (31.2%)	18 (85.7%)
Patients from Center 2	16 (22.2%)	2 (2.8%)	5 (31.2%)	3 (14.3%)
Primary lung cancer n (%)	29 (40.3)	–	8 (50)	–
Primary breast cancer n (%)	13 (18.0)	–	3 (18)	–
Melanoma n (%)	9 (12.5)	–	2 (12.5)	–
Primary colo-rectal cancer n (%)	5 (6.9)	–	0 (0)	–
Primary Clair cell carcinoma n (%)	4 (5.6)	–	1 (6.3)	–
Other primary cancer * n (%)	12 (16.7)	–	2 (12.5)	–

**Primary rare cancer: choriocarcinoma, sarcoma, salivary gland carcinoma, papillary carcinoma of the thyroid.*

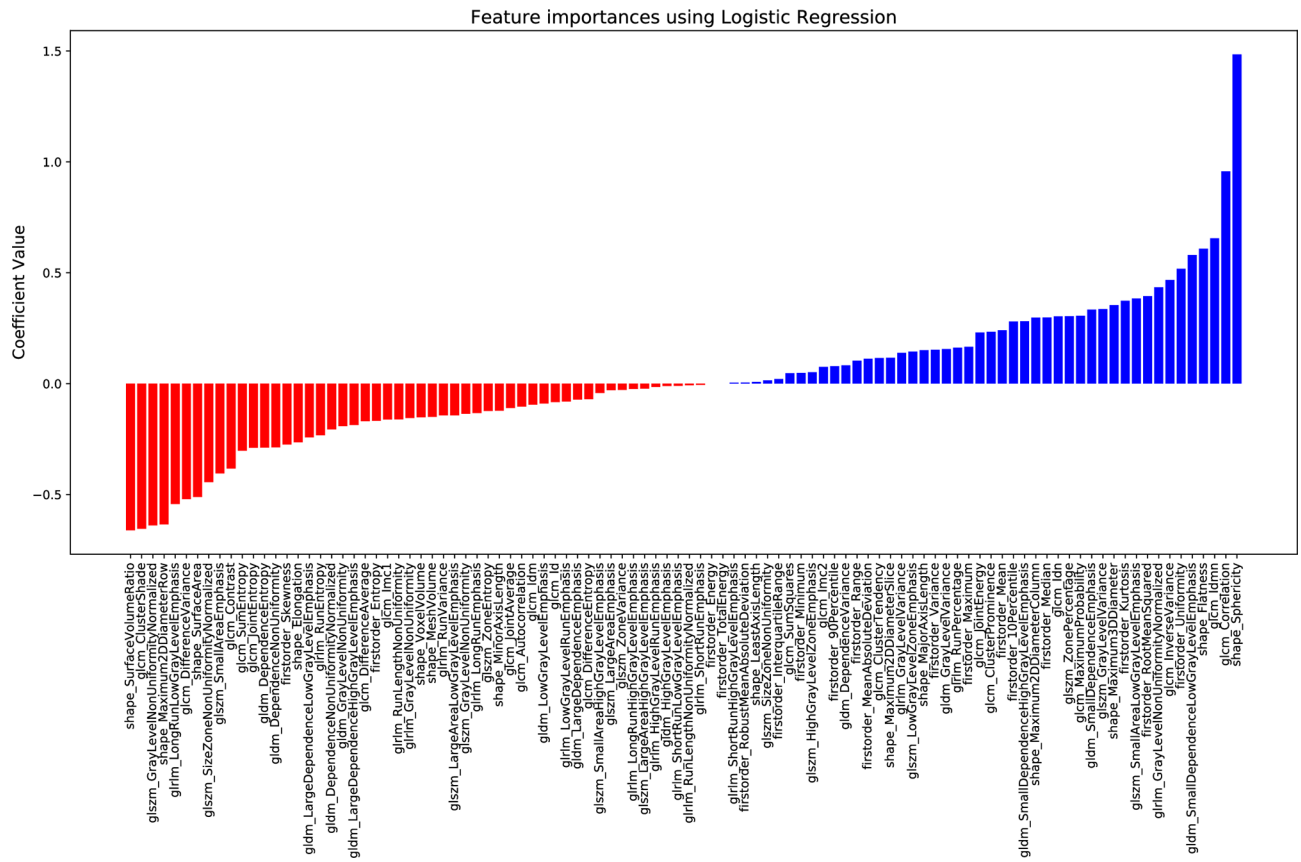


FIGURE 3 | Coefficient of each radiomic feature in the decision function for the proposed logistic regression model.

features-based classifier in contrast with the concept of the "black box" in some ML models, where even its designers cannot explain why the artificial intelligence reaches a decision (39). It introduces the notion of analyzing a tumor with its representation in 3D to differentiate solitary BM and GBM, which is usually not available during conventional reading of sectional imaging. Indeed, sphericity is a 3D shape feature representing a measure of roundness of the tumor, with a value ranging from 0 to 1, where 1 indicates a perfect sphere. The classifier showed that GBM have lower sphericity than BM (**Figure 6**), which was expected given the morphological characteristics of BM and GBM on histopathological slides. The more spherical the lesion is, the more likely it is to be a BM. Thus, the radiomic features-based classifier is consistent with current

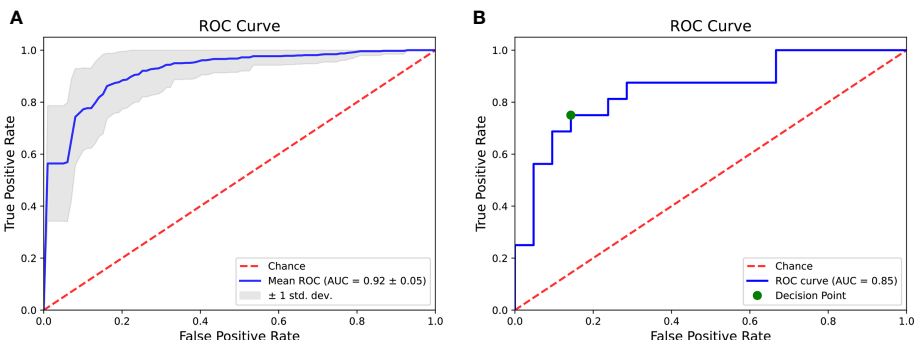


FIGURE 4 | Areas under the receiver operating characteristics curve of the radiomic classifier after ten-repeated 5-fold cross-validation **(A)** and on the test set **(B)**.

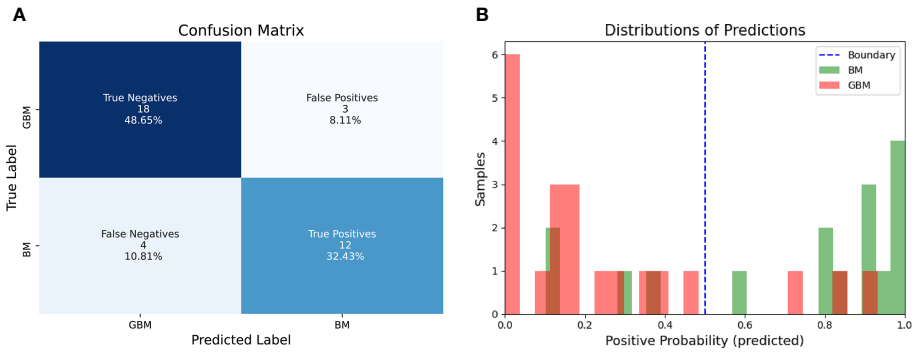


FIGURE 5 | Confusion Matrix of the radiomic model on the test set **(A)** and distribution of probabilities as predicted by the logistic regression model compared to ground truth **(B)**.

TABLE 2 | Sensitivities, specificities, balanced accuracies, positive predictive values, negative predictive values of the radiomic classifier and of the neuroradiologists (R1, R2, R3, R4, R5) on the test set.

Reader	Se*	Sp*	Balanced Accuracy	PPV*	PNV*	Se p-value*	Sp p-value*
Radiomic classifier	0.75	0.86	0.8	0.8	0.82	–	–
R1	0.88	0.86	0.87	0.82	0.9	0.41	1
R2	0.94	0.95	0.94	0.94	0.95	0.08	0.16
R3	0.69	0.76	0.72	0.76	0.69	0.65	0.41
R4	0.63	0.81	0.72	0.71	0.74	0.48	0.65
R5	0.81	0.95	0.88	0.93	0.87	0.65	0.16

**Se, Sensitivity; Sp, Specificity; PPV, Positive Predictive Value; PNV, Positive Negative Value; Se p-value, p-value (calculated with McNemar's test) of the difference between the sensibility of the radiomic classifier and the sensibility of the reader; Sp p-value, p-value (calculated with McNemar's test) of the difference between the specificity of the radiomic classifier and the specificity of the reader.*

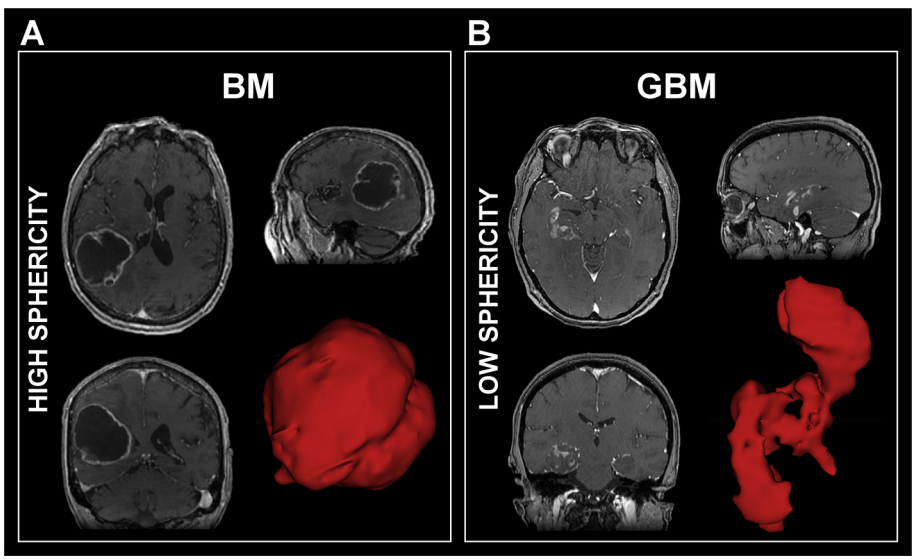


FIGURE 6 | Examples of 3D representation of a brain metastasis **(A)** for which the sphericity was equal to 0.76 and a glioblastoma **(B)** for which the sphericity was equal to 0.45. GBM, Glioblastoma; BM, Brain Metastasis.

morphological characteristics between BM and GBM, also adding further information regarding tumor heterogeneity imperceptible to the human eye, as the radiomic classifier is also based on other texture and intensity features. This result is in line with a pioneering paper (40) that described in 2012 2D circularity as one of the best morphological features to differentiate BM from GBM on the basis of a cohort of 50 patients.

In our study, we trained the ML classifier using a nested cross-validation and a ten-repeated 5-fold cross-validation on the training set in order to minimize overfitting. In addition to limit the extraction to 100 features (shape, first order and second order features) that we thought to be the most meaningful and interpretable, we selected a classifier model which could embed feature selection. For this model, L1 and L2 regularization methods were tested as hyperparameters. The L2 method provided the best performance in the cross-validation (CV) process, validating the usefulness of the 100 features. The selected classifier was then applied on a test set of data, which demonstrates that the high performances obtained were not random but generalizable. In the test set, 12/16 BM were correctly classified leading to a sensitivity of 75%. Among the four BM incorrectly classified, two had leptomeningeal enhancement, one had ventriculitis adjacent to the lesion and the fourth one had a multilocular lesion (**Figure 7**). The first three elements were absent from BM of the training set, which might have misled the classifier, suggesting the need for a larger training set which extensively reproduces all clinical situations encountered in clinic.

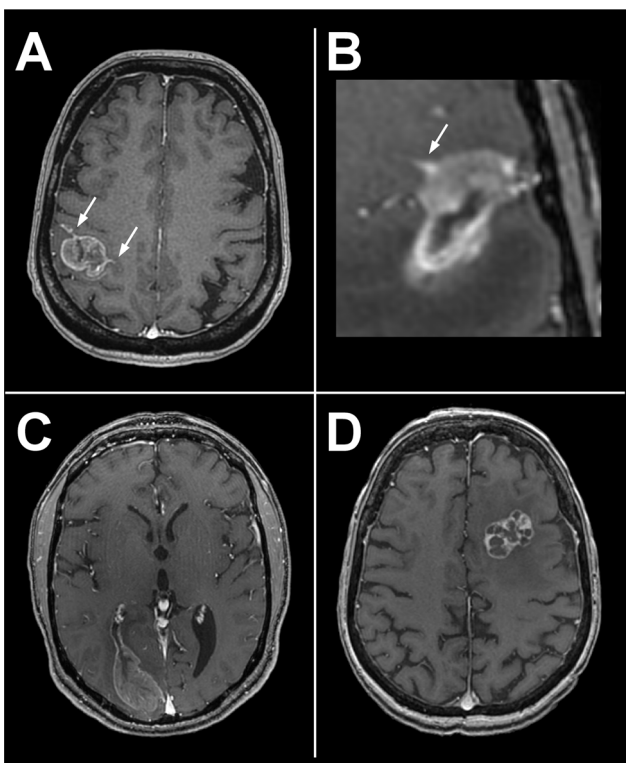


FIGURE 7 | Four incorrectly classified BM of the test set. Two of them presented tumoral leptomeningitis (arrows, **A, B**), one a metastatic ventriculitis **(C)** and the forth one a multilocular lesion **(D)**. Leptomeningitis and ventriculitis may have interfered with spatial delineation of tumor boundaries.

The results of our study are consistent with the results of three previous studies which also used radiomic features-based classifiers on post-contrast 3D T1 MR images to differentiate BM from GBM. Among these studies, Chen et al. (41) achieved diagnostic performance slightly lower than our on 134 patients, however without applying image pre-processing (42–44) nor evaluation on a test set. Artzi et al. (45) built a radiomics-based classifier on 358 patients and evaluated its performance on a test set of 88 patients. Excellent performances were achieved on the test set. However, the radiomic analysis was carried out on three central slices only to simplify the segmentation process, which did not allow 3D shape features such as sphericity, to be taken into account. Moreover, there was no comparison to human performance. In 2019, Qian et al. (46) used a cohort of 227 patients to train a ML classifier using cross-validation and evaluated it on an independent test set of 185 patients. Despite high diagnostic performances, there were biases in the study considering several radiomic features-based classifiers were evaluated on the test set. Finally, in 2020, Bae et al. (47) developed a Deep Neural Network (DNN) classifier based on post-contrast 3D T1-weighted and T2-weighted MR images, which outperformed the best-performing traditional machine learning model. Results showed excellent performance on an independent test set (AUC of 0.956 on the test set) and outperformed scores of two trained neuro radiologists. However, comparing the literature is not a trivial task due to the use of different data sets, each with varying degrees of complexity, suggesting the need for publicly available data sets.

Our study had a few limitations. First, we chose to build the radiomic features-based classifier on imaging data acquired on the same model of MR scanner with acquisitions performed with the same parameters in order to minimize inter-acquisition variability. This choice limited the number of patients included in the study. Several methods are available today to compensate for differences in image quality between scanners (36, 48), which should allow the applicability of our signature in other centers. In addition, no spatial resampling was applied to the MR images prior to feature extraction. Although this step is mandatory to obtain rotationally invariant features, no bias was introduced in the machine learning pipeline, as the entire cohort had exactly the same imaging parameters. The developed signature can finally be generalized to new patients with MR images of different voxel sizes by integrating an additional resampling step [resampling at a voxel size of (0.8 mm × 0.8 mm × 1.2 mm)]. Third, a semi-automatic method was used for tumor delineation and a single radiologist specialized in neurology performed the contouring of the lesions. Perturbation of the contours would have been an alternative to multiple segmentations to evaluate the robustness of the model developed to segmentation (49). However, the semi-automatic contouring process has been shown to be reliable between raters

for brain tumors (50). An integrated diagnostic support system should include an automatic segmentation of the volumes of interest to be considered for radiomics analysis. The automation of this step is now possible with high performance as demonstrated by the recent results of the BRATS challenge (51). Then, the radiomic only features-based classifier takes into account imaging data. The addition of the patient's age, gender, and medical history elements would lead to holistic models enabling to analyze the correlations between radiomic/non-radiomic features, and to better assess the added value of such a signature compared to more readily available clinical features (49). As well, only post-contrast 3DT1 MR images were considered. A more complex classifier combining data from other sequences such as FLAIR, T2 (47) or perfusion MR sequences may improve classification performance. Finally, a larger cohort of lesions studied would enable its generalizability.

In conclusion, we developed a radiomic features-based classifier based on post-contrast 3DT1 MR images that helps in differentiating GBM and solitary metastatic brain tumors with high diagnosis performance. The performance of the radiomic classifier equals that of neuroradiologists however needs to be improved in further studies including feature extraction applied on FLAIR and perfusion sequences. An interesting point is that the radiomic feature with the highest coefficient value in the classifier, namely sphericity, allows an explainability of the developed model. Future studies using this model on larger sets of patients may clarify its role and its benefit in differentiating these two lesions, particularly by a prospective study registered in a trial database.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Data can be available on demand. Requests to access these datasets should be directed to CR ch.robert@gustaveroussy.fr and ME myriam.edjlali@gmail.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by College de neurochirurgie, Paris. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AdC, AC, ME, SA, SR, and CR designed the research. AdC, AC, ME, SA, SR, and CR performed the research, analyzed, and interpreted the data, and wrote the paper. AT-E and PV reviewed histopathological data. AR, EDez, JP, AT-E, PV, and FD took care of the patients and retrieved the data. AdC, AC, AR, AT-E, SA, EDez, FD, SR, EDeu, CO, PV, JP, ME, and CR revised and approved the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This material is based upon work supported by ITMO PhysiCancer, the Fondation pour la Recherche Médicale (FRM; No. DIC20161236437), and Amazon Web Services (AWS). Amazon Web Services was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

We would like to thank Armelle Lesaunier, MD; Violeta Fridjoi, MD; Philippe Beyssen, MD; Juliette Fayard, MD; Corentin Provost, MD-MSc; Marjorie Latrasse, MD; Céline Corcy, MD, Joseph Benzakoun, MD-MSc; Wagih Ben Hassen, MD-MSc, Grégoire Boulouis, MD-MSc, Emmanuèle Lechapt-Zalcman, MD-PhD and Jean-François Meder, MD-PhD; for their precious help and advice in this study. **Figure 1** has been designed using resources from Flaticon.com.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.638262/full#supplementary-material

REFERENCES

1. Lemke DM. Epidemiology, Diagnosis, and Treatment of Patients With Metastatic Cancer and High-Grade Gliomas of the Central Nervous System. *J Infus Nurs: Off Publ Infus Nurs Soc* (2004) 27:263–9. doi: 10.1097/00129804-200407000-00012
2. Achrol AS, Rennert RC, Anders C, Soffietti R, Ahluwalia MS, Nayak L, et al. Brain Metastases. *Nat Rev Dis Primers* (2019) 5:5. doi: 10.1038/s41572-018-0055-y
3. Ostrom QT, Gittleman H, Farah P, Ondracek A, Chen Y, Wolinsky Y, et al. CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2006-2010. *Neuro-Oncology* (2013) 15 Suppl 2:ii1–56. doi: 10.1093/neuonc/not151
4. Server A, Josefsen R, Kulle B, Maehlen J, Schellhorn T, Gadmar Ø, et al. Proton Magnetic Resonance Spectroscopy in the Distinction of High-Grade Cerebral Gliomas From Single Metastatic Brain Tumors. *Acta Radiol (Stockholm Sweden: 1987)* (2010) 51:316–25. doi: 10.3109/02841850903482901
5. Benzakoun J, Robert C, Legrand L, Pallud J, Meder JF, Oppenheim C, et al. Anatomical and Functional MR Imaging to Define Tumoral Boundaries and Characterize Lesions in Neuro-Oncology. *Cancer Radiother: J la Soc Fr Radiother Oncol* (2020) 24:453–62. doi: 10.1016/j.canrad.2020.03.005
6. Daumas-Duport C, Meder JF, Monsaingeon V, Missir O, Aubin ML, Szikla G. Cerebral Gliomas: Malignancy, Limits and Spatial Configuration. Comparative Data From Serial Stereotaxic Biopsies and Computed Tomography (a Preliminary Study Based on 50 Cases). *J Neuroradiol = J Neuroradiol* (1983) 10:51–80.
7. Petrella JR, Provenzale JM. MR Perfusion Imaging of the Brain: Techniques and Applications. *AJR Am J Roentgenol* (2000) 175:207–19. doi: 10.2214/ajr.175.1.1750207
8. Lin L, Xue Y, Duan Q, Sun B, Lin H, Huang X, et al. The Role of Cerebral Blood Flow Gradient in Peritumoral Edema for Differentiation of

Glioblastomas From Solitary Metastatic Lesions. *Oncotarget* (2016) 7:69051–9. doi: 10.18632/oncotarget.12053
9. Blasel S, Jurcoane A, Franz K, Morawe G, Pellikan S, Hattingen E. Elevated Peritumoural rCBV Values as a Mean to Differentiate Metastases From High-Grade Gliomas. *Acta Neurochirurg* (2010) 152:1893–9. doi: 10.1007/s00701-010-0774-7
10. Galanaud D, Nicoli F, Figarella-Branger D, Roche P, Confort-Gouny S. Le Fur Y, Et al. Spectroscopie Par Résonance Magnétique Des Tumeurs Cérébrales. *J Radiol* (2006) 87:822–32. doi: 10.1016/S0221-0363(06)74090-2
11. Tsolaki E, Svolos P, Kousi E, Kapsalaki E, Fountas K, Theodorou K, et al. Automated Differentiation of Glioblastomas From Intracranial Metastases Using 3T MR Spectroscopic and Perfusion Data. *Int J Comput Assist Radiol Surg* (2013) 8:751–61. doi: 10.1007/s11548-012-0808-0
12. Tsougos I, Svolos P, Kousi E, Fountas K, Theodorou K, Fezoulidis I, et al. Differentiation of Glioblastoma Multiforme From Metastatic Brain Tumor Using Proton Magnetic Resonance Spectroscopy, Diffusion and Perfusion Metrics at 3 T. *Cancer Imaging* (2012) 12:423–36. doi: 10.1102/1470-7330.2012.0038
13. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. Perfusion MRI as a Diagnostic Biomarker for Differentiating Glioma From Brain Metastasis: A Systematic Review and Meta-Analysis. *Eur Radiol* (2018) 28:3819–31. doi: 10.1007/s00330-018-5335-0
14. Nguyen LN, Maor MH, Oswald MJ. Brain Metastases as the Only Manifestation of an Undetected Primary Tumor. *Cancer* (1998) 83:2181–4. doi: 10.1002/(sici)1097-0142(19981115)83:10<2181::aid-cncr17>3.0.co;2-j
15. Rudà R, Borgognone M, Benech F, Vasario E, Soffietti R. Brain Metastases From Unknown Primary Tumour: A Prospective Study. *J Neurol* (2001) 248:394–8. doi: 10.1007/s004150170180
16. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A Summary. *Acta Neuropathol* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1
17. Lin X, DeAngelis LM. Treatment of Brain Metastases. *J Clin Oncol: Off J Am Soc Clin Oncol* (2015) 33:3475–84. doi: 10.1200/JCO.2015.60.9503
18. Weller M, van den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E, et al. European Association for Neuro-Oncology (EANO) Guideline on the Diagnosis and Treatment of Adult Astrocytic and Oligodendroglial Gliomas. *Lancet Oncol* (2017) 18:e315–29. doi: 10.1016/S1470-2045(17)30194-8
19. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
20. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer (Oxford England: 1990)* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
21. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and Challenges for the Implementation of Computational Medical Imaging (Radiomics) in Oncology. *Ann Oncol: Off J Eur Soc Med Oncol* (2017) 28:1191–206. doi: 10.1093/annonc/mdx034
22. Gillies RJ, Anderson AR, Gatenby RA, Morse DL. The Biology Underlying Molecular Imaging in Oncology: From Genome to Anatome and Back Again. *Clin Radiol* (2010) 65:517–21. doi: 10.1016/j.crad.2010.04.005
23. Hajianfar G, Shiri I, Maleki H, Oveisi N, Haghparast A, Abdollahi H, et al. Noninvasive O6 Methylguanine-DNA Methyltransferase Status Prediction in Glioblastoma Multiforme Cancer Using Magnetic Resonance Imaging Radiomics Features: Univariate and Multivariate Radiogenomics Analysis. *World Neurosurg* (2019) 132:e140–61. doi: 10.1016/j.wneu.2019.08.232
24. Nicolasjilwan M, Hu Y, Yan C, Meerzaman D, Holder CA, Gutman D, et al. Addition of MR Imaging Features and Genetic Biomarkers Strengthens Glioblastoma Survival Prediction in TCGA Patients. *J Neuroradiol = J Neuroradiol* (2015) 42:212–21. doi: 10.1016/j.neurad.2014.02.006
25. Kotrotsou A, Zinn PO, Colen RR. Radiomics in Brain Tumors: An Emerging Technique for Characterization of Tumor Environment. *Magnet Reson Imaging Clinics North America* (2016) 24:719–29. doi: 10.1016/j.mric.2016.06.006
26. Lohmann P, Galldiks N, Kocher M, Heinzel A, Filss CP, Stegmayr C, et al. Radiomics in Neuro-Oncology: Basics, Workflow, and Applications. *Methods (San Diego Calif)* (2021) 188:112–21. doi: 10.1016/j.ymeth.2020.06.003
27. Kickingereder P, Neuberger U, Bonekamp D, Piechotta PL, Götz M, Wick A. Et al. Radiomic Subtyping Improves Disease Stratification Beyond Key Molecular, Clinical, and Standard Imaging Characteristics in Patients With Glioblastoma. *Neuro-Oncology* (2018) 20:848–57. doi: 10.1093/neuonc/nox188
28. Park JE, Kickingereder P, Kim HS. Radiomics and Deep Learning From Research to Clinical Workflow: Neuro-Oncologic Imaging. *Korean J Radiol* (2020) 21:1126–37. doi: 10.3348/kjr.2019.0847
29. Shboul ZA, Chen J, MIftekharuddin K. Prediction of Molecular Mutations in Diffuse Low-Grade Gliomas Using MR Imaging Features. *Sci Rep* (2020) 10:3711. doi: 10.1038/s41598-020-60550-0
30. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE Trans Med Imaging* (2010) 29:1310–20. doi: 10.1109/TMI.2010.2046908
31. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. *NeuroImage* (2011) 54:2033–44. doi: 10.1016/j.neuroimage.2010.09.025
32. Smith SM. Fast Robust Automated Brain Extraction. *Hum Brain Mapp* (2002) 17:143–55. doi: 10.1002/hbm.10062
33. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. *arXiv:1811.02629 [cs stat]* (2019). doi: 10.17863/CAM.38755
34. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339
35. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295:328–38. doi: 10.1148/radiol.2020191145
36. Carré A, Klausner G, Edjlali M, Lerousseau M, Briend-Diop J, Sun R, et al. Standardization of Brain MR Images Across Machines and Protocols: Bridging the Gap for MRI-Based Radiomics. *Sci Rep* (2020) 10:12340. doi: 10.1038/s41598-020-69298-z
37. Duron L, Balvay D, Vande Perre S, Bouchouicha A, Savatovsky J, Sadik JC, et al. Gray-Level Discretization Impacts Reproducible MRI Radiomics Texture Features. *PloS One* (2019) 14:e0213459. doi: 10.1371/journal.pone.0213459
38. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-Learn: Machine Learning in Python. *arXiv:1201.0490 [cs]* (2018).
39. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and Explainability of Artificial Intelligence in Medicine. *WIREs Data Min Knowl Discovery* (2019) 9:e1312. doi: 10.1002/widm.1312
40. Mouthuy N, Cosnard G, Abarca-Quinones J, Michoux N. Multiparametric Magnetic Resonance Imaging to Differentiate High-Grade Gliomas and Brain Metastases. *J Neuroradiol* (2012) 39:301–7. doi: 10.1016/j.neurad.2011.11.002
41. Chen C, Ou X, Wang J, Guo W, Ma X. Radiomics-Based Machine Learning in Differentiation Between Glioblastoma and Metastatic Brain Tumors. *Front Oncol* (2019) 9:806. doi: 10.3389/fonc.2019.00806
42. Kuo MD, Jamshidi N. Behind the Numbers: Decoding Molecular Phenotypes With Radiogenomics–Guiding Principles and Technical Considerations. *Radiology* (2014) 270:320–5. doi: 10.1148/radiol.13132195
43. Aerts HJWL. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. *JAMA Oncol* (2016) 2:1636–42. doi: 10.1001/jamaoncol.2016.2631
44. Reuzé S, Schernberg A, Orlhac F, Sun R, Chargari C, Dercle L, et al. Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls, and Challenges. *Int J Radiat Oncol Biol Phys* (2018) 102:1117–42. doi: 10.1016/j.ijrobp.2018.05.022
45. Artzi M, Bressler I, Bashat DB. Differentiation Between Glioblastoma, Brain Metastasis and Subtypes Using Radiomics Analysis. *J Magnet Reson Imaging* (2019) 50:519–28. doi: 10.1002/jmri.26643
46. Qian Z, Li Y, Wang Y, Li L, Li R, Wang K, et al. Differentiation of Glioblastoma From Solitary Brain Metastases Using Radiomic Machine-Learning Classifiers. *Cancer Lett* (2019) 451:128–35. doi: 10.1016/j.canlet.2019.02.054
47. Bae S, An C, Ahn SS, Kim H, Han K, Kim SW, et al. Robust Performance of Deep Learning for Distinguishing Glioblastoma From Single Brain Metastasis

Using Radiomic Features: Model Development and Validation. *Sci Rep* (2020) 10:12110. doi: 10.1038/s41598-020-68980-6

48. Orlhac F, Lecler A, Savatovski J, Goya-Outi J, Nioche C, Charbonneau F, et al. How Can We Combat Multicenter Variability in MR Radiomics? Validation of a Correction Procedure. *Eur Radiol* (2021) 31:2272–80. doi: 10.1007/s00330-020-07284-9
49. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141
50. Tixier F, Um H, Young RJ, Veeraraghavan H. Reliability of Tumor Segmentation in Glioblastoma: Impact on the Robustness of MRI-Radiomic Features. *Med Phys* (2019) 46:3582–91. doi: 10.1002/mp.13624
51. Henry T, Carre A, Lerousseau M, Estienne T, Robert C, Paragios N, et al. Brain Tumor Segmentation With Self-Ensembled, Deeply-Supervised 3D U-Net Neural Networks: A BraTS 2020 Challenge Solution. *arXiv:2011.01045 [cs eess]* (2020). doi: 10.1007/978-3-030-72084-1_30

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

frontiers in Oncology

SYSTEMATIC REVIEW
published: 14 July 2021
doi: 10.3389/fonc.2021.628408

Updated Trends in Imaging Practices for Pancreatic Neuroendocrine Tumors (PNETs): A Systematic Review and Meta-Analysis to Pave the Way for Standardization in the New Era of Big Data and Artificial Intelligence

Ephraïm Partouche[1], Randy Yeh[2], Thomas Eche[1], Laura Rozenblum[3], Nicolas Carrere[4], Rosine Guimbaud[5], Lawrence O. Dierickx[6], Hervé Rousseau[1], Laurent Dercle[7†] and Fatima-Zohra Mokrane[1†]*

[1] Radiology Department, Rangueil University Hospital, Toulouse, France, [2] Memorial Sloan Kettering Cancer Center, Molecular Imaging and Therapy Service., New York, NY, United States, [3] Sorbonne Université, Service de Médecine Nucléaire, AP-HP, Hôpital La Pitié-Salpêtrière, Paris, France, [4] Surgery Department, Toulouse University Hospital, Toulouse, France, [5] Oncology Department, Toulouse University Hospital, Toulouse, France, [6] Nuclear Medicine Department, IUCT-Oncopole, Toulouse, France, [7] Department of Radiology, New York Presbyterian Hospital, Columbia University Vagellos College of Physicians and Surgeons, New York, NY, United States

OPEN ACCESS

Edited by:
Natalie Julie Serkova, University of Colorado, United States

Reviewed by:
Zhongqiu Wang, Affiliated Hospital of Nanjing University of Chinese Medicine, China
Giuseppe Esposito, MedStar Georgetown University Hospital, United States

***Correspondence:**
Fatima-Zohra Mokrane
mokrane.fz@chu-toulouse.fr

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 11 November 2020
Accepted: 25 June 2021
Published: 14 July 2021

Citation:
Partouche E, Yeh R, Eche T, Rozenblum L, Carrere N, Guimbaud R, Dierickx LO, Rousseau H, Dercle L and Mokrane F-Z (2021) Updated Trends in Imaging Practices for Pancreatic Neuroendocrine Tumors (PNETs): A Systematic Review and Meta-Analysis to Pave the Way for Standardization in the New Era of Big Data and Artificial Intelligence. Front. Oncol. 11:628408. doi: 10.3389/fonc.2021.628408

Purpose: Medical imaging plays a central and decisive role in guiding the management of patients with pancreatic neuroendocrine tumors (PNETs). Our aim was to synthesize all recent literature of PNETs, enabling a comparison of all imaging practices.

Methods: based on a systematic review and meta-analysis approach, we collected; using MEDLINE, EMBASE, and Cochrane Library databases; all recent imaging-based studies, published from December 2014 to December 2019. Study quality assessment was performed by QUADAS-2 and MINORS tools.

Results: 161 studies consisting of 19852 patients were included. There were 63 'imaging' studies evaluating the accuracy of medical imaging, and 98 'clinical' studies using medical imaging as a tool for response assessment. A wide heterogeneity of practices was demonstrated: imaging modalities were: CT (57.1%, n=92), MR (42.9%, n=69), PET/CT (13.3%, n=31), and SPECT/CT (9.3%, n=15). International imaging guidelines were mentioned in 2.5% (n=4/161) of studies. In clinical studies, imaging protocol was not mentioned in 30.6% (n=30/98) of cases and only mentioned imaging modality without further information in 63.3% (n=62/98), as compared to imaging studies (1.6% (n=1/63) of ($p<0.001$)). QUADAS-2 and MINORS tools deciphered existing biases in the current literature.

Conclusion: We provide an overview of the updated current trends in use of medical imaging for diagnosis and response assessment in PNETs. The most commonly used

imaging modalities are anatomical (CT and MRI), followed by PET/CT and SPECT/CT. Therefore, standardization and homogenization of PNETs imaging practices is needed to aggregate data and leverage a big data approach for Artificial Intelligence purposes.

Keywords: pancreatic neuroendocrine tumors (pNETs), imaging practices, meta-analysis, systematic review, computed tomogaphy, MRI, PET - positron emission tomography

INTRODUCTION

Pancreatic neuroendocrine tumors (PNETs) are relatively uncommon tumors, with an increasing incidence due to widespread use of cross-sectional imaging (1, 2). PNETs represent a heterogeneous entity, characterized by a wide variation in clinical presentation and prognosis due to tumor functional status, possible genetic context, and variable aggressiveness, making the management of PNETs highly challenging (3–6).

Medical imaging plays a critical role in guiding PNETs patients management (7, 8). Computed tomography (CT) is often the initial modality used to evaluate pancreatic lesions, mostly because of its high spatial and temporal resolution (9), and correlation with histological prognostic factors (10). Magnetic resonance imaging (MRI) also plays a major role in pancreatic tumor characterization (11, 12) and demonstrates imaging features that can be correlated with tumor aggressiveness and grade (13, 14). A wide range of molecular imaging techniques are also used in PNET patients, as Somatostatin receptor (SSTR) imaging with single photon emission tomography/CT (SPECT/CT) and positron emission tomography/CT (PET/CT) (15), ^{18}F-DOPA (16) and ^{18}F-FDG with PET/CT (17). Molecular imaging techniques have shown a high association with tumor grade and are critical for theranostic approaches (18–21). Increasingly, a multimodal imaging strategy, combining anatomical and molecular techniques, are leveraged for imaging-guided approaches to personalize and optimize patient management (22, 23).

PNETs present four characteristics that make imaging evaluation challenging. First, PNETs are hypervascular slow-growing tumors and therefore, limiting thus the value of using Response-Evaluation-Criteria-in-Solid-Tumors (RECIST) because tumor burden remains stable rather than decreased in patients with the best survival (24, 25). Second, tumor size measurements may vary with contrast medium injection protocols on either CT or MRI (26). Third, new targeted cytostatic agents are used in PNETs treatment and alternative imaging criteria are needed, as tumor density change on perfusion CT (25). Fourth, immune-checkpoint modulators (ICMs) are currently being evaluated in several PNETs clinical trials (27, 28). Because of their mechanisms of action, radiologists should be aware of new patterns of response and progression to immune therapies, as well as immune Related Adverse Events (iRAE) (29–34). In addition, treatment beyond progression is allowed and immune RECIST (iRECIST) criteria should be used (35). This new era of immunotherapy makes tumor response assessment in PNETs even more difficult.

One of the key concepts unique to medical imaging is that the relevance and clinical utility of information derived from imaging depends on technical imaging parameters and acquisition. Therefore, using poor quality imaging techniques in clinical routine or in scientific studies may lead to inaccurate and biased results. Imaging examinations need to be technically adequate, uniform and homogeneous, which is even more salient in imaging PNETs since a majority of PNETs are hypervascular and up to 20% of PNETs measure 2 cm or less. Therefore, CT or MRI scans without an arterial phase acquisition or thin slices drastically reduces the sensitivity of the examination (36). Molecular imaging is also sensitive to technical parameters, with optimal patient preparation, administered radiotracer activity, and acquisition time as essential elements for high-quality molecular examinations (37, 38). Thus, imaging standardization is critically important in both clinical practice and in medical research, which encompasses clinical therapeutic trials and imaging research studies (i.e. diagnostic accuracy studies, comparison of imaging techniques, etc.). With respect to clinical trials, survival assessment, therapeutic response or prognostic value of a therapeutic effect are mostly dependent on the tumor imaging response mostly based on tumor size variations assessed by medical imaging. In an effort to harmonize and standardize clinical practice, the European Neuroendocrine Tumor Society (ENETS) published consensus guidelines for the standards of care in 2009 (38), which was updated in 2017 (37) and emphasized the importance of PNETs diagnostic procedures and technical quality of imaging methods.

In order to unravel the potential "imaging databases" that exist at the international level, we have conducted an updated review on the current imaging trends in clinical practice and research, based on a systematic review and meta-analysis approach, evaluating standardization of medical imaging in international PNETs studies during the last five years. The aim of this study was to evaluate the methodology and level of standardization of imaging in the recent literature of PNETs. We have focused this review on recent literature in order to reflect updated and current practices in imaging of PNETs, especially given the growth of literature in newer imaging techniques and theranostics.

MATERIALS AND METHODS

A preliminary step was conducted before stating this study, in which we have reviewed all available literature using different international guidelines in this area. This search is summarized in **Supplementary Table 1**.

Literature Search Strategy and Study Selection

The study protocol was developed and previously registered in PROSPERO (study number: *CRD42020168542*).

In order to review the entire recent published literature on PNETs during the last five years, a systematic search of major reference databases MEDLINE (PubMed), CENTRAL (Cochrane Central Register of Controlled Trials) and EMBASE was undertaken in December 2019, according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (39). PRISMA checklist is shown in **Supplementary Table 2**.

Key search terms included "pancreatic neuroendocrine tumor/tumor/neoplasm/carcinoma", and "islet cell adenoma/tumor". Study selection focused on recent literature, from December 1, 2014 to December 1, 2019. Details of search terms used for each database is reported on **Supplementary Table 3**.

After removal of duplicate studies and publications including only an abstract, non-English and non-human studies were automatically excluded from the study selection, as were case reports, systematic or non-systematic reviews and meta-analysis. The bibliographic management commercial software used was EndNote X9.3.1.

Inclusion and Exclusion Criteria

Titles and abstracts of studies were initially screened to select eligible publications, by removing studies dealing with the following topics: (1) Publications with other NETs than PNETs or non-neuroendocrine neoplasms/*Not exclusively about PNETs*; (2) Inherited syndromes and mutated population (MEN-1, VHL, NF-1); (3) Studies investigating the diagnostic value of Ultrasound (US), endoscopic-US (EUS), EUS-guided fine needle aspiration (EUS-FNA); (4) Basic science.

Then, full-text studies of eligible publications were reviewed to select all of imaging-based publications, separated into 2 subgroups: 1) imaging diagnostic accuracy studies and related studies: comparison of two imaging techniques, evaluation of prognosis value of a subtype of imaging technique; 2) clinical studies, including observational and cohort studies, experimental studies and clinical trials, for therapeutic or prognostic purposes, whose results are themselves based on therapeutic responses and survival endpoints (Progression-Free Survival PFS, Disease-Free Survival DFS, Disease Control Rate DCR, Recurrence-Free Survival RFS, Objective Response Rate ORR), fully dependent on imaging.

All studies identified by the search were screened for eligibility by two independent authors (E.P and F.Z.M), blinded to each other's decisions. In case of disagreement, a consensus was reached after review with a third reviewer (L.D).

Data Extraction

The two reviewers (E.P and F.Z.M) who performed the initial literature search independently extracted relevant data for each selected imaging-based publication, using a standardized form. This includes *(a)* General publication data (title, authors, journal and year of publication, country of origin); *(b)* Study design characteristics; *(c)* Demographic, clinical and pathological variables; *(d)* Type of imaging-based survival endpoint assessed (PFS, RFS, DFS, DCR, ORR) *(e)* Any imaging available data (imaging modality used, response evaluation criteria used); *(f)* Technical characteristics and acquisition parameters of each imaging modality; *(g)* Reference or mentioning of an imaging technical guideline from international NET societies. **Table 1** summarizes all extracted data. Two investigators (E.P, F.Z.M) working in duplicate independently assessed all studies. Discordances were discussed with a third reviewer (L.D) and resolved by consensus.

Methodological Quality: Risk of Bias and Quality of Evidence

We assessed the risk of bias for all included studies. First, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool was used without modification as provided by the QUADAS-2 group (40). This tool aims to evaluate the methodological quality applied to four "risk of bias" domains and three "concerns regarding applicability" domains (a total of 7 items to assess). Then, we also used the Methodological Index for Non-Randomized Studies (MINORS) grading score for clinical studies (41). MINORS score is a validated tool which uses eight graded questions for non-comparative studies. We judged each domain as presenting low, high, or unclear risk of bias by using a numeric score: each item can be scored as 0 (not reported), 1 (reported but inadequate) and 2 (reported and adequate). Ideal global score varies from 16 for non-comparative studies and 24 for comparative ones.

Statistical Analysis

Analyses were conducted using Microsoft Excel (v2019, Microsoft, USA, 2019) and open-source R software (version 3.6.2; R Foundation for Statistical Computing, Vienna, Austria). A p-value less than 0.05 was considered to indicate statistical significance (α=0.05).

RESULTS

Identification and Selection of Studies

The literature search yielded 9982 studies. Following the removal of duplicates, 6509 studies remained. Among them, 4189 records including only an abstract (n=149), non-English (n=293) and non-human studies (n=2115) were automatically excluded from the study selection, as were case reports (n=1058), systematic or non-systematic reviews and meta-analysis (n=574). Afterwards, 2320 studies were screened on the basis of title and abstract. Among them, 1846 were excluded: studies not exclusively dealing with PNETs (n=951), inherited syndromes (n=133), studies evaluating ultrasound (US) (n=250), basic science studies (n=317), and unrelated studies (n=195).

474 publications were included and fully reviewed, of which 161 were identified as imaging-based studies and included in the final analysis: 63 *imaging studies* on diagnostic accuracy studies and related studies and 98 *clinical studies* based on therapeutic responses and survival endpoints, fully dependent on imaging.

TABLE 1 | Extracted relevant data.

General publication data	Title	
	Authors	
	Journal and year of publication	
	Geographical origin	
Study design and characteristics	Diagnostic accuracy study	
	Cohort study	
	Clinical trial (with phase)	
	Case-control study	
	Data collection method	Prospective Retrospective
	Comparative	
	Randomized	
	Institutional design	Monocentric National multicentric International multicentric
Duration of study	Time interval	Data collection start date Data collection end date Duration of time interval
Demographic and clinical variables	Number of patients	
	Mean age	
	Gender	Male Female
	Inherited syndrome	
	Metastatic disease	
Tumor functional status	Functional	Insulinoma Gastrinoma Glucagonoma VIPoma Other
	Nonfunctional	
Pathologic features	ENETS/WHO grading	G1 G2 G3
	TNM/UICC staging	Stage 1 Stage 2 Stage 3 Stage 4
Imaging modality	CT	
	MRI	
	SPECT/CT	
	PET/CT	
	Available technical acquisition parameters	
	Number of equipment	
Anatomical imaging practices	Detailed acquisition protocol	
	Contrast agent administration	
	Use of bolus tracking	
	Slice thickness	
Nuclear medicine imaging practices	Time before acquisition	
	Time acquisition	
	Molecular radiotracer	^{68}Ga-DOTA ^{18}F-FDG ^{18}F-DOPA GLP-1R

(Continued)

TABLE 1 | Continued

Place of imaging in clinical studies	Radiotracer dose Type of imaging-based survival endpoint Imaging response evaluation criteria Mention imaging technical guidelines

The PRISMA flow chart of literature search and study selection process is shown in **Figure 1**.

Study Characteristics

Between December 2014 and December 2019, the average [range] annual number of publications was 12.6 [2;22] for imaging studies and 19.6 [1;24] for clinical studies (**Figure 2A**). Sixty-three imaging studies (diagnostic accuracy studies and related studies) and ninety-eight clinical studies have been identified. Eighty-nine percent (56/63) of the imaging studies, and eighty-five percent (83/98) of clinical ones, were retrospective (**Table 2**, **Figure 2B**). The most common study design was retrospective cohorts (n=84, 85.7%). Only 3.2% (n=2) of imaging publications and 4.1% (n=4) of clinical publications were randomized. Similarly, a minority of the studies was comparative: 14.3% (n=9) of imaging publications and 15.3% (n=15) of clinical publications.

A majority of imaging studies (n=60, 95.2%) were performed only in a single institution while 4.8% (n=3) were multi-institutional studies from a single country. Strikingly, no studies were international collaborations. Range of multi-institutional from a single country studies upon clinical studies was higher (27.6% (n=27)), and 11.2% (n=11) were international studies (p=<0.001) (**Figure 2C**). In addition, 92.9% (n=91) of clinical studies were non-blinded, against only 4.8% (n=3) for imaging studies (p=<0.001) (**Table 2**).

The mean duration of retrospective studies (time interval chosen in the database used) was 5.7 years [0;27] (**Figure 2D**). Moreover, 71.4% (n=70/98) of clinical studies started before 2005, while 53.9% (n=34/63) of imaging studies used data collected after 2005 (p=0.007).

Finally, the geographical origin per continent of the populations studied is firstly represented by Asia (n=72, 41.1%), followed by Europe (n=64, 36.3%), and North America (n=36, 20.6%) (**Figure 2E**).

Demographic, Clinical and Pathological Variables

In this study, 19852 patients were included, with 15728 patients in "clinical studies" and 4124 patients in "imaging studies. The mean [SD] number of patients was lower in imaging studies than clinical studies: 65.5 [52.7] and 160.5 [345.3] respectively (p=0.032). For imaging studies, 51.8% of included patients were female, with a mean (SD) age of 49.4 (15.9). In clinical studies, 49.1% were female with a mean (SD) age of 53.4 (13.0).

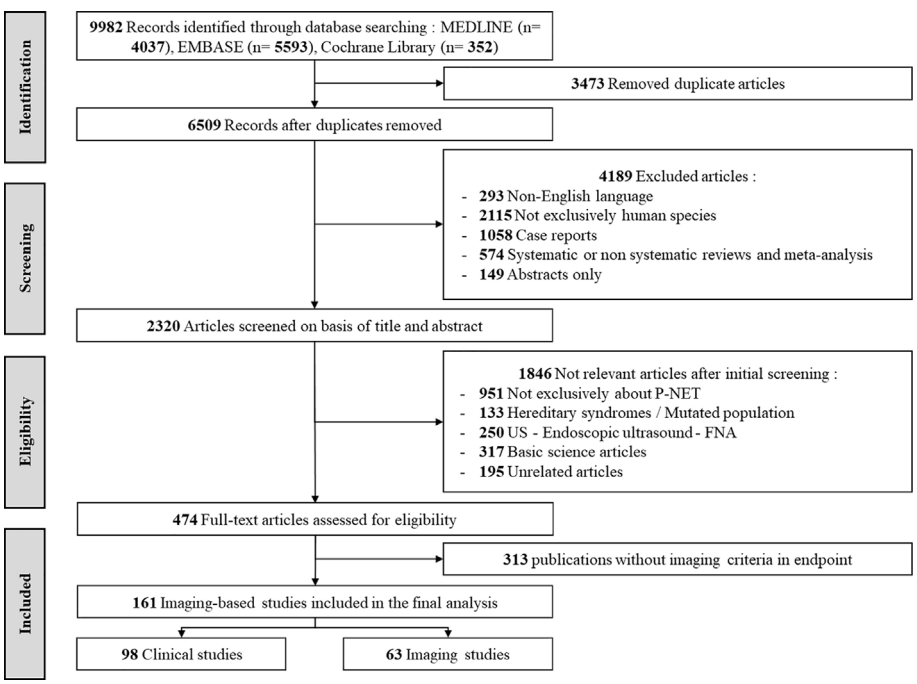


FIGURE 1 | PRISMA flow chart of literature search and study selection process.

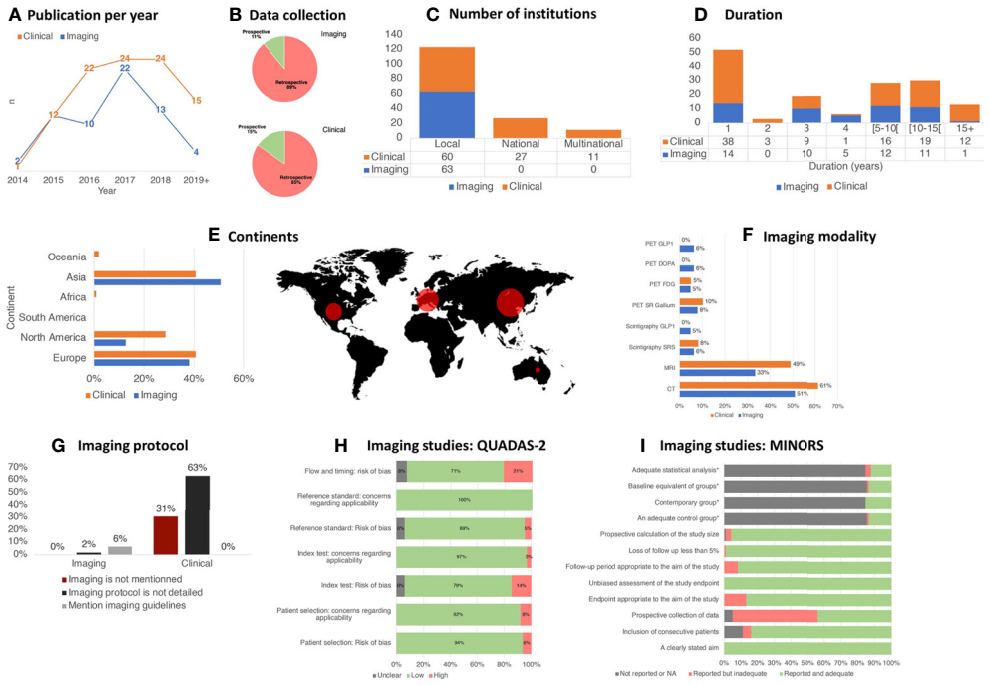


FIGURE 2 | Overview of trends in imaging practices for PNETs.

All pathological variables, especially the tumor functional status, type of PNET if functional, rate of mutated population and metastatic disease, ENETS/WHO grading and TNM/UICC staging are summarized in **Table 3**.

Recent Trends in PNET Imaging

The two imaging modalities most frequently used in the recent PNET literature are CT and MRI representing 57.1% (n=92/161) and 42.9% (n=69/161) of studies, respectively. Nuclear medicine imaging was less frequently used with 19.3% (n=31/161) of studies utilizing PET/CT and 9.3% (n=15/161) utilizing planar scintigraphy and/or SPECT/CT. No significant difference was observed between imaging and clinical studies. Detailed repartition of imaging modalities per subgroup is illustrated in **Figure 2F**.

Standardization of Practice: Reporting of Imaging Technical Parameters

In imaging studies, available information on imaging methods, examination protocol and technical details for each imaging modality were collected and summarized in **Table 4**. With respect to CT, most studies reported a detailed imaging acquisition protocol (93.8%), with almost all studies using multiphase contrast-enhanced acquisition, except one reporting single phase acquisition. CT slice thickness parameters were

TABLE 2 | Study characteristics per subgroup.

Studies characteristics		Imaging studies (n=63)	Clinical studies (n=98)	*P* value
Data collection	Prospective (%)	7/63 (11.1%)	15/98 (15.3%)	0.602
	Retrospective (%)	56/63 (88.9%)	83/98 (84.7%)	
Design	Diagnostic accuracy studies (%)	63/63 (100%)	–	N.A
	Cohort (%)	–	84/98 (85.7%)	N.A
	Clinical trial (%)	–	13/98 (13.3%)	N.A
Method	Randomization (%)	2/63 (3.2%)	4/98 (4.1%)	1
	Comparative (%)	9/63 (14.3%)	15/98 (15.3%)	1
Institutional design	Monocentric (%)	60/63 (95.2%)	60/98 (61.2%)	**<0.001**
	National multicentric (%)	3/63 (4.8%)	27/98 (27.6%)	
	International multicentric (%)	–	11/98 (11.2%)	
Blinding	Non-blinded (%)	3/63 (4.8%)	91/98 (92.9%)	**<0.001**
	Simple blinded (%)	60/63 (95.2%)	5/98 (5.1%)	
	Double blinded (%)	–	2/98 (2.0%)	

The bold values represent significant values.

TABLE 3 | Demographic, clinical and pathological variables per subgroup.

	Imaging studies (n=63)	Clinical studies (n=98)	*P* value
Number of patients	65.5 (52.7)	160.5 (345.3)	**0.032**
Age	49.4 (15.9)	53.4 (13.0)	0.077
Gender*			
Male	29.8 (27.3); 48.2%	79.7 (179.0); 50.9%	**0.03**
Female	32.1 (28.8); 51.8%	76.9 (166.7); 49.1%	**0.036**
Tumor functional status			
Functional	9.9 (18.4)	20.4 (52.3)	0.129
Insulinoma	6.3 (13.0)	7.9 (27.5)	0.657
Gastrinoma	1.3 (5.8)	3.2 (15.8)	0.359
Glucagonoma	0.4 (0.8)	1.3 (6.8)	0.269
VIPoma	0.1 (0.2)	0.8 (3.0)	0.062
Other functional PNET	0.2 (1.0)	0.6 (1.9)	0.12
Nonfunctional	22.9 (43.3)	100.1 (295.7)	**0.042**
Not available	1.4 (11.3)	0.9 (4.8)	0.712
Inherited syndrome	2.6 (8.3)	4.2 (11.2)	0.326
Metastatic disease	9.4 (17.3)	43.3 (124.5)	**0.034**
Pathologic features			
ENETS/WHO grading			
G1	33.1 (36.8)	54.1 (89.1)	0.077
G2	18.0 (17.9)	38.8 (50.1)	**0.003**
G3	3.9 (5.6)	8.0 (11.7)	**0.009**
Not available	0.5 (1.7)	34.0 (278.9)	0.342
TNM/UICC staging			
Stage 1	6.1 (14.9)	26.5 (148.4)	0.28
Stage 2	4.1 (13.4)	15.6 (64.6)	0.166
Stage 3	1.7 (4.6)	6.6 (18.8)	**0.047**
Stage 4	2.3 (6.1)	19.1 (121.6)	0.275
Not available	0.1 (0.5)	0.6 (3.0)	0.128

*Data are expressed as: mean (SD). *data are expressed as: mean (SD); percentage.*
The bold values represent significant values. SD Standard deviation, ENETS European Neuroendocrine Tumor Society, WHO World Health Organization, TNM Tumour Node Metastasis, UICC Union for International Cancer Control.

available in 75.0% of studies, with a mean slice thickness of 2.6 mm. Contrast administration details were reported in most studies (81.3%), with 56.3% using contrast bolus tracking. Majority of CT studies were performed on 2 or more different types of CT-equipments (68.8%).

For MRI, most studies reported an imaging acquisition protocol (95.2%, n=20/21), with almost all studies acquiring multiphase contrast-enhanced images (80%, n=16/20), T1- and T2-weighted images (90.5%), and majority of studies obtaining DWI images (76.2%). MRI slice thickness parameters were available in 81.0% of studies, with a mean slice thickness of 3.1 mm. Details of contrast administration were reported in 66.7% of studies. Among studies in which MRI scanner details were reported (n =20/21), half of them were performed on one single MRI scanner, while other half were performed on 2 or more different scanners. For magnet field strength, 33.3% were performed on a 1.5 T system, 28.6% on a 3.0 T system, and 33.3% used both 1.5 and 3.0 T systems.

For PET/CT imaging, the most common radiotracer studied in PNETs is ^{68}Ga-DOTA (31.3%), followed by the GLP-1R agonist (25%), ^{18}F-DOPA (25%), and ^{18}F-FDG (18.6%). Most studies were performed on one scanner (68.8%, n=11/16), while the rest were performed on 2 or more scanners (25%, n=4/16), with 1 study not reporting scanner details. For SPECT/CT, 57.1% of studies evaluated SSTR radiotracers, while 42.9% studied GLP-1R. Most studies reported scanner details (71.4%, n=5/7), with all of them performing it on a single scanner.

There was significant heterogeneity regarding the reporting of imaging modalities and imaging acquisition protocols (**Figure 2G**). For example, 30.6% (n=30/98) of clinical studies did not describe which imaging modalities were used, in contrast to imaging studies, which specified the imaging modality in 100% of cases (n=63/63) (p<0.001). Additionally, 63.3% of clinical studies (n=62/98) reported the imaging modalities used, however, no details on the imaging protocol were reported, while only one imaging study did not report the specific imaging protocol used (1.6%, p<0.001). In only 11.2% (n=11) of clinical studies, injection phase was specified, and in 4.1% (n=4), the multiphase injection phase was clearly specified.

Studies rarely mentioned international guidelines with no clinical study and only four (6.3%) imaging studies referring to guidelines, all of them exclusively referencing ENETS 2009 guidelines (n=4) (16, 42–44) with no studies referencing the most updated 2017 ENETS technical guidelines (37).

Methodological Quality: Risk of Bias Assessment

Imaging Studies

For imaging studies (QUADAS-2: **Figure 2H**), no study was considered to be at low risk of bias for all domains. In 6.3% of included studies, a high risk of bias for patient selection was due to non-consecutive or random enrollment. Regarding the patient flow and timing, 20.6% of imaging studies used a combination of histopathologic findings and clinical follow-up in reference

TABLE 4 | Imaging methods, examination protocols and technical details for each imaging modality.

A. Anatomical imaging		
Modality	**CT**	**MRI**
Number of studies per modality (%)	32/63 (50.8%)	21/63 (33.3%)
Detailed acquisition protocol (%)	30/32 (93.8%)	20/21 (95.2%)
Multiphase contrast-enhanced acquisition (%)	29/30 (96.7%)	16/20 (80.0%)
Single-phase acquisition (%)	1/30 (3.3%)	3/20 (15.0%)
T1-weighted imaging	–	19/21 (90.5%)
T2-weighted imaging	–	19/21 (90.5%)
Diffusion-weighted imaging	–	16/21 (76.2%)
CT-slice thickness available information (%)	24/32 (75%)	17/21 (81.0%)
CT-slice thickness (mm) (mean ± SD)	2.6 (± 2.0)	3.1 (± 1.0)
Details of the contrast agent administration (%)	26/32 (81.3%)	14/21 (66.7%)
Bolus tracking (%)	18/32 (56.3%)	–
Available technical acquisition parameters (%)	21/32 (65.6%)	17/21 (81.0%)
Number of CT/MR-system used		
One equipment (%)	4/32 (12.5%)	10/21 (47.6%)
Two or more equipments (%)	22/32 (68.8%)	10/21 (47.6%)
Not available (%)	6/32 (18.8%)	1/21 (4.8%)
Magnetic Field Strength (Tesla)		
1.5 T-system (%)	–	7/21 (33.3%)
3.0 T-system (%)	–	6/21 (28.6%)
1.5 and 3.0 T-system (%)	–	7/21 (33.3%)
Not available (%)	–	1/21 (4.8%)
B. Molecular imaging		
Modality	**SPECT**	**PET**
Number of studies per modality (%)	7/63 (11.1%)	16/63 (25.4%)
Molecular radiotracer		
^{68}Ga-DOTA (%)	4/7 (57.1%)	5/16 (31.3%)
GLP-1R (%)	3/7 (42.9%)	4/16 (25%)
^{18}F-FDG (%)	–	3/16 (18.6%)
^{18}F-DOPA (%)	–	4/16 (25%)
Radiotracer dose		
^{68}Ga-DOTA (MBq) (mean ± SD)	227.1 (± 100.3)	145.2 (± 33.1)
GLP-1R (MBq) (mean ± SD)	299.8 (± 381.2)	88.8 (± 11.4)
18F-FDG (MBq) (mean ± SD)	–	328.1 (± 106.3)
18F-DOPA (MBq) (mean ± SD)	–	263.8 (± 18.9)
Time before acquisition (min) (mean ± SD)	1181 (± 972.8)	57.5 (± 42.6)
Time acquisition		
Not available (%)	6/7 (85.7%)	–
Available (%)	1/7 (14.3%)	–
Number of systems used		
One equipment (%)	5/7 (71.4%)	11/16 (68.8%)
Two or more equipments (%)	–	4/16 (25%)
Not available (%)	2/7 (28.6%)	1/16 (6.3%)

A: *imaging methods, examination protocol and technical details for each imaging modality: anatomical imaging. CT Computed Tomography, MRI Magnetic Resonance Imaging.*
B: *imaging methods, examination protocol and technical details for each imaging modality: molecular imaging, SPECT Single-Photon Emission Computed Tomography, PET Positron Emission Tomography, SSTR Somatostatin Receptor, 18F-FDG Fluorodeoxyglucose, 18F-DOPA Fluorodeoxyphenylalanine, GLP-1R Glucagonlike Peptide-1 Receptor, SD Standard Deviation.*

standard, introducing a high risk of bias. Lastly, in 4.8% of studies, a high risk of bias was recorded due to a non-blinded nature of reference standard assessment of imaging test results.

Clinical Studies

Using MINORS index for clinical studies (**Figure 2I**) allowed highlighting the fact that the main bias was introduced by the lack of prospective collection of data in 56.1% of the time. The second major bias was the lack of information on the consecutive nature of the inclusion of patients (16.3%). Of note, only a small proportion of the clinical studies was comparative (13.3%, n=15/98), which precluded the possibility to evaluate the four additional criteria (adequate control group, contemporary groups, baseline equivalence of groups, adequate statistical analyses).

DISCUSSION

Medical imaging plays a decisive role in PNETs management, a highly challenging disease (22), and is represented by a large panel of imaging tools available to physicians. With a purpose of standardizing practices, ENETS 2017 guidelines emphases on the need for technical quality of imaging methods (37, 38, 45). To optimize treatment strategies, it is often necessary to combine data from several centers in clinical therapeutic trials. In the new era of big data and artificial intelligence, harmonization of imaging practices is especially important to find relevant imaging biomarkers. This also explains the importance of assessing practice heterogeneity, in order to unravel the potential "imaging databases" that exist in this field. Based on this unmet need, the first

objective of this systematic review was to assess the level of standardization of imaging practices in the recent PNETs literature.

In this study, we demonstrated the existence of a significant lack of standardization and homogenization of methodological imaging practices in the recent PNETs literature. Study selection resulted in 161 imaging-based manuscripts and allowed the creation of two different sub-groups of publications in the final analysis: 63 imaging studies and 98 clinical studies.

The choice of studying each sub-group separately can be explained by our assumptions about the differences in conduction of each type of studies. In imaging studies subgroup, we expected to have all the necessary details because the purpose of these studies is to evaluate diagnostic accuracy. We wanted to assess the degree of homogeneity and compare this information with the international guidelines. For clinical studies in which the therapeutic evaluation is obtained by radiological assessment, we have hypothesized a very small amount of technical details since clinical outcomes were the primary endpoints.

Our study is the first to evaluate imaging standardization in PNETs. Beyond the overall lack of methodological standardization and homogenization, six key concepts were identified in this study.

First, overall methodology quality remains suboptimal. Indeed, the vast majority of the studies was retrospective (n=139/161; 86.3%) and non-randomized (n=155/161; 96.3%). However, there was a significant difference between the two subgroups in terms of institutional design, with multicentric nature in 38.8% of clinical studies, versus less than 5% of imaging studies (p=<0.001). This point may indicate that clinical studies are generally more qualitative, in a methodological point of view. At the opposite, clinical studies were mostly non-blinded, against only 4.8% of imaging studies (p=<0.001), making thus imaging diagnostic accuracy studies' evaluation methodologically valid. While clinical studies are prospective and multicentric, there is limited reporting of and lack of standardization of the imaging acquisition in these studies, which may lead to heterogeneous image quality. Imaging studies have more homogeneous and better described imaging techniques, but the level of evidence is limited by the fact that studies are monocentric and retrospective.

Second, there is a mismatch between types of data used for clinical or imaging studies. While 71.4% of clinical studies started collecting data before 2005 (n=70/98), 53.9% (n=34/63) of imaging studies used data collected after 2005. This point highlights the possible difference between results obtained with clinical studies as compared to imaging ones. Indeed, molecular imaging in the field of NET has been extensively developed this last decade, vastly improving the performance of imaging techniques through more accurate methods, such as Ga-68 DOTATATE PET/CT imaging (46, 47). The two most common imaging modalities reported in the recent PNETs literature are CT and MRI, despite significant progress in nuclear medicine imaging, with the advent of newer high-performance PET radiotracers and its increased availability. Therefore, we predict there will be a future rebalancing in the partition of different PNETs imaging modalities.

Third, geographical distribution of populations in the current literature shows a lack of representation of patients from South America, Africa and Oceania, although the prevalence of PNETs in these parts of the world is not different (48). In other terms, international societies need to encourage research in these countries in order to obtain worldwide results, and better homogenize PNET patients' management, both in clinical routine and for research purposes.

Fourth, imaging procedures are described more frequently and in better detail in imaging studies than in clinical studies, even in large multicentric international clinical studies. Moreover, the radiological assessment is also of better quality in imaging studies, with a significantly higher rate of blinded assessors. In 90.8% of clinical studies, imaging assessment was not clearly stated. Paradoxically, in this study, multicentric international studies, which are supposed to be methodologically better, presented lower quality in terms of radiological methodology. This can be explained by a lack of standardization between each center. For instance, RADIANT-3, a large prospective, randomized, phase 3 clinical trial, published in 2011, demonstrated improvement of everolimus in progression-free survival (PFS) compared with placebo for patients with advanced PNETs (49). Contrasting with the vast majority of the recent clinical studies analyzed, imaging technical details were fully described in their supplementary materials. Since progression-free survival is in part an imaging-based clinical endpoint, this fact confirms that this study is methodologically correct in terms of technical quality and imaging protocol and has a high evidence-based value.

Fifth, adherence to international guidelines is very low in the included studies, as shown by low rates of reference to international imaging technical guidelines (2.5% (n=4/161) of all selected studies). In these 4 cases (published between 2015 and 2018), ENETS 2009 technical guidelines were mentioned. We noticed that ENETS 2017 technical guidelines were never mentioned in the 102 selected articles published since 2017, although it was the most recent and detailed guidelines.

Last, there was a lack of imaging quality assessment tools. Indeed, many tools and indexes are available for methodological quality evaluation of studies and assessment of risks of bias. We have chosen to use MINOR and QUADAS-2 because of their simplicity and their reliability, as demonstrated by the rigorous and evidence-based process to develop these tools. However, neither of these tools were specifically designed to assess how standardized imaging procedures are performed, which can be essential in some areas. Therefore, it seems important that future work focuses on a methodological quality assessment tool that incorporates the evaluation of how imaging techniques are performed.

Based on a systematic review and meta-analysis approach, using strict inclusion criteria, we applied state-of-the-art methodology in this research. We have chosen to restrict our search strategy to the last five years. Indeed, we focused on imaging technical parameters, a field of medicine and technology that is constantly evolving and changing. For example, thin CT sections were not routinely systematic before 2009 ENETS technical guidelines.

These results showed the difficulty of pooling all data for a big data approach. Qualitative assessment of potential "imaging

databases", theoretically accessible to Datamining using AI in recent PNETs literature shows an excessive data heterogeneity. This is exacerbated by the use of many different machines and equipment, which increases input data variability. Initiatives like the EARL FDG PET/CT accreditation program provide a way to limit the data heterogeneity and facilitate multicenter research projects with accurate and reproducible imaging data.

Results expressed in this study might have major implications for clinicians, researchers, and guideline committees. Clinical decisions should be based on the best available imaging technique, using rigorously the recommended technical properties for each technique. A non-optimal imaging acquisition or reconstruction should be repeated before taking any clinical decision.

Similarly, precision should be requested in imaging-based studies. In addition, as a quality guarantee, affirmation of the use of imaging examinations in accordance with reference guidelines should be at least mentioned before envisioning any future publication. Another approach to improve practices would be to modify prospective databases from which a majority of studies collect their information. Technical imaging data, radiological protocols and acquisition methods should be mentioned, and only patients who have benefited from appropriate imaging examinations in accordance with international guidelines should be included. A proposal to expand these databases to include imaging technical information would also allow better selection of patients with technically correct imaging.

Herein, this systematic review of the recent literature on PNETs, with a special emphasis on imaging, demonstrates the lack of rigorous reporting and standardization of imaging techniques in clinical practice and research. Indeed, a clear gap in imaging information in clinical studies was demonstrated, particularly for types of modalities used, radiological protocol applied, and imaging assessment. This lack of information seems more intriguing, when it comes to clinical studies whose results are mainly based on radiological evaluation. Even when technical details were available in imaging studies, there is a significant heterogeneity of practices and a lack of references to established international guidelines. This non-uniformity makes it difficult to envision a pooled use of data for AI datamining and big data purposes since AI needs absolute homogeneity and standardization of clinical practices, that will perhaps allow identifying new biomarkers for treatment effectiveness, and thus a higher optimization of PNETs patients' management.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The study was designed, directed and coordinated by F-ZM and LD. EP conducted the study, wrote the manuscript and designed the figures. LD and RY performed the statistics LR and LOD analyzed nuclear medicine data. TE, RY, NC, RG, and HR reviewed the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.628408/full#supplementary-material

Supplementary Table 1 | Available international guidelines for PNETs technical imaging.

Supplementary Table 2 | PRISMA 2009 checklist.

Supplementary Table 3 | Search strategies.

REFERENCES

1. Smith JK, Ng SC, Hill JS, Simons JP, Arous EJ, Shah SA, et al. Complications After Pancreatectomy for Neuroendocrine Tumors: A National Study. *J Surg Res* (2010) 163:63–8. doi: 10.1016/j.jss.2010.04.017
2. Canellas R, Lo G, Bhowmik S, Ferrone C, Sahani D. Pancreatic Neuroendocrine Tumor: Correlations Between MRI Features, Tumor Biology, and Clinical Outcome After Surgery. *J Magn Reson Imaging* (2018) 47:425–32. doi: 10.1002/jmri.25756
3. Bilimoria KY, Tomlinson JS, Merkow RP, Stewart AK, Ko CY, Talamonti MS, et al. Clinicopathologic Features and Treatment Trends of Pancreatic Neuroendocrine Tumors: Analysis of 9,821 Patients. *J Gastrointest Surg* (2007) 11:1460–9. doi: 10.1007/s11605-007-0263-3
4. Strosberg JR, Cheema A, Weber JM, Ghayouri M, Han G, Hodul PJ, et al. Relapse-Free Survival in Patients With Nonmetastatic, Surgically Resected Pancreatic Neuroendocrine Tumors: An Analysis of the AJCC and ENETS Staging Classifications. *Ann Surg* (2012) 256:321–5. doi: 10.1097/SLA.0b013e31824e6108
5. Jensen RT, Berna MJ, Bingham DB, Norton JA. Inherited Pancreatic Endocrine Tumor Syndromes: Advances in Molecular Pathogenesis, Diagnosis, Management, and Controversies. *Cancer* (2008) 113:1807–43. doi: 10.1002/cncr.23648
6. Hashim YM, Trinkaus KM, Linehan DC, Strasberg SS, Fields RC, Cao D, et al. Regional Lymphadenectomy Is Indicated in the Surgical Treatment of Pancreatic Neuroendocrine Tumors (PNETs). *Ann Surg* (2014) 259:197–203. doi: 10.1097/SLA.0000000000000348
7. Sahani DV, Bonaffini PA, Fernández–Del Castillo C, Blake MA. Gastroenteropancreatic Neuroendocrine Tumors: Role of Imaging in Diagnosis and Management. *Radiology* (2013) 266:38–61. doi: 10.1148/radiol.12112512
8. Kocha W, Maroun J, Kennecke H, Law C, Metrakos P, Ouellet JF, et al. Consensus Recommendations for the Diagnosis and Management of Well-Differentiated Gastroenterohepatic Neuroendocrine Tumours: A Revised Statement From a Canadian National Expert Group. *Curr Oncol* (2010) 17:49–64. doi: 10.3747/co.v17i3.484
9. Horton KM, Hruban RH, Yeo C, Fishman EK. Multi–Detector Row CT of Pancreatic Islet Cell Tumors. *RadioGraphics* (2006) 26:453–64. doi: 10.1148/rg.262055056
10. Okabe H, Hashimoto D, Chikamoto A, Yoshida M, Taki K, Arima K, et al. Shape and Enhancement Characteristics of Pancreatic Neuroendocrine Tumor on Preoperative Contrast-Enhanced Computed Tomography May

be Prognostic Indicators. *Ann Surg Oncol* (2017) 24:1399–405. doi: 10.1245/s10434-016-5630-4
11. Rozenblum L, Mokrane F-Z, Yeh R, Sinigaglia M, Besson F, Seban R-D, et al. The Role of Multimodal Imaging in Guiding Resectability and Cytoreduction in Pancreatic Neuroendocrine Tumors: Focus on PET and MRI. *Abdom Radiol* (2019) 44:2474–93. doi: 10.1007/s00261-019-01994-5
12. Lee L, Ito T, Jensen RT. Imaging of Pancreatic Neuroendocrine Tumors: Recent Advances, Current Status, and Controversies. *Expert Rev Anticancer Ther* (2018) 18:837–60. doi: 10.1080/14737140.2018.1496822
13. Semelka RC, Custodio CM, Cem Balci N, Woosley JT. Neuroendocrine Tumors of the Pancreas: Spectrum of Appearances on MRI. *J Magn Reson Imaging* (2000) 11:141–8. doi: 10.1002/(sici)1522-2586(200002)11:2<141::aid-jmri10>3.0.co;2-u
14. Kulali F, Semiz-Oysu A, Demir M, Segmen-Yilmaz M, Bukte Y. Role of Diffusion-Weighted MR Imaging in Predicting the Grade of Nonfunctional Pancreatic Neuroendocrine Tumors. *Diagn Interv Imaging* (2018) 99:301–9. doi: 10.1016/j.diii.2017.10.012
15. Breeman WA, de Blois E, Chan HS, Konijnenberg M, Kwekkeboom DJ, Krenning EP. 68Ga-Labeled DOTA-Peptides and 68Ga-Labeled Radiopharmaceuticals for Positron Emission Tomography: Current Status of Research, Clinical Applications, and Future Perspectives. *Semin Nucl Med* (2011) 41(4):314–21. doi: 10.1053/j.semnuclmed.2011.02.001
16. Imperiale A, Sebag F, Vix M, Castinetti F, Kessler L, Moreau F, et al. 18F-FDOPA PET/CT Imaging of Insulinoma Revisited. *Eur J Nucl Med Mol Imaging* (2015) 42:409–18. doi: 10.1007/s00259-014-2943-z
17. Binderup T, Knigge U, Loft A, Federspiel B, Kjaer A. 18F-Fluorodeoxyglucose Positron Emission Tomography Predicts Survival of Patients With Neuroendocrine Tumors. *Clin Cancer Res* (2010) 16:978–85. doi: 10.1158/1078-0432.CCR-09-1759
18. Bozkurt MF, Virgolini I, Balogova S, Beheshti M, Rubello D, Decristoforo C, et al. Guideline for PET/CT Imaging of Neuroendocrine Neoplasms With 68Ga-DOTA-Conjugated Somatostatin Receptor Targeting Peptides and 18F–DOPA. *Eur J Nucl Med Mol Imaging* (2017) 44:1588–601. doi: 10.1007/s00259-017-3728-y
19. Sadowski SM, Neychev V, Millo C, Shih J, Nilubol N, Herscovitch P, et al. Prospective Study of ^{68} Ga-DOTATATE Positron Emission Tomography/Computed Tomography for Detecting Gastro-Entero-Pancreatic Neuroendocrine Tumors and Unknown Primary Sites. *JCO* (2016) 34:588–96. doi: 10.1200/JCO.2015.64.0987
20. Buchmann I, Henze M, Engelbrecht S, Eisenhut M, Runz A, Schäfer M, et al. Comparison of 68Ga-DOTATOC PET and 111In-DTPAOC (Octreoscan) SPECT in Patients With Neuroendocrine Tumours. *Eur J Nucl Med Mol Imaging* (2007) 34:1617–26. doi: 10.1007/s00259-007-0450-1
21. Bushnell DL, Baum RP. Standard Imaging Techniques for Neuroendocrine Tumors. *Endocrinol Metab Clin North Am* (2011) 40:153–62. doi: 10.1016/j.ecl.2010.12.002
22. Rozenblum L, Mokrane F-Z, Yeh R, Sinigaglia M, Besson FL, Seban R-D, et al. Imaging-Guided Precision Medicine in Non-Resectable Gastro-Entero-Pancreatic Neuroendocrine Tumors: A Step-by-Step Approach. *Eur J Radiol* (2020) 122:108743. doi: 10.1016/j.ejrad.2019.108743
23. Sinigaglia M, Assi T, Besson FL, Ammari S, Edjlali M, Feltus W, et al. Imaging-Guided Precision Medicine in Glioblastoma Patients Treated With Immune Checkpoint Modulators: Research Trend and Future Directions in the Field of Imaging Biomarkers and Artificial Intelligence. *EJNMMI Res* (2019) 9:78. doi: 10.1186/s13550-019-0542-5
24. Dromain C, de Baere T, Lumbroso J, Caillet H, Laplanche A, Boige V, et al. Detection of Liver Metastases From Endocrine Tumors: A Prospective Comparison of Somatostatin Receptor Scintigraphy, Computed Tomography, and Magnetic Resonance Imaging. *J Clin Oncol* (2004) 23:70–8. doi: 10.1200/JCO.2005.01.013
25. Martins D, Spada F, Lambrescu I, Rubino M, Cella C, Gibelli B, et al. Predictive Markers of Response to Everolimus and Sunitinib in Neuroendocrine Tumors. *Target Oncol* (2017) 12:611–22. doi: 10.1007/s11523-017-0506-5
26. Garcia-Carbonero R, Garcia-Figueiras R, Carmona-Bayonas A, Sevilla I, Teule A, Quindos M, et al. Imaging Approaches to Assess the Therapeutic Response of Gastroenteropancreatic Neuroendocrine Tumors (GEP-NETs): Current Perspectives and Future Trends of an Exciting Field in Development. *Cancer Metastasis Rev* (2015) 34:823–42. doi: 10.1007/s10555-015-9598-5
27. Patel SP, Othus M, Chae YK, Giles FJ, Hansel DE, Singh PP, et al. A Phase II Basket Trial of Dual Anti–CTLA-4 and Anti–PD-1 Blockade in Rare Tumors (DART SWOG 1609) in Patients With Nonpancreatic Neuroendocrine Tumors. *Clin Cancer Res* (2020) 26:2290–6. doi: 10.1158/1078-0432.CCR-19-3356
28. Yao JC, Strosberg J, Fazio N, Pavel ME, Ruszniewski P, Bergsland E, et al. Activity & Safety of Spartalizumab (PDR001) in Patients (Pts) With Advanced Neuroendocrine Tumors (NET) of Pancreatic (Pan), Gastrointestinal (GI), or Thoracic (T) Origin, & Gastroenteropancreatic Neuroendocrine Carcinoma (GEP NEC) Who Have Progressed on Prior Treatment (Tx). *Ann Oncol* (2018) 29:viii467–8. doi: 10.1093/annonc/mdy293.001
29. Dercle L, Mokrane F-Z, de Colella JMS, Stamatoullas A, Morschhauser F, Brice P, et al. Unconventional Immune-Related Phenomena Observed Using 18F-FDG PET/CT in Hodgkin Lymphoma Treated With Anti PD-1 Monoclonal Antibodies. *Eur J Nucl Med Mol Imaging* (2019) 46:1391–2. doi: 10.1007/s00259-019-04310-x
30. Dercle L, Seban R-D, Lazarovici J, Schwartz LH, Houot R, Ammari S, et al. 18F-FDG PET and CT Scans Detect New Imaging Patterns of Response and Progression in Patients With Hodgkin Lymphoma Treated by Anti–Programmed Death 1 Immune Checkpoint Inhibitor. *J Nucl Med* (2018) 59:15–24. doi: 10.2967/jnumed.117.193011
31. Dercle L, Ammari S, Seban R-D, Schwartz LH, Houot R, Labaied N, et al. Kinetics and Nadir of Responses to Immune Checkpoint Blockade by Anti-PD1 in Patients With Classical Hodgkin Lymphoma. *Eur J Cancer* (2018) 91:136–44. doi: 10.1016/j.ejca.2017.12.015
32. Champiat S, Dercle L, Ammari S, Massard C, Hollebecque A, Postel-Vinay S, et al. Hyperprogressive Disease Is a New Pattern of Progression in Cancer Patients Treated by Anti-PD-1/PD-L1. *Clin Cancer Res* (2017) 23:1920–8. doi: 10.1158/1078-0432.CCR-16-1741
33. Dercle L, Ammari S, Champiat S, Massard C, Ferté C, Taihi L, et al. Rapid and Objective CT Scan Prognostic Scoring Identifies Metastatic Patients With Long-Term Clinical Benefit on Anti-PD-1/-L1 Therapy. *Eur J Cancer* (2016) 65:33–42. doi: 10.1016/j.ejca.2016.05.031
34. Dercle L, Chisin R, Ammari S, Gillebert Q, Ouali M, Jaudet C, et al. Nonsurgical Giant Cell Tumour of the Tendon Sheath or of the Diffuse Type: Are MRI or 18 F-FDG PET/CT Able to Provide an Accurate Prediction of Long-Term Outcome? *Eur J Nucl Med Mol Imaging* (2015) 42:397–408. doi: 10.1007/s00259-014-2938-9
35. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: Guidelines for Response Criteria for Use in Trials Testing Immunotherapeutics. *Lancet Oncol* (2017) 18:e143–52. doi: 10.1016/S1470-2045(17)30074-8
36. Kuo EJ, Salem RR. Population-Level Analysis of Pancreatic Neuroendocrine Tumors 2 cm or Less in Size. *Ann Surg Oncol* (2013) 20:2815–21. doi: 10.1245/s10434-013-3005-7
37. Sundin A, Arnold R, Baudin E, Cwikla JB, Eriksson B, Fanti S, et al. ENETS Consensus Guidelines for the Standards of Care in Neuroendocrine Tumors: Radiological, Nuclear Medicine and Hybrid Imaging. *Neuroendocrinology* (2017) 105:212–44. doi: 10.1159/000471879
38. Kwekkeboom DJ, Krenning EP, Scheidhauer K, Lewington V, Lebtahi R, Grossman A, et al. ENETS Consensus Guidelines for the Standards of Care in Neuroendocrine Tumors: Somatostatin Receptor Imaging With 111In-Pentetreotide. *Neuroendocrinology* (2009) 90:184–9. doi: 10.1159/000225946
39. PRISMA-P Group, Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 Statement. *Syst Rev* (2015) 4:1. doi: 10.1186/2046-4053-4-1
40. Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* (2011) 155:529. doi: 10.7326/0003-4819-155-8-201110180-00009
41. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological Index for Non-Randomized Studies (MINORS): Development and Validation of a New Instrument. *ANZ J Surg* (2003) 73:712–6. doi: 10.1046/j.1445-2197.2003.02748.x
42. McGovern JM, Singhi AD, Borhani AA, Furlan A, McGrath KM, Zeh HJIII, et al. CT Radiogenomic Characterization of the Alternative Lengthening of Telomeres Phenotype in Pancreatic Neuroendocrine Tumors. *Am J Roentgenol* (2018) 211(5):1020–5. doi: 10.2214/AJR.17.19490

43. De Robertis R, D'Onofrio M, Zamboni G, Martini PT, Gobbo S, Capelli P, et al. Pancreatic Neuroendocrine Neoplasms: Clinical Value of Diffusion-Weighted Imaging. *Neuroendocrinology* (2016) 103:758–70. doi: 10.1159/000442984
44. Jilesen APJ, Hoefnagel SJM, Busch ORC, Bennink RJ, Gouma DJ, van Dijkum EN. The Influence of Somatostatin Receptor Scintigraphy During Preoperative Staging of Non-Functioning Pancreatic Neuroendocrine Tumours. *Clin Radiol* (2016) 71:537–42. doi: 10.1016/j.crad.2016.01.015
45. Sundin A, Vullierme M-P, Kaltsas G, Plockinger U. ENETS Consensus Guidelines for the Standards of Care in Neuroendocrine Tumors: Radiological Examinations. *Neuroendocrinology* (2009) 90:167–83. doi: 10.1159/000184855
46. Hofman MS, Lau WE, Hicks RJ. Somatostatin Receptor Imaging With 68Ga DOTATATE PET/CT: Clinical Utility, Normal Patterns, Pearls, and Pitfalls in Interpretation. *Radiographics* (2015) 35:500–16. doi: 10.1148/rg.352140164
47. Chan DL, Pavlakis N, Schembri GP, Bernard EJ, Hsiao E, Hayes A, et al. Dual Somatostatin Receptor/FDG PET/CT Imaging in Metastatic Neuroendocrine Tumours: Proposal for a Novel Grading Scheme With Prognostic Significance. *Theranostics* (2017) 7:1149. doi: 10.7150/thno.18068
48. Gunavathy M, Rohana AG, Norlela S, Nor Azmi K. A Ten-Year Retrospective Analysis of Gastroenteropancreatic Neuroendocrine Tumors (GEP-NETs) in Malaysia. *Med J Malaysia* (2014) 69:133–7.
49. Yao JC, Shah MH, Ito T, Bohas CL, Wolin EM, Van Cutsem E, et al. Everolimus for Advanced Pancreatic Neuroendocrine Tumors. *N Engl J Med* (2011) 364:514–23. doi: 10.1056/NEJMoa1009290

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 24 August 2021
doi: 10.3389/fonc.2021.692973

The Impact of Artificial Intelligence CNN Based Denoising on FDG PET Radiomics

Cyril Jaudet[1]*, *Kathleen Weyts*[2], *Alexis Lechervy*[3], *Alain Batalla*[1], *Stéphane Bardet*[2] *and Aurélien Corroyer-Dulmont*[1,4]*

[1] Medical Physics Department, CLCC François Baclesse, Caen, France, [2] Nuclear Medicine Department, CLCC François Baclesse, Caen, France, [3] UMR GREYC, Normandie Univ, UNICAEN, ENSICAEN, CNRS, Caen, France, [4] Normandie Univ, UNICAEN, CEA, CNRS, ISTCT/CERVOxy group, GIP CYCERON, Caen, France

Background: With a constantly increasing number of diagnostic images performed each year, Artificial Intelligence (AI) denoising methods offer an opportunity to respond to the growing demand. However, it may affect information in the image in an unknown manner. This study quantifies the effect of AI-based denoising on FDG PET textural information in comparison to a convolution with a standard gaussian postfilter (EARL1).

Methods: The study was carried out on 113 patients who underwent a digital FDG PET/CT (VEREOS, Philips Healthcare). 101 FDG avid lesions were segmented semi-automatically by a nuclear medicine physician. VOIs in the liver and lung as reference organs were contoured. PET textural features were extracted with pyradiomics. Texture features from AI denoised and EARL1 *versus* original PET images were compared with a Concordance Correlation Coefficient (CCC). Features with CCC values ≥ 0.85 threshold were considered concordant. Scatter plots of variable pairs with R2 coefficients of the more relevant features were computed. A Wilcoxon signed rank test to compare the absolute values between AI denoised and original images was performed.

Results: The ratio of concordant features was 90/104 (86.5%) in AI denoised *versus* 46/104 (44.2%) with EARL1 denoising. In the reference organs, the concordant ratio for AI and EARL1 denoised images was low, respectively 12/104 (11.5%) and 7/104 (6.7%) in the liver, 26/104 (25%) and 24/104 (23.1%) in the lung. SUVpeak was stable after the application of both algorithms in comparison to SUVmax. Scatter plots of variable pairs showed that AI filtering affected more lower *versus* high intensity regions unlike EARL1 gaussian post filters, affecting both in a similar way. In lesions, the majority of texture features 79/100 (79%) were significantly ($p<0.05$) different between AI denoised and original PET images.

Conclusions: Applying an AI-based denoising on FDG PET images maintains most of the lesion's texture information in contrast to EARL1-compatible Gaussian filter. Predictive features of a trained model could be thus the same, however with an adapted threshold. Artificial intelligence based denoising in PET is a very promising approach as it adapts the denoising in function of the tissue type, preserving information where it should.

Keywords: denoising, AI, PET, radiomics, medical imaging, convolutional neural network, VEREOS

OPEN ACCESS

Edited by:
Laurent Dercle, Columbia University Irving Medical Center, United States

Reviewed by:
Suyash P. Awate, Indian Institute of Technology Bombay, India
Ziren Kong, Chinese Academy of Medical Sciences, China

***Correspondence:**
Cyril Jaudet
c.jaudet@baclesse.unicancer.fr
Aurélien Corroyer-Dulmont
a.corroyerdulmont@baclesse.unicancer.fr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *09 April 2021*
Accepted: *26 July 2021*
Published: *24 August 2021*

Citation:
Jaudet C, Weyts K, Lechervy A, Batalla A, Bardet S and Corroyer-Dulmont A (2021) The Impact of Artificial Intelligence CNN Based Denoising on FDG PET Radiomics. Front. Oncol. 11:692973. doi: 10.3389/fonc.2021.692973

INTRODUCTION

Imaging modalities are nowadays an essential diagnostic tool in medicine. From 2009 to 2019 the number of exams in the USA has increased by about 18%, 42% and 105% for CT, MRI and PET respectively (1). This increasing demand has exceeded the actual offer leading to unreasonable delay, weeks or even months for MRI and PET scans in France/Europe (2). An appropriate image denoising may help to reduce scanning time or even reduce injected dose for PET. It may allow to increase the number of examinations without impacting too much working hours or requiring the installation of new medical imaging devices. Deep learning as a subdivision of artificial intelligence (AI) allows to build promising denoising models.

We focused on PET imaging as it will benefit of denoising because of its long scanning time. Although many studies are actually investigating the clinical performance of this method, it may also impact other emerging fields such as imaging based predictive models, radiomics and other AI applications (3).

Medical images are basically a visual representation of different grey levels based on their density (CT), magnetic properties (MRI) or functional information (PET/SPECT). The distribution of the grey values characterizes the heterogeneity of the information. A fast-evolving field called radiomics provide a methodology to extract different features based on intensity, shape, texture from images in order to build predictive models (4). This approach holds great promises as being able to predict patient outcomes. They might allow personalized treatment. As an example, an overall survival predictive model including radiomics features was computed in lung cancer (5) This field is increasing with an annual growth rate of published papers of 177.82% between 2013 and 2018 (6). The models are very promising but there are still some efforts to be made to translate and implement them in a routine clinical setting (7).

Artificial intelligence is in the early phase of application in medical imaging. In this article, we used deep learning and more specifically convolutional neural network approaches which represent a subdivision of AI techniques. Today deep learning has a key role in image reconstruction, processing (denoising, segmentation), analysis and predictive modelling. These applications will develop even more in the future (8). In most of these tasks, they often outperformed a more traditional approach (9). A comparison of this type of AI based denoising algorithm on a PET/MR with clinical data show an increase of the contrast over noise ratio by $46.80 \pm 25.23\%$ compared to $18.16 \pm 10.02\%$ for a Gaussian filter only (10)]. Other methods studied in (10) like guided nonlocal means, block matching 4D or deep decoder improve the CNR oby$24.35 \pm 16.30\%$, $38.31 \pm 20.26\%$ and $41.67 \pm 22.28\%$ respectively. Denoising may also be performed during reconstruction, however this cannot be implemented on an existing machine. The most important limitation is the lack of FDA or CE certification of all those approaches. We focus our study on Subtle PET™ (Subtle Medical, Stanford, USA provided by Incepto, France). It is a post-processing FDA and CE approved denoising software for FDG PET (11), based on convolutional neural networks (CNN), the most common deep learning architecture for image processing.

AI denoising and radiomics are two very promising fields in medical imaging. However, we are the first, to the best of our knowledge, to try to combine these two approaches for PET Imaging. We question whether a radiomics model using PET [^{18}F] FDG trained on classical data is still valid after an AI denoising method. This study measured the stability of basic and radiomics PET features in lesions and normal reference organs when applying an AI denoising solution. We also wanted to provide an intuitive understanding on how images are affected by AI compared to a reference gaussian post filter routinely used in our center to generate EARL1 compatible PET series.

MATERIALS AND METHODS

This retrospective study was approved by the local institution review board. 113 patients referred to our oncological institution for an initial or follow-up [^{18}F] FDG PET/CT exam between January and March 2020 were retrospectively included. We obtained an informed consent (non-opposition) from all patients. This observational study was in line with MR 004, a national French institution (INDS) defining health research conduct guidelines. The study population characteristics are shown in **Table 1**.

Our PET center is accredited by EANM research limited (EARL) (12) and EANM imaging guidelines (13) were respected. The patients were injected with 4MBq/kg of [^{18}F] FDG IV. PET images from skull base to mid-thighs were acquired on a digital

TABLE 1 | Description of the patient cohort.

Patients (N)		**133**	**Number**
Sex		68%	female%
Age(Y)		61.5±13.5	mean±SD
		[24-89]	[range]
Weight (kg)		74±16	mean±SD
		[35-110]	[range]
BMI(kg/m2)		27±6	mean±SD
		[15-42]	[range]
Indication			
Oncologic	95 (84%)		Number (%)
	Breast	36 (32%)	
	Lung	17 (15%)	
	Gynaecologic (except breast)	14 (12%)	
	other malignancies (lymphoma, anal, colorectal, bladder, thyroid, head and neck, melanoma, myeloma) or mixed	28 (25%)	
Diagnostic benign versus malignant	14 (12%)		
Miscellaneous	5 (4%)		

PET/CT (VEREOS 2018, Philips Healthcare) during 1min/bed position. Once acquired, PET images were reconstructed with an 3D OSEM algorithm, 4 iterations, 4 subsets with point spread function (PSF) correction. Scatter and attenuation correction was applied. The spacing and matrix size were respectively of 2x2x2 mm^3 and 288x288 pixels. An EARL1 reconstruction was also generated with the same parameters but convolved with a gaussian post filter of 7.2mm. CT scan parameters were 100-140 kV (BMI adaptive), with variable mAs according to an index dose right of 14 and an iterative reconstruction I dose 4; 64x0.625mm slice collimation, pitch of 0.83, rotation time 0.5 s, 3D modulation, matrix 512x512 and voxel size 0.97x0.97x 3 mm^3. The PET mean dose was 5.32 mSv for a patient of 70 Kg. CT had a CTDI median value of 4.8 mGy and a DLP of 431.5 mGy.cm.

The originally reconstructed PET images (with PSF modelling) were denoised with a convolutional neural network (CNN) approach by a commercially available software, Subtle PET® by Subtle Medical. SubtlePET™ uses a multi-slice 2.5D (5 slices) encoder-decoder U-Net DCNN to perform denoising. The software takes a low count PET image as the input and generates a high-quality PET image (close to full dose image) as the output. Accreditation from FDA and CE required robustness. The denoising model was trained on PET images from different centers and vendors. It employs a CNN-based method in a pixel's neighborhood to reduce noise and increase image quality. Using a residual learning approach and optimized for quantitative (L1 norm) as well as structural similarity (SSIM), the software learns to separate and suppress the noise components while preserving and enhancing non-noise components. The images were directly sent from the PET console to a specific local server. Once transferred they were anonymized, denoised, deanonymized and pushed back to a clinical viewer. The mean treatment time was 45 s on a NVIDIA 1080 GPU processor.

All contours were performed in 3D slicer version 4.10 (14) on original PSF PET images and copied on AI denoised and EARL1 PET series. Spherical volumes of interest (VOI) were drawn in the reference organs: liver (3 cm radius, avoiding upper parts, tissue boundaries and major vessels) and lung (1.5 cm radius, drawn in the upper parts). Up to five FDG avid lesions per patient (including only the most intense ones), in total 101 lesions, were segmented by an experienced nuclear medicine physician. Segmented lesions consisted only of authentic malignant primary and metastatic lesions in solid tumors or lymphoma. A semi-automatic tool was employed to segment lesions. A VOI was created by clicking on the original PET image. This 3DSlicer module (PETTumors Segmentation) is based on a highly automated optimal surface segmentation approach, which is a variant of the layered optimal graph image segmentation of multiple objects and surfaces segmentation (15). The VOI was than inspected and manually adjusted with a brush if needed. An automatic donut of 2 voxels diameters was grown around the lesion to calculate the lesion over background ratio. The mean analyzed metabolic volume was 20 (1-162) ml. The same VOI were used for original, AI denoised and EARL1like images.

The extraction of radiomics features was automatically carried out with the pyradiomics package (16) thus mostly complying with the Image Biomarker Standardisation Initiative (17). Images had a native isotropic spacing of 2x2x2 mm^3 so an interpolation step was not necessary. As there is no consensus about the intensity discretization, a fixed bin number of 64 was used (18). A python code using simpleITK (19) was developed to extract all the radiomics features and is accessible in the supplementary information. Eight groups of radiomics features were computed. The intensity class contains first-order data, describing the distribution of voxel intensities within the image region defined by the VOI. They are commonly used and basic images metrics. The shape class is constituted of the 3D size and shape of the VOI. These shape features were excluded as the VOI was the same in all the images. A Grey Level Co-occurrence Matrix (GLCM) class describes the second-order joint probability function of an image region. Grey Level Size Zone Matrix (GLSZM) features quantify grey level zones in an image. A grey level zone is defined as the number of connected voxels that share the same grey level intensity (3D). The Grey Level Run Length Matrix (GLRLM) class testifies of grey level runs, which are defined as the length in number of pixels, of consecutive pixels that have the same grey level value (1D). Neighboring Grey Tone Difference Matrix (NGTDM) is a descriptor of the difference between a grey value and the average grey value of neighbors. A Grey Level Dependence Matrix (GLDM) characterizes the number of connected voxels within a distance from the center voxel in function of their grey level. Most features used in this study are in compliance with Imaging Biomarker Standardization Initiative (IBSI) (IBSI reference manual).IQ wavelets class contains two features, a local analyzing just the VOI and a global of the whole image. These metrics characterize image quality as the ratio between high and low wavelet frequencies.

The Concordance Correlation Coefficients (CCC) (20) were evaluated comparing the post processing IA denoised and EARL1 images to the original PET. CCC values of +/-1 describe a perfect positive/negative correlation and 0 no correlation. Features with a minimum CCC of 0.85 were considered as statistically reproducible and concordant (21). Scatter plots of variable pairs with R^2 value was displayed for the coefficient of variation (CV) and mean SUV values to understand the difference of CCC's in lesions and in liver when an AI denoising or EARL1 filter are used on original images. Mean SUV in lesions is presented using boxplots with minimum, maximum, 1st quartile and 3rd quartile to highlight the difference between the 3 series. A paired Wilcoxon signed rank test was used to compare features in original and AI denoised, and original and EARL1 images. P-values <0.05 were considered statistically significant. All the statistical analyzes were performed using python (22) and scipy.stats library. All the data and the python code of the analysis are available on https://github.com/AurelienCD/RadiomicsIA_PET_Depository_Manuscript-ID-692973.

RESULTS

A visual comparison of AI denoised (B) *versus* original images (A) shows that the AI approach seems to decrease noise in healthy tissues while preserving the intensity distribution in the

lesion in **Figure 1**. In the EARL1-PET image (C) background noise is reduced, but also in the lesion the uptake intensity and distribution are affected. Similar observations can be deduced from the second patient's images (**Figures 1D–F**).

The concordance correlation coefficient (CCC) testifying of the stability of the features comparing denoised to original images is presented in **Figure 2**. In lesions, 90/104 (86.5%) with AI and 46/104 (44%) with EARL1 denoising stayed stable. All stable features in the EARL1 images were also stable in AI images. For the basic intensity class parameters, SUVpeak, SUVmean and SUVmedian kept a CCC≥0.85 in the two denoising approaches. SUVmax and SUVmin CCC values stayed stable for the AI denoised images in the lesions, but fell below the significant threshold for EARL1 images. The NGTDM features were less affected by both denoising methods. In the reference organs, for AI and EARL1 respectively, 12/104 (11.5%) and 7/104 (6.7%) in liver and 26/104 (25%) and 24/104 (23.1%) in lung had a CCC value at least of 0.85. The majority of the features in reference organs are less stable then in lesions for the two denoising methods. For the basic intensity parameters, SUV mean was overall stable for both denoising methods while SUV peak in both liver and lung for AI denoising, *versus* only in the lung for EARL1.

Concerning AI denoising, CV values in lesions before and after processing were very similar. In the lesions the values were slightly below and parallel to the identity line with R^2 = 0.992. EARL1 showed a lower correlation and greater distance from the identity line(**Figure 3A**). In healthy liver (**Figure 3B**), the behavior was different. CV was reduced by a magnitude order of 2 for both denoising methods. With IA denoising, the points were also more scattered for liver (R^2 = 0.884) than for lesions (R^2 = 0.992). EARL1 denoising showed less differences (R^2 = 0.851 *vs* 0.893). The SUV mean value displayed in **Figures 3C, D** showed high correlation in lesions as well as in the healthy tissue. In **Figure 3D** SUV mean in liver is not modified by a EARL1 gaussian postfilter (R^2 = 1). Scatter plots of variable pairs for all the features are accessible in the supplementary materials.

Figure 4 testifies of the difference of SUV mean in lesions between AI and EARL1 denoising compared to the original images. AI denoising not significantly modified the SUV mean values with a p=0.06. EARL1 post filter led to a significantly lower mean SUV in lesions (p<0.001).

The results of the paired Wilcoxon signed rank test between original and AI denoised images are presented in **Table 2**. Almost all the features 79/100 (79%) were significantly different. Wavelets features were not studied. In the intensity class only 4/27 were not significantly different. SUV mean and median values were not significantly different between the AI denoised image and the original one. **Table 2** shows in blue the 18 features that had a CCC>0.85 and were not significantly different.

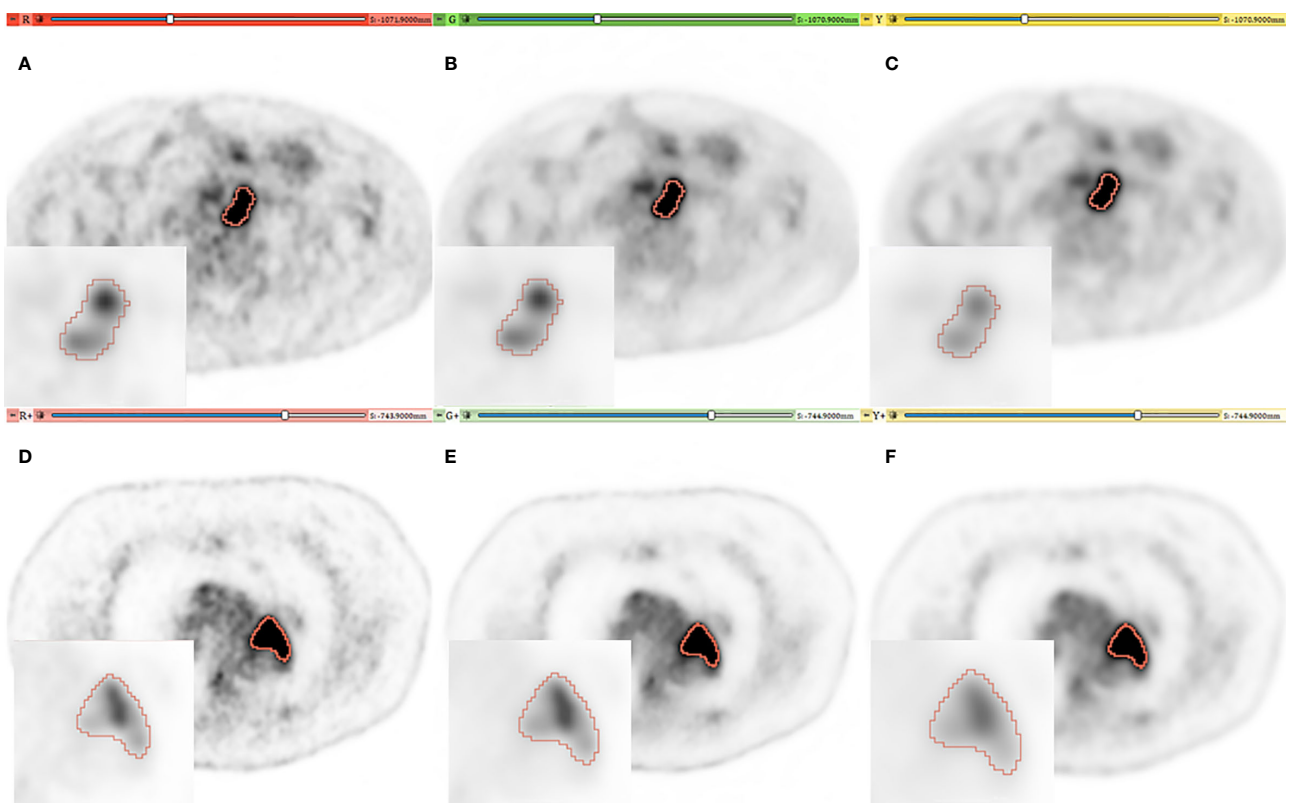


FIGURE 1 | Representative PET imaging of two lesions in different patients with a SUV windowing of (0–5). **(A, D)** [red], **(B, E)** [green], **(C, F)** [yellow] for original, AI and EARL images respectively. A Zoom is added on each image with SUV windowing between (0–25) and (0–30) for the first and second patient.

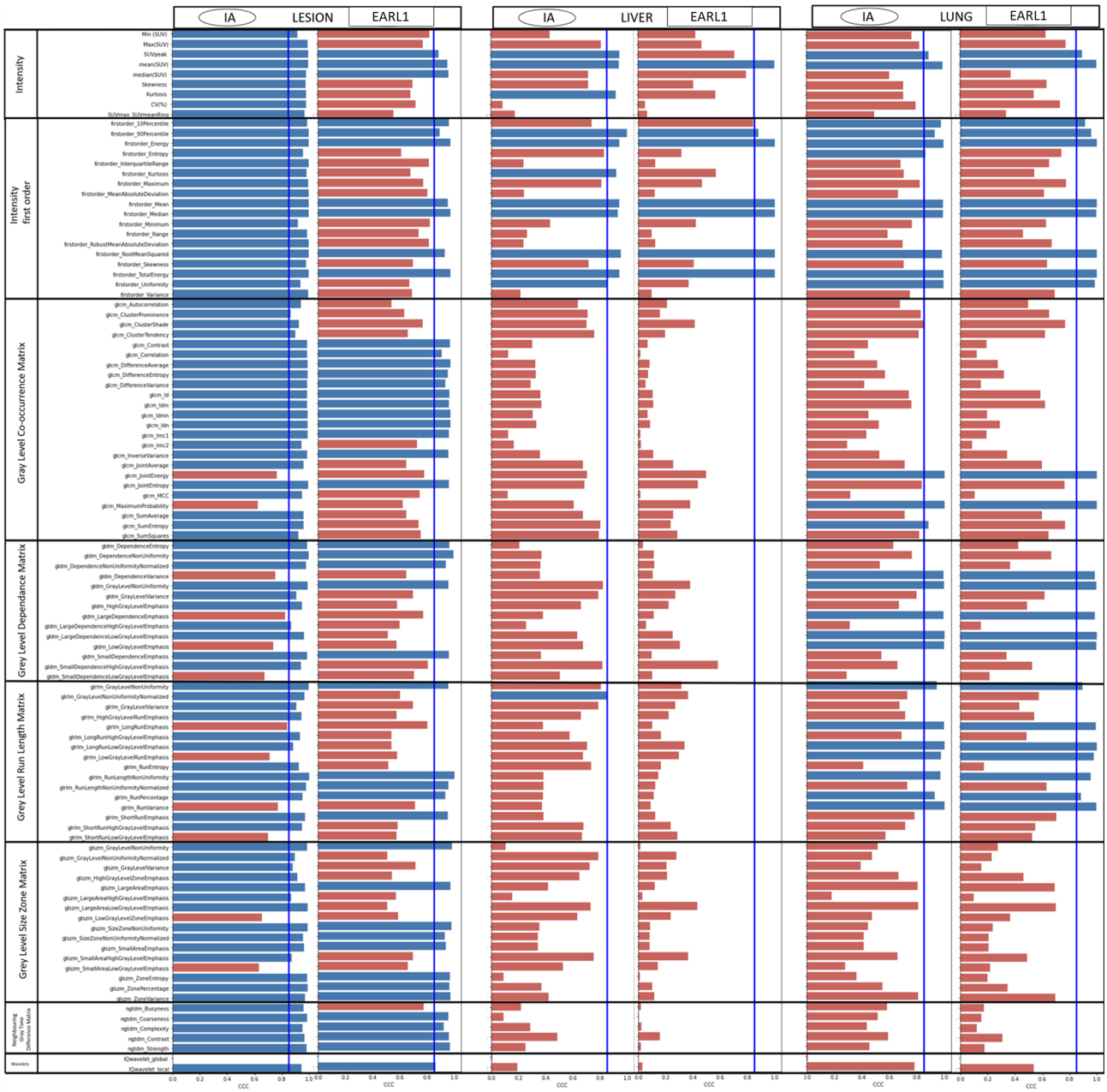


FIGURE 2 | CCC of all the features from AI and EARL *versus* original images. A threshold is display by a line with a CCC = 0.85. Blue bar indicates a CCC≥0.85, red bar CCC < 0.85.

DISCUSSION

We evaluated the impact of AI denoising on the stability of radiomics features computed in FDG PET images, the standard being the clinical images. We also concurrently evaluated the effect of EARL1 gaussian filtering. To the best of our knowledge, it is the first clinical study on the impact of artificial intelligence denoising on PET radiomics.

Texture features used in radiomics models describe the pattern distribution of voxels and quantify intra-tumor heterogeneity in all 3 dimensions (4). 86.5% showed a stable behavior for intensity and radiomics classes. The stability criterium was based on a CCC≥0.85 (20). In lesions, values were significantly different in 71.1% of the features after AI denoising. An AI denoising approach like CNN seems to change the absolute values of most of the features but keep the correlation between them.

Advanced applications aim at the correlation of image features, like radiomics, with clinical endpoints (4, 23). Radiomics models derived from CT correlated with a prognostic value, overall survival in lung cancer patient (5). In baseline PET of locally advanced rectal

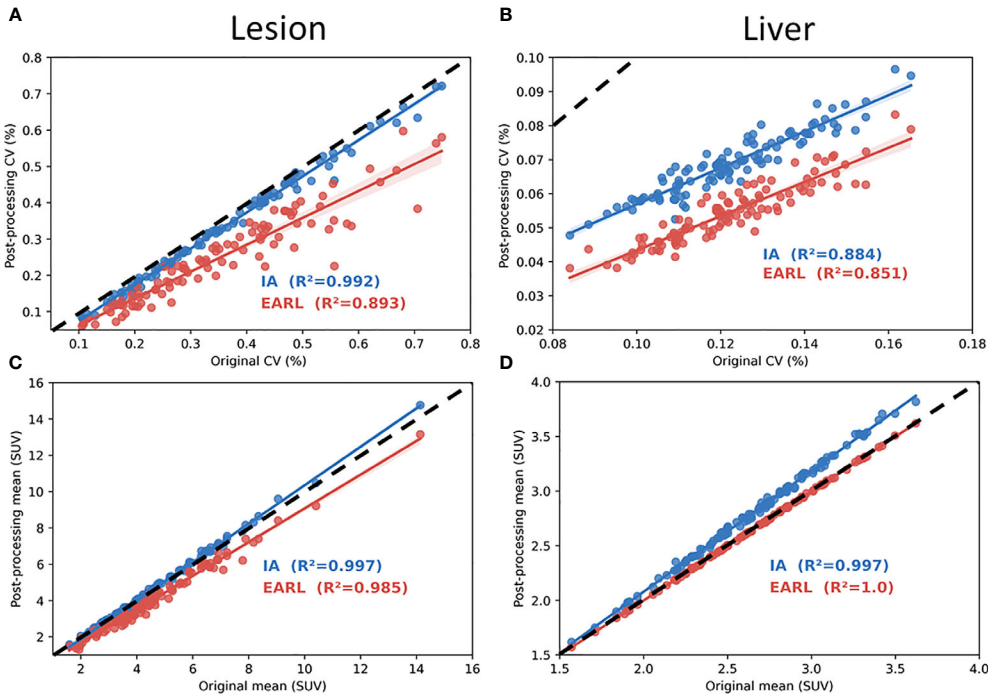


FIGURE 3 | Coefficient of correlation plot with R^2 value in lesions **(A, C)** and healthy liver **(B, D)**. **(A–D)** show respectively coefficient of variation (CV) and mean SUV calculated from AI and EARL1-like image in function of original images. Dotted line represents the identity line.

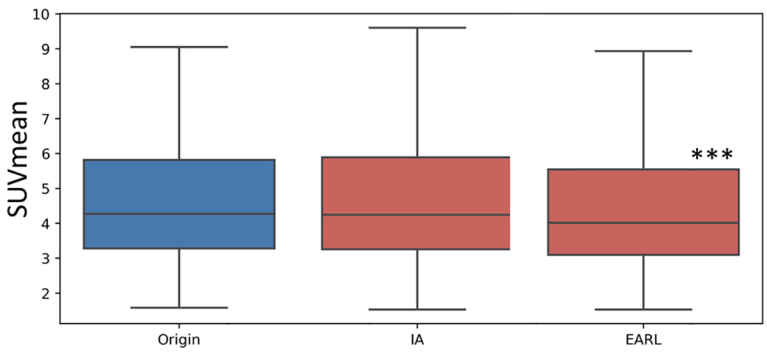


FIGURE 4 | Box plot of the mean SUV value in lesion in Original, AI and EARL1-like images. The distribution difference between the original images and AI is not significant (p=0.06) while it is significant (p<0.001 ***) between original and EARL1-like images.

cancer ^{18}F-FDG PET/CT texture features provide strong independent predictors of survival in patients (24). These models are very promising however there are several pitfalls to overcome (25) such as study design, data acquisition, segmentation, features calculation and modeling by the radiomics community. This study allows a better understanding of the behavior of predictive models when an AI denoising method is employed. A predictive model based on this type of information can be built from MRI, PET, CT or a combination of image modalities.

Deep learning AI techniques have been used to perform denoising on PET images for example by generating a full-dose PET images from low-dose images (26) or to directly filter reconstructed PET images (27). We used an AI denoising approach based on DCNN (11). This approach seems to be able to reduce the acquisition time activity product by a factor of 2 to 4. We used it directly on the studied PET image without activity or time reduction because we want to characterize the effect of AI denoising while not compensating for count losses.

Denoising will be more and more used but may also generate pitfalls to build a radiomics predictive model as the 3d texture information may be modified. Studying the stability of features with a test-retest approach has been performed in PET (28). The number

TABLE 2 | Result of Wilcoxon signed rank test of all the features between AI denoised and original images in lesions.

Classes	Features	P value	Median [min; max]	Classes	Features	P value	Median [min; max]
Intensity	SUVmin	<0.001 (***)	0.25 [-0.13 ; 1.09]	Gray Level Size Zone Matrix	glszm_GrayLevelNonUniformity	<0.001 (***)	-0.48 [-24.09 ; 1.34]
	SUVmax	<0.001 (***)	-0.53 [-4.52 ; 0.0]		glszm_GrayLevelNonUniformityNormalized	<0.001 (***)	-0.0 [-0.01 ; 0.01]
	SUVpeak (SUV)	<0.001 (***)	-0.15 [-1.3 ; 0.91]		glszm_GrayLevelVariance	<0.001 (***)	15.79 [-21.5 ; 60.88]
	SUVmean	0.13	0.03 [-0.34 ; 0.64]		glszm_HighGrayLevelZoneEmphasis	0.11	13.64 [-132.09 ; 206.28]
	SUVmedian	0.17	-0.04 [-0.75 ; 1.18]		glszm_LargeAreaEmphasis	0.11	0.02 [-95.71 ; 1321.6]
	Skewness	<0.01 (**)	-0.02 [-0.34 ; 0.21]		glszm_LargeAreaHighGrayLevelEmphasis	<0.001 (***)	162.72 [-1058.51 ; 112710.32]
	Kurtosis	<0.001 (***)	-0.13 [-1.46 ; 0.15]		glszm_LargeAreaLowGrayLevelEmphasis	<0.001 (***)	0.01 [-16.91 ; 46.07]
	CV(%)	<0.001 (***)	-0.02 [-0.09 ; -0.0]		glszm_LowGrayLevelZoneEmphasis	<0.001 (***)	0.0 [-0.03 ; 0.03]
	SUVmax/SUVmeanRing	<0.001 (***)	-0.55 [-3.74 ; -0.18]		glszm_SizeZoneNonUniformity	0.84	1.69 [-182.77 ; 68.57]
	firstorder_10Percentile	<0.001 (***)	0.1 [-0.15 ; 0.58]		glszm_SizeZoneNonUniformityNormalized	<0.05 (*)	0.0 [-0.1 ; 0.13]
	firstorder_90Percentile	<0.01 (**)	-0.11 [-0.67 ; 0.83]		glszm_SmallAreaEmphasis	<0.05 (*)	0.0 [-0.08 ; 0.07]
	firstorder_Energy	<0.01 (**)	-45.92 [-3961.59 ; 133754.46]		glszm_SmallAreaHighGrayLevelEmphasis	0.4	8.44 [-159.75 ; 258.22]
	firstorder_Entropy	<0.001 (***)	0.04 [-0.19 ; 0.23]		glszm_SmallAreaLowGrayLevelEmphasis	<0.001 (***)	0.0 [-0.03 ; 0.03]
	firstorder_InterquartileRange	<0.001 (***)	-0.08 [-0.43 ; 0.25]		glszm_ZoneEntropy	0.27	0.0 [-0.26 ; 0.18]
	firstorder_Kurtosis	<0.001 (***)	-0.13 [-1.47 ; 0.15]		glszm_ZonePercentage	0.55	0.0 [-0.07 ; 0.08]
	firstorder_Maximum	<0.001 (***)	-0.53 [-4.52 ; 0.0]		glszm_ZoneVariance	<0.05 (*)	0.04 [-95.59 ; 1276.34]
	firstorder_MeanAbsoluteDeviation	<0.001 (***)	-0.07 [-0.35 ; 0.18]	Gray Level Run Length Matrix	glrlm_GrayLevelNonUniformity	<0.001 (***)	-0.94 [-56.51 ; 4.11]
	firstorder_Mean	0.13	0.03 [-0.34 ; 0.64]		glrlm_GrayLevelNonUniformityNormalized	<0.001 (***)	-0.0 [-0.01 ; 0.0]
	firstorder_Median	<0.01 (**)	0.03 [-0.3 ; 0.71]		glrlm_GrayLevelVariance	<0.001 (***)	18.76 [-11.94 ; 65.78]
	firstorder_Minimum	<0.001 (***)	0.25 [-0.13 ; 1.09]		glrlm_HighGrayLevelRunEmphasis	<0.001 (***)	22.09 [-110.68 ; 235.67]
	firstorder_Range	<0.001 (***)	-0.81 [-5.13 ; -0.13]		glrlm_LongRunEmphasis	0.1	0.0 [-0.04 ; 1.12]
	firstorder_RobustMeanAbsoluteDeviati	<0.001 (***)	-0.04 [-0.15 ; 0.16]		glrlm_LongRunHighGrayLevelEmphasis	<0.001 (***)	34.26 [-118.94 ; 310.16]
	firstorder_RootMeanSquared	0.62	-0.0 [-0.46 ; 0.58]		glrlm_LongRunLowGrayLevelEmphasis	<0.001 (***)	0.0 [-0.04 ; 0.06]
	firstorder_Skewness	<0.01 (**)	-0.02 [-0.34 ; 0.22]		glrlm_LowGrayLevelRunEmphasis	<0.001 (***)	0.0 [-0.03 ; 0.04]
	firstorder_TotalEnergy	<0.01 (**)	367.4 [-31692.67 ; 1070036.9		glrlm_RunEntropy	<0.001 (***)	0.04 [-0.13 ; 0.28]
	firstorder_Uniformity	<0.001 (***)	-0.0 [-0.01 ; 0.02]		glrlm_RunLengthNonUniformity	0.85	0.9 [-772.5 ; 39.45]
	firstorder_Variance	<0.001 (***)	-0.16 [-5.37 ; 1.74]		glrlm_RunLengthNonUniformityNormalized	0.43	0.0 [-0.08 ; 0.01]
Gray Level Co-occurrence Matrix	glcm_Autocorrelation	<0.001 (***)	32.8 [-108.96 ; 302.57]		glrlm_RunPercentage	0.63	-0.0 [-0.1 ; 0.01]
	glcm_ClusterProminence	<0.001 (***)	3901.61 [-175021.65 ; 674754.		glrlm_RunVariance	<0.05 (*)	0.0 [-0.02 ; 0.64]
	glcm_ClusterShade	<0.001 (***)	1452.16 [-1515.44 ; 6965.97]		glrlm_ShortRunEmphasis	0.56	0.0 [-0.05 ; 0.01]
	glcm_ClusterTendency	<0.001 (***)	68.57 [-48.76 ; 212.39]		glrlm_ShortRunHighGrayLevelEmphasis	<0.001 (***)	20.21 [-109.38 ; 228.93]
	glcm_Contrast	<0.001 (***)	4.22 [-20.95 ; 63.31]		glrlm_ShortRunLowGrayLevelEmphasis	<0.001 (***)	0.0 [-0.03 ; 0.04]
	glcm_Correlation	<0.001 (***)	0.02 [-0.06 ; 0.09]	Neighbouring GrayTone Difference Matrix	ngtdm_Busyness	<0.01 (**)	0.01 [-0.19 ; 0.12]
	glcm_DifferenceAverage	<0.001 (***)	0.2 [-0.55 ; 1.95]		ngtdm_Coarseness	<0.01 (**)	-0.0 [-0.0 ; 0.01]
	glcm_DifferenceEntropy	<0.001 (***)	0.02 [-0.16 ; 0.25]		ngtdm_Complexity	<0.001 (***)	214.16 [-1087.88 ; 3109.34]
	glcm_DifferenceVariance	<0.001 (***)	1.0 [-7.23 ; 21.33]		ngtdm_Contrast	<0.001 (***)	0.03 [-0.65 ; 0.34]
	glcm_Id	<0.01 (**)	-0.0 [-0.03 ; 0.07]		ngtdm_Strength	<0.001 (***)	0.73 [-8.51 ; 6.65]
	glcm_Idm	<0.05 (*)	-0.0 [-0.02 ; 0.08]	Gray Level Dependence Matrix	gldm_DependenceEntropy	<0.001 (***)	0.04 [-0.19 ; 0.23]
	glcm_Idmn	<0.001 (***)	-0.0 [-0.01 ; 0.0]		gldm_DependenceNonUniformity	0.9	1.43 [-188.75 ; 44.87]
	glcm_Idn	<0.001 (***)	-0.0 [-0.02 ; 0.01]		gldm_DependenceNonUniformityNormalize	0.16	0.0 [-0.05 ; 0.09]
	glcm_Imc1	<0.001 (***)	-0.01 [-0.06 ; 0.02]		gldm_DependenceVariance	<0.05 (*)	0.04 [-1.53 ; 27.27]
	glcm_Imc2	<0.001 (***)	0.01 [-0.0 ; 0.03]		gldm_GrayLevelNonUniformity	<0.001 (***)	-0.94 [-58.39 ; 148.9]
	glcm_InverseVariance	<0.05 (*)	-0.0 [-0.04 ; 0.03]		gldm_GrayLevelVariance	<0.001 (***)	19.73 [-12.98 ; 66.4]
	glcm_JointAverage	<0.01 (**)	0.37 [-2.3 ; 4.37]		gldm_HighGrayLevelEmphasis	<0.001 (***)	22.35 [-109.79 ; 239.0]
	glcm_JointEnergy	<0.01 (**)	-0.0 [-0.0 ; 0.02]		gldm_LargeDependenceEmphasis	0.19	0.02 [-1.63 ; 64.99]
	glcm_JointEntropy	<0.001 (***)	0.03 [-0.56 ; 0.23]		gldm_LargeDependenceHighGrayLevelEmph	<0.001 (***)	382.84 [-802.27 ; 5417.16]
	glcm_MCC	<0.001 (***)	0.02 [-0.02 ; 0.07]		gldm_LargeDependenceLowGrayLevelEmph	<0.05 (*)	0.0 [-2.3 ; 3.47]
	glcm_MaximumProbability	0.34	-0.0 [-0.0 ; 0.09]		gldm_LowGrayLevelEmphasis	<0.001 (***)	0.0 [-0.03 ; 0.04]
	glcm_SumAverage	<0.01 (**)	0.74 [-4.6 ; 8.74]		gldm_SmallDependenceEmphasis	0.41	0.0 [-0.05 ; 0.08]
	glcm_SumEntropy	<0.001 (***)	0.06 [-0.2 ; 0.25]		gldm_SmallDependenceHighGrayLevelEmph	0.38	3.4 [-125.21 ; 227.65]
	glcm_SumSquares	<0.001 (***)	18.8 [-17.43 ; 66.07]		gldm_SmallDependenceLowGrayLevelEmph	<0.001 (***)	0.0 [-0.03 ; 0.02]

*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. Yellow filling means that the features are not significantly different $p>0.05$ and have a CCC>0.85.

of features selected based on their stability was 71% (CCC>0.8) in PET NSCLC patients. In this study the stability of FDG PET radiomics features in lesions was 86.5% (CCC≥0.85) between AI denoised and original images. These values are at least of the same order of magnitude highlighting the performance of AI for denoising in PET imaging. As a consequence, a predictive model built on standard PET images could be transposed on AI denoised images, especially concerning the features we have shown as stable in this study. However, the threshold values will have to be recomputed. On the other hand, in healthy tissues as liver and lung most of the features were unstable. Stable features were even less frequent in the liver (11.5%) than in the lung (25%). The effect of denoising on these tissues seems more drastic than on lesions. We hypothesize that the AI algorithm recognizes similar healthy features and changes their intensity value and distributions. As a consequence, the ratio of the lesion over liver uptake should be

transposed with care in clinical PET evaluation, as this ratio is altered for AI denoised versus original PET images.

The difference of behaviors in lesions and healthy tissue is one of the main advantages of AI based methods compared to an EARL1 gaussian post filter method. AI denoising maintains in the lesion the textural information and FDG uptake more stable while modifying healthy tissue. CV measures noise but also grey levels and is correlated to NECR/image quality in PET (29). As shown in **Figures 3A, B**, AI denoising had almost no effect on CV in lesions but reduced it in liver. On the contrary, EARL1 Gaussian postfilter reduced CV similarly in lesions and liver. Gaussian post filter will apply denoising accounting for neighbor all over the images whereas AI may be more selective in amplitude of denoising depending on noise vs non-noise components. The distribution of SUV mean in lesions has a different behavior between AI (paired t-test p=0.06) and EARL1 post filter (paired t-test p<0.001) compared to original images.

Interestingly EARL1 gaussian postfilter led to no modification of SUV mean values in liver (**Figure 3D**). It is mainly due to the increase of point spread function caused by the application of a gaussian postfilter. In a large homogenous area SUVmean was not modified while the noise (CV) was reduced. In smaller, more heterogeneous areas it will melt the grey levels of the different neighbors, lesions and healthy voxels (30). The modification of all the tissues in the image by the gaussian postfilter also appeared in **Figure 2**. Even in lesions only 44% of the features remained stable in EARL1 compatible versus original PET images.

In this analysis we tried to minimize the bias inherent of a radiomics workflow. We use pyradiomics which is mostly compatible with IBSI initiative. Each AI and EARL1 denoised images were extracted from the same images. The same VOI were used on all the series. One could however point out the use of the same contours for lesions in the 3 images as a possible study drawback. Re-segmentation of lesions on each image could have led to different contours and feature values. There is no gold standard for a segmentation method in PET radiomics. It remains also unclear to which extent this can affect radiomics values and predictive models (31). We chose a resampling of 64 bins instead of a fixed bin width (32) even if it showed a better reproducibility. As we directly compared images before and after denoising (minimum and maximum values of the image changed) resampling with a fixed bin width could lead to a different number of bins just due to noise reduction and not to texture based information. In a future work we would apply the same methodology with bin width resampling to strengthen our outcome. We didn' t split the data into training, validation and test cohorts in this study due to the relatively small number of patients and lesions (33). A test-retest radiomics study on patient in CT showed that 446/542 features had a higher CCC for patients with lung cancer than for those with rectal cancer (34). Our study was based on 113 patients, which is a small number. Pooling however different primary malignancies and lesions' nature and size might have helped to reduce overfitting. The main next challenge will be to validate our findings on different and heterogenous patient cohorts and other PET protocols and systems. It might be very risky to apply the same selection of features on other PET or even MRI or CT systems (25). Also, the mechanism of AI denoising recognizing successfully non noise *versus* noise components has to be further investigated on other camera types and PET protocols.

Numerical PET/CT's have a better spatial and temporal resolution leading to a more contrasted activity distribution in lesions than analog systems (35). As this study was carried out on a digital PET/CT we can expect that it will have been more sensitive to variations in texture compared to one on an analog system.

CONCLUSION

Applying an AI, CNN denoising on FDG PET images maintains most of the lesion's texture information in contrast to a EARL1-compatible Gaussian postfilter. The predictive texture features of a trained model could be transposed, however with an adapted threshold. Artificial intelligence in PET is a very promising approach as it adapts the denoising for noise *versus* non-noise components preserving information where it should.

DATA AVAILABILITY STATEMENT

All the data and the python code of the analysis are available on https://github.com/AurelienCD/RadiomicsIA_PET_Depository_Manuscript-ID-692973.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

CJ has designed the study, computed the radiomics and written the majority of the manuscript. KW included study patients, performed image analysis including segmentation, wrote manuscript parts and substantially revised it. AC-D has computed the statistics and written manusript sections. AL gave feedback on the AI algorithm. AB helped to design the study. SB gave advice and feedback on the study. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Henry Austins for the diligent correction of the English language.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.692973/full#supplementary-material

REFERENCES

1. OECD statistique. Available at: https://stats.oecd.org/.
2. SNITEM. Available at: https://www.snitem.fr/en/node/69277.
3. Forghani R, Savadjiev P, Chatterjee A, Muthukrishnan N, Reinhold C, Forghani B. Radiomics and Artificial Intelligence for Biomarker and Prediction Model Development in Oncology. *Comput Struct Biotechnol J* (2019) 17:995. doi: 10.1016/j.csbj.2019.07.001
4. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036. C., Gillies, R., Boellard, R. Dekker, A., and Aerts, HJ.
5. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5006. 6, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR. Dekker A, Quackenbush J, Gillies RJ, Lambin P.
6. Song J, Yin Y, Wang H, Chang Z, Liu Z, Cui L. A Review of Original Articles Published in the Emerging Field of Radiomics. *Eur J Radiol* (2020) 108991. doi: 10.1016/j.ejrad.2020.108991
7. Rogers W, Thulasi Seetha S, Refaee TA, Lieverse RI, Granzier RW, Ibrahim A, et al. Radiomics: From Qualitative to Quantitative Imaging. *Br J Radiol* (2020) 93(1108):20190948. doi: 10.1259/bjr.20190948
8. Visvikis D, Le Rest CC, Jaouen V, Hatt M. Artificial Intelligence, Machine (Deep) Learning and Radio (Geno) Mics: Definitions and Nuclear Medicine Imaging Applications. *Eur J Nucl Med Mol Imaging* (2019) 46(13):2630–7. doi: 10.1007/s00259-019-04373-w
9. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A Survey on Deep Learning in Medical Image Analysis. *Med Image Anal* (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005
10. Cui J, Gong K, Guo N, Wu C, Meng X, Kim K, et al. PET Image Denoising Using Unsupervised Deep Learning. *Eur J Nucl Med Mol Imaging* (2019) 46 (13):2780–9. doi: 10.1007/s00259-019-04468-4
11. Subtlepet. . Available at: https://subtlemedical.com/usa/subtlepet/.
12. Boellaard R, Hristova I, Ettinger S, Stroobants S, Chiti A, Bauer A, et al. Initial Experience With the EANM Accreditation Procedure of FDG PET/CT Devices. *Eur J Cancer* (2011) 47(Suppl. 4):S8 Abstract. doi: 10.1016/S0959-8049(11)72621-1
13. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. Fdg PET/CT: EANM Procedure Guidelines for Tumour Imaging: Version 2.0. *Eur J Nucl Med Mol Imaging* (2015) 42(2):328–54. doi: 10.1007/s00259-014-2961-x
14. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3d Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* (2012) 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001
15. Beichel RR, Van Tol M, Ulrich EJ, Bauer C, Chang T, Plichta KA, et al. Semiautomated Segmentation of Head and Neck Cancers in 18F-FDG PET Scans: A Just-Enough-Interaction Approach. *Med Phys* (2016) 43 (6Part1):2948–64. doi: 10.1118/1.4948679
16. Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.CAN-17-0339
17. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJ, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-throughput Image-Based Phenotyping. *Radiology* (2016) 295(2):328–38. doi: 10.1148/radiol.2020191145
18. Park JE, Park SY, Kim HJ, Kim HS. Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives. *Korean J Radiol* (2019) 20(7):1124. doi: 10.2967/jnumed.110.082404
19. Yaniv Z, Lowekamp BC, Johnson HJ, Beare R. SimpleITK Image-Analysis Notebooks: A Collaborative Environment for Education and Reproducible Research. *J Digital Imaging* (2018) 31(3):290–303. doi: 10.1007/s10278-017-0037-8. 31(3):290-303.
20. Lawrence I, Lin K. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* (1989) 45(1):255–68. doi: 10.2307/2532051
21. Peerlings J, Woodruff HC, Winfield JM, Ibrahim A, Van Beers BE, Heerschap A, et al. Stability of Radiomics Features in Apparent Diffusion Coefficient Maps From a Multi-Centre Test-Retest Trial. *Sci Rep* (2019) 9(1):1–10. doi: 10.1038/s41598-019-41344-5
22. Anaconda Software Distribution. *Anaconda Documentation*. Anaconda Inc. (2020). Available at: https://docs.anaconda.com/.
23. Seifert R, Weber M, Kocakavuk E, Rischpler C, Kersting D. AI and Machine Learning in Nuclear Medicine: Future Perspectives. In: *Seminars in Nuclear Medicine*. WB Saunders (2020). doi: 10.1053/j.semnuclmed.2020.08.003
24. Lovinfosse P, Polus M, Van Daele D, Martinive P, Daenen F, Hatt M, et al. Fdg PET/CT Radiomics for Predicting the Outcome of Locally Advanced Rectal Cancer. *Eur J Nucl Med Mol Imaging* (2018) 45(3):365–75. doi: 10.1007/s00259-017-3855-5
25. Hatt M, Le Rest CC, Antonorsi N, Tixier F, Tankyevych O, Jaouen V, et al. Radiomics in PET/CT: Current Status and Future Ai-Based Evolutions. In: *Seminars in Nuclear Medicine*. WB Saunders Seminars in Nuclear Medicine (2020). doi: 10.1053/j.semnuclmed.2020.09.002
26. Kaplan S, Zhu YM. Full-Dose PET Image Estimation From Low-Dose PET Image Using Deep Learning: A Pilot Study. *J Digital Imaging* (2019) 32 (5):773–8. doi: 10.1007/s10278-018-0150-3
27. Gong K, Guan J, Liu CC, Qi J. PET Image Denoising Using a Deep Neural Network Through Fine Tuning. *IEEE Trans Radiat Plasma Med Sci* (2018) 3 (2):153–61. doi: 10.1109/TRPMS.2018.2877644
28. Leijenaar RT, Carvalho S, Velazquez ER, Van Elmpt WJ, Parmar C, Hoekstra OS, et al. Stability of FDG-PET Radiomics Features: An Integrated Analysis of Test-Retest and Inter-Observer Variability. *Acta Oncol* (2013) 52(7):1391–7. doi: 10.3109/0284186X.2013.812798
29. Reynés-Llompart G, Sabaté-Llobera A, Llinares-Tello E, Martí-Climent JM, Gámez-Cenzano C. Image Quality Evaluation in a Modern PET System: Impact of New Reconstructions Methods and a Radiomics Approach. *Sci Rep* (2019) 9(1):1–9. doi: 10.1038/s41598-019-46937-8
30. Soret M, Bacharach SL, Buvat I. Partial-volume Effect in PET Tumor Imaging. *J Nucl Med* (2007) 48(6):932–45. doi: 10.2967/jnumed.106.035774
31. Cook GJ, Azad G, Owczarczyk K, Siddique M, Goh V. Challenges and Promises of PET Radiomics. *Int J Radiat Oncol Biol Phys* (2018) 102 (4):1083–9. doi: 10.1016/j.ijrobp.2017.12.268
32. Leijenaar RT, Nalbantov G, Carvalho S, Van Elmpt WJ, Troost EG, Boellaard R, et al. The Effect of SUV Discretization in Quantitative FDG-PET Radiomics: The Need for Standardized Methodology in Tumor Texture Analysis. *Sci Rep* (2015) 5(1):1–10. doi: 10.1038/srep11075
33. Steyerberg EW. Validation in Prediction Research: The Waste by Data Splitting. *J Clin Epidemiol* (2018) 103:131–3. doi: 10.1016/j.jclinepi.2018.07.010
34. van Timmeren JE, Leijenaar RT, van Elmpt W, Wang J, Zhang Z, Dekker A, et al. Test–Retest Data for Radiomics Feature Stability Analysis: Generalizable or Study-Specific? *Tomography* (2016) 2(4):361. doi: 10.18383/j.tom.2016.00208
35. Van Sluis J, De Jong J, Schaar J, Noordzij W, Van Snick P, Dierckx R, et al. Performance Characteristics of the Digital Biograph Vision PET/CT System. *J Nucl Med* (2019) 60(7):1031–6. doi: 10.2967/jnumed.118.215418

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 18 October 2021
doi: 10.3389/fonc.2021.726865

Accurate Tumor Delineation *vs.* Rough Volume of Interest Analysis for ^{18}F-FDG PET/CT Radiomics-Based Prognostic Modeling in Non-Small Cell Lung Cancer

Shima Sepehri[1], *Olena Tankyevych*[1,2], *Andrei Iantsen*[1], *Dimitris Visvikis*[1], *Mathieu Hatt*[1*†] *and Catherine Cheze Le Rest*[1,2†]

[1] LaTIM, INSERM, UMR 1101, Univ Brest, Brest, France, [2] University Hospital Poitiers, Nuclear Medicine Department, Poitiers, France

OPEN ACCESS

Edited by:
Florent L. Besson,
Université Paris-Saclay,
France

Reviewed by:
Subathra Adithan,
Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), India
Kun Zheng,
Peking Union Medical College Hospital (CAMS), China

***Correspondence:**
Mathieu Hatt
Mathieu.hatt@inserm.fr

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 17 June 2021
Accepted: 20 September 2021
Published: 18 October 2021

Citation:
Sepehri S, Tankyevych O, Iantsen A, Visvikis D, Hatt M and Cheze Le Rest C (2021) Accurate Tumor Delineation vs. Rough Volume of Interest Analysis for ^{18}F-FDG PET/CT Radiomics-Based Prognostic Modeling in Non-Small Cell Lung Cancer. Front. Oncol. 11:726865. doi: 10.3389/fonc.2021.726865

Background: The aim of this work was to investigate the ability of building prognostic models in non-small cell lung cancer (NSCLC) using radiomic features from positron emission tomography and computed tomography with 2-deoxy-2-[fluorine-18]fluoro-D-glucose (^{18}F-FDG PET/CT) images based on a "rough" volume of interest (VOI) containing the tumor instead of its accurate delineation, which is a significant time-consuming bottleneck of radiomics analyses.

Methods: A cohort of 138 patients with stage II–III NSCLC treated with radiochemotherapy recruited retrospectively (n = 87) and prospectively (n = 51) was used. Two approaches were compared: firstly, the radiomic features were extracted from the delineated primary tumor volumes in both PET (using the automated fuzzy locally adaptive Bayesian, FLAB) and CT (using a semi-automated approach with 3D Slicer™) components. Both delineations were carried out within previously manually defined "rough" VOIs containing the tumor and the surrounding tissues, which were exploited for the second approach: the same features were extracted from this alternative VOI. Both sets for features were then combined with the clinical variables and processed through the same machine learning (ML) pipelines using the retrospectively recruited patients as the training set and the prospectively recruited patients as the testing set. Logistic regression (LR), random forest (RF), and support vector machine (SVM), as well as their consensus through averaging the output probabilities, were considered for feature selection and modeling for overall survival (OS) prediction as a binary classification (either median OS or 6 months OS). The resulting models were compared in terms of balanced accuracy, sensitivity, and specificity.

Results: Overall, better performance was achieved using the features from delineated tumor volumes. This was observed consistently across ML algorithms and for the two clinical endpoints. However, the loss of performance was not significant, especially when a consensus of the three ML algorithms was considered (0.89 *vs*. 0.88 and 0.78 *vs*. 0.77).

Conclusion: Our findings suggest that it is feasible to achieve similar levels of prognostic accuracy in radiomics-based modeling by relying on a faster and easier VOI definition, skipping a time-consuming tumor delineation step, thus facilitating automation of the whole radiomics workflow. The associated cost is a loss of performance in the resulting models, although this loss can be greatly mitigated when a consensus of several models is relied upon.

Keywords: segmentation, radiomics, non-small cell lung cancer, machine learning, prognosis

1 INTRODUCTION

Non-small cell lung cancer (NSCLC) benefited from several improvements in diagnosis, staging, and treatment, but remains a deadly disease as the first cause of cancer death for men and the second for women (1). On the one hand, significant differences in the outcomes of patients have been observed depending on the clinical stage; hence, physicians rely on that factor to select a therapeutic strategy (i.e., concomitant or sequential combination of surgery, chemotherapy, and radiotherapy) (2). On the other hand, among patients with a similar stage, especially for stages II and III, highly variable outcomes (i.e., response to therapy and survival) have been reported.

Several studies showed the usefulness and the value of positron emission tomography/computed tomography (PET/CT) image modality using 2-deoxy-2-[^{18}F]fluoro-D-glucose (^{18}F-FDG) radiotracer for NSCLC staging, treatment planning, and monitoring (3). The clinical relevance of some of the new response metrics, such as the metabolically active tumor volume (MATV) and total lesion glycolysis (TLG), are under investigation. Commonly, the response to treatment is predominantly measured using the maximum standardized uptake value (SUV_{max}) obtained within a tumor. However, it has many shortcomings: firstly, SUV_{max} is not capable of characterizing all types of uptake changes and associated responses. It can only precisely measure those responses that occur when there is a global change in the tracer uptake, i.e., when the uptake changes in the tumor are spatially homogeneous. Since SUV_{max} only involves a single voxel, it cannot capture changes in the shape of the tumor or in its spatial uptake distribution properties.

In recent years, various handcrafted quantitative features, known today as radiomics, have been introduced and investigated for their potential to quantify the intensity, shape, and heterogeneity of tracer uptake within the tumor volume on PET/CT images (4, 5).

Because radiomic features are typically extracted from a previously delineated tumor volume, the impact of the segmentation step on the resulting intrinsic value of radiomics has been examined in several studies. The robustness of a subset of textural features used to quantify ^{18}F-FDG PET uptake, depending on the segmentation technique was first investigated in esophageal cancer treated with radiochemotherapy (6). A later study (7) investigated the test–retest variability of radiomic features in a dataset of 11 NSCLC patients with repeated scans and the inter-observer delineation variability in a set of 23 patients. Later, the impact of reconstruction and delineation was studied using 11 NSCLC full-body ^{18}F-FDG PET/CT scans in order to investigate the repeatability and the effects of the reconstruction methods and delineation (8). The repeatability of the radiomic features to explore sensitivity to image reconstruction, noise, and the delineation method was further considered by the same team (9). The impact of tumor segmentation on the robustness of the features (10), on the reproducibility and non-redundancy of the features (11), or on the resulting prognostic value (12) were also investigated recently. On the one hand, all these studies showed that the choice of segmentation techniques can lead to substantial variations for some radiomic features, but all investigated the impact within the context of using the most accurate tumor volume to extract features. On the other hand, several studies recently compared the use of features extracted from delineated tumors *versus* these extracted from specifically different (larger or smaller) volumes of interest (VOIs), i.e., not necessarily containing the entire tumor or limited to the tumor extent.

A first study in the context of cervical cancer and FDG PET imaging investigated the predictive value of features (volume and total lesion glycolysis) extracted from VOIs of varying sizes by considering various thresholds from 30% to 70% of the SUV_{max}, determining a variability of performance in the resulting models (13). A second work compared different segmentation volumes in differentiating uterine sarcoma from leiomyoma with preoperative imaging (14). The study compared three volumes: the tumor only, the tumor and the surrounding tissues, and the entire uterus. The best models were obtained by relying on features from the entire uterus [area under the receiver operating characteristic curve (AUC) = 0.876)] compared to the two other smaller VOIs (0.830 and 0.853 for tumor only and for tumor and the surrounding tissues, respectively). A third study investigated the impact of segmentation margin on machine learning (ML)-based high-dimensional quantitative CT texture analysis in the context of differentiating between low- and high-grade renal cancer (15). Two VOIs were compared: contour-focused *vs.* margin shrinkage of 2 mm. Features from the VOI with margin shrinkage were more reproducible than those from contour-focused VOI (93.2% *vs.* 86.2%); however, models combining contour-focused-derived features had better performance (AUC = 0.865–0.984 *vs.* 0.745–0.887).

One advantage of using larger VOIs containing the tumor could be to alleviate the need for accuracy in defining the VOI, hence facilitating and accelerating the whole radiomics analysis. Indeed, the accurate delineation of the tumor is often considered a significant time-consuming bottleneck step of the radiomics workflow.

The aim of this work was thus to investigate the ability of building prognostic models in NSCLC using radiomic features from ^{18}F-FDG PET/CT images based on a "rough" VOI containing the tumor volume instead of the accurately delineated tumor. We hypothesized that a combination of features extracted from this larger VOI may capture the relevant information in a different manner than those calculated in the delineated tumor volume and still enable prediction of the outcome, sparing the cost of the delineation step. In that context, it is expected that shape features might become less informative in the rough VOIs compared to those calculated on the delineated tumor and that additional and/or alternative intensity or textural features will be selected in the models instead.

2 MATERIALS AND METHODS

2.1 Patient Cohort

Since stage 1 patients have a very different (and more favorable) prognosis compared to those with stage II or III disease, mostly driven by treatment [(surgery *vs*. (chemo)radiotherapy], we focused here on patients with stage 2 and 3 tumors, where the potential impact of radiomics is likely to be the most important (16).

The inclusion criteria were confirmed NSCLC, stage 2 or 3; curative (chemo)radiotherapy treatment, and pretreatment FDG PET/CT imaging. Data from 138 NSCLC patients treated at the University Hospital of Poitiers, France, were collected (**Table 1**). The data of the first 87 patients were collected retrospectively, whereas the next 51 patients were recruited prospectively within the PRINCE project (INCa, PRTK-2015, registered trial NCT03199599). The study was conducted according to the guidelines of the Declaration of Helsinki. Ethical review and approval were waived for this study because the data were already collected for routine patient management before analysis, in which patients provided informed consent. No additional data were specifically collected for the present study. The exact same cohort of patients was recently analyzed in another study focusing on the comparison and fusion of ML algorithms, so the present results are directly comparable with that previous work (17).

2.2 PET/CT Imaging

All patients underwent a combined ^{18}F-FDG PET/CT acquisition as part of the diagnosis and staging before treatment. A Biograph mCT 40 ToF with axial field of view of 21.6 cm (Siemens, Erlangen, Germany) was used, relying on the routine clinical protocol. PET/CT acquisition began after 6 h of fasting and 60 ± 5 min after injection of 2.5 MBq/kg of ^{18}F-FDG (421 ± 98 MBq, range = 220–695 MBq). Non-contrast-enhanced, non-respiratory-gated (free breathing) CT images were acquired (120 kVp; Care Dose® current modulation system) with an in-plane resolution of 0.853×0.853 mm^2 and a 5-mm slice thickness. PET data were acquired using 3.5 min per bed position, and images were reconstructed using a CT-based attenuation correction and the standard routine clinical protocol, as we recently showed no improvement in the prognostic value of radiomic features when using different settings (either smaller voxels or smaller full width at half maximum of the Gaussian post-filtering) (18): OSEM-TrueX-TOF algorithm, with time-of-flight and spatial resolution modeling (three iterations and 21 subsets, 5-mm 3D Gaussian post-filtering; voxel size, $4 \times 4 \times 4$ mm^3).

2.3 Radiomics Analysis

2.3.1 Preprocessing

As the PET images were reconstructed on a matrix with isotropic voxels, no further image interpolation was performed. CT images were interpolated to isotropic $1 \times 1 \times 1$ mm^3 voxels using linear interpolation.

PET images were converted into SUV using patient weight. Low-dose CT images were processed in Hounsfield unit (HU).

2.3.2. VOI Definition and Segmentation

Only the primary tumors were considered. PET and CT images were segmented independently by a single expert. The first step consisted of manually defining a "rough" VOI containing the

TABLE 1 | Patient characteristics.

	Characteristics	No. of patients (*N* = 138)	Training/validation set (*N* = 87)	Test set (*N* = 51)
Gender	Male	106	62	44
	Female	32	25	7
Age (years)	Range	46–94	46–94	46–89
	Mean ± SD	71.43 ± 9.44	71.35 ± 9.37	71.55 ± 10.00
Treatment	Radiotherapy only	68	30	28
	Chemoradiotherapy	70	57	23
Histology	Adenocarcinoma	82	51	29
	Squamous cell carcinoma	56	36	22
Clinical stage	I	0	0	0
	II	43	26	17
	III	95	61	34
	IV	0	0	0

tumor and its surroundings in both modalities. This is the usual first step in facilitating the automated or semi-automated tumor delineation by excluding the surrounding physiological uptakes or normal structures that should not be included in the tumor-only analysis. The tumor metabolic volume was then obtained in PET by applying the FLAB algorithm (19, 20) (MIRAS v1.0, LaTIM INSERM UMR 1101, Brest, France) in the manually defined "rough" VOI. The anatomical volume was obtained from the low-dose CT rough VOI semi-automatically by relying on the *Growcut effect* function of 3D Slicer™ (21). All delineations were checked and validated by an expert physician (C. Cheze Le Rest). **Figure 1** illustrates this process.

For the rest of the radiomics workflow, two different volumes for both PET and CT images were thus considered: the delineated tumor volume and the rough VOI.

2.3.3 Radiomic Feature Extraction

Seventy-three radiomic features (14 shape, 10 intensity, and 49 textural) (see **Supplemental Table 1**) compliant with the most up-to-date imaging biomarker standardization initiative (IBSI) benchmark (22) were extracted using homemade software (MIRAS v1.0, LaTIM INSERM UMR 1101, Brest, France). Three different grey-level discretization methods [fixed bin number (FBN) with 64 bins, fixed bin width (FBW) with 0.5 SUV or 10 HU, and histogram equalization with 64 bins] were considered for second- and higher-order textural features. Note that the FBN and FBW discretization schemes are IBSI-compliant, but the histogram equalization, although mentioned by the IBSI, is not yet a standard. Texture matrices were implemented in 3D following the merging strategy (i.e., considering all 13 directions simultaneously). More details on the entire radiomics workflow are provided in **Supplemental Table 2**. A total of 147 features ($10 + 14 + 49 \times 3$) were thus extracted from each tumor volume in both PET and CT, leading to 294 image-derived variables for each patient. These 294 features were extracted from both the tumor delineated volumes and the rough VOI.

2.3.4 Modeling

All available clinical variables (age, gender, stage, treatment, and histology) and the PET and CT radiomic features were grouped into a single set to be processed by each of the ML pipelines. The two different sets corresponding to the two approaches (delineated tumor *vs.* rough VOI) were processed independently using the exact same data split and ML pipeline for a fair comparison.

Data were split into a training/validation set (n = 87 retrospectively recruited patients, 63%) and a test set (n = 51 prospectively recruited patients, 37%) (**Table 1**).

The classification task was set as a binary identification of patients with overall survival (OS) below 6 months (unbalanced, n = 15 in the training set and n = 9 in the test set) or below the median OS (balanced). In the case of the 6-month prediction, the synthetic minority oversampling technique (SMOTE) was implemented to facilitate the training of models.

The ML pipelines consisted of three algorithms with embedded feature selection and a consensus: support vector machine (SVM) with recursive feature elimination (RFE), random forest (RF) with embedded wrapper (EW), and logistic regression (LR) with features selected using least absolute shrinkage and selection

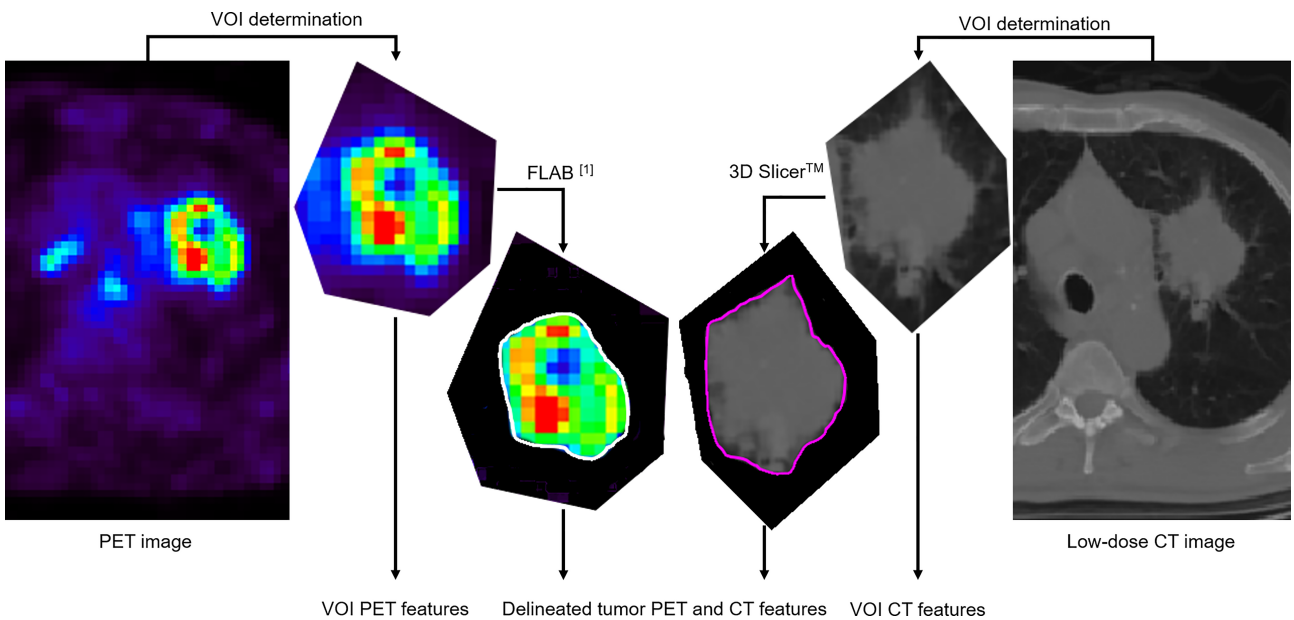


FIGURE 1 | Both PET and low-dose CT images of the primary tumor are processed in the same manner: a volume of interest (VOI) containing the tumor is first manually determined. Radiomic features are extracted from this VOI (denoted "VOI features"). Then, segmentation of the tumor volume is carried out within the VOI with a (semi)automated algorithm. Radiomic features are then extracted from the delineated volume (this is the usual workflow).

operator (LASSO). Hyperparameters of the algorithms (e.g., the number of trees in RF) were optimized through five-fold cross-validation in the training/validation set. To generate a consensus model, the output probabilities of each of the three algorithms were averaged and binarized (> or ≤0.5), as this approach provided better results than did majority voting in our previous work (17).

The performance of the models in the training/validation set was assessed using accuracy, balanced accuracy (BAcc) in the case of the 6-month OS prediction, and the combination of sensitivity (Se) and specificity (Sp), favoring models with higher Sp and with a smaller number of features (lower complexity and higher potential for generalizability) for similar levels of BAcc. For models with similar overall accuracy values, a higher specificity is more clinically relevant, as patients who would be falsely identified as having poor prognosis might be offered palliative (or intensified) treatment when the standard treatment would actually benefit them. Importantly, none of the data from the test set were used in the training and optimization step of any of the models under comparison (either each of the models or the ensemble through averaging). All models were first finalized and optimized in the training set before final evaluation without any further modifications in the test set.

The best models obtained through each ML algorithm in training/validation set were then applied to the test set for final evaluation and to allow for relevant comparisons (i.e., models trained using input features from the accurately delineated tumor *vs*. those from the "rough" VOI). In order to provide some reference comparison, we also determined the accuracy reached by using only the clinical features as input to the ML pipelines or by relying only on clinical staging (stage 2 *vs*. 3), as previously reported (17).

Finally, although the present work focuses on the question of the input VOI for the performance of the models rather than the actual development of a prognostic model, we nonetheless auto-evaluated our study using the radiomics quality score (RQS) (23).

3 RESULTS

Our study scored moderately on the RQS (see **Supplemental Table 3**) at 16 (19 when the data will be made available) out of 36, which is nonetheless higher than that of the average of studies reported recently (23–25).

The average follow-up was 41 months, with a minimum of 1.1 months and a maximum of 95 months. Median OS was 14.4 months, ranging between 1.1 and 50 months.

All results from the different models and the two outcome prediction tasks are presented for the training and test sets in **Table 2**.

Models trained using only the clinical variables as input did not significantly improve the performance over clinical stage alone (BAcc <0.60 for all ML pipelines and both endpoints in the training set and <0.55 in the test set).

Overall, the level of accuracy achieved by the models relying on radiomic features was superior to that of clinical stage alone (BAcc values of 0.58 and 0.53 using stage 2 *vs*. stage 3 classification, respectively, as previously reported) (17), and the models were better at predicting very poor prognosis (6-month OS endpoint) than median OS. Some of the radiomics models included one or two clinical variables (staging and/or treatment), but mostly relied on the histogram, shape (except for the models trained using rough VOI features), and textural features. The drop of performance between the training/validation and test sets also suggests some overfitting.

Regarding the question addressed in this work, it was observed that the radiomic features extracted from the delineated primary tumor volume were slightly more

TABLE 2 | Performance comparison of the ML techniques using either features from the delineated tumor (D) or from the rough VOI (V), in addition to the available clinical factors.

ML	Task	VOI[a]	Training set			No. of features	Test set		
			Se	Sp	BAcc		Se	Sp	BAcc
LR	Median OS	D	0.67	0.77	0.72	37	0.54	0.75	0.63
		V	0.58	0.68	0.63	24	0.59	0.57	0.58
	6-month OS	D	0.81	0.87	0.84	45	0.8	0.76	0.78
		V	0.74	0.78	0.76	32	0.61	0.65	0.63
RF	Median OS	D	0.87	0.91	0.89	25	0.60	0.75	0.67
		V	0.75	0.86	0.87	23	0,53	0.59	0.56
	6-month OS	D	1	1	1	47	0.74	0.86	0.80
		V	0.83	0.89	0.86	58	0.73	0.75	0.74
SVM	Median OS	D	1	1	1	27	0.53	0.73	0.64
		V	0.82	0.82	0.82	20	0.56	0.60	0.58
	6-month OS	D	0.88	0.96	0.92	38	0.76	0.74	0.75
		V	0.84	0.90	0.87	43	0.75	0.77	0.76
Fusion (average of output probabilities)	Median OS	D	1	1	1	-	0.76	0.80	0.78
		V	0.93	0.89	0.90	-	0.76	0.78	0.77
	6-month OS	D	1	1	1	-	0.91	0.87	0.89
		V	0.88	0.94	0.91	-	0.98	0.78	0.88

ML, machine learning; VOI, volume of interest; Se, sensitivity; Sp, specificity; BAcc, balanced accuracy; LR, logistic regression; RF, random forest; SVM, support vector machine.
[a]D stands for the accurately delineated tumor and V for the "rough" VOI.

informative than those extracted from a rough VOI. Indeed, models built with the three ML pipelines combining rough VOI features obtained a slightly lower performance (BAcc values of 0.57 ± 0.01 and 0.71 ± 0.07 for median OS and 6-month OS, respectively) than those exploiting delineated tumor features (BAcc values of 0.65 ± 0.02 and 0.78 ± 0.03). In both cases, the differences were not significant at the $p < 0.01$ or 0.05 levels ($p = 0.059$ for median OS and $p = 0.286$ for 6-month OS).

These models, however, relied on a similar number of features, i.e., models using rough VOI features did not need a larger number of features. Notably, shape features were not included in the models based on the rough VOI, contrary to those relying on the delineated tumor.

These observations were consistent for both endpoints (median OS and 6-month OS).

However, when looking at the consensus models (fusion of the output probabilities of each of the three pipelines, as the average of the outputs), the advantage of relying on the delineated tumor rather than on the rough VOI was greatly reduced, the two showing almost the same performances, with improved predictive ability compared to each independent ML algorithm, as previously reported (17): 0.89 *vs.* 0.88 for the 6-month OS endpoint and 0.78 *vs.* 0.77 for median OS, respectively.

4 DISCUSSION

The main finding of our work is that, although radiomic features extracted from delineated tumor seemed more informative than those extracted from a simple rough VOI containing the tumor, almost as good results can be achieved without the need for the tedious and time-consuming (semi)automated delineation in the radiomics workflow, especially in the context of relying on a consensus of several ML techniques (in that case, the performance was almost equal). This could imply consequences regarding the way radiomics analyses are carried out since avoiding the need for actual tumor delineation before feature extraction could simplify and facilitate the whole process, at a very small cost in the resulting performance of the built models.

As expected, no shape features were used by the models to predict outcomes when features were extracted from the rough VOI, contrary to when then are extracted from the delineated tumor. Although there is an obvious correlation between the size and shape of the VOI and that of the tumor (larger, more complex tumors require larger and more complex VOIs to encompass them), there are most likely fewer differences between the various VOIs shapes to allow for patient differentiation. These features were replaced by alternative intensity and/or textural metrics in the VOI models. Although some of the models retained clinical variables (only clinical staging and treatment being selected), relying only on clinical factors provided the models with only limited accuracy (<0.60 in training and <0.55 in testing), and only models incorporating radiomic features had good performance in the test set.

Our work has several limitations. Firstly, the cohort used was collected from a single center. It allowed us to focus on the question at hand without having to deal with harmonization issues (26, 27) since all patients had their PET/CT acquisition in the exact same PET/CT system, with no variability in the acquisition protocol or reconstruction settings. However, this means that our findings will need to be validated in external datasets, for which we will implement harmonization techniques for handling the multicenter nature of the data (28). Although our cohort included both retrospectively (for training) and prospectively (for testing) recruited patients, the size of the test set was small as we could not include all available patients because a minimum follow-up duration was not reached and the prospective recruitment is still ongoing. This limited the statistical power for comparing the different results obtained with or without tumor delineation. However, the observed trends were systematic across all ML techniques and their consensus, strengthening our confidence in the potential generalizability of our results. The VOI determination and the tumor delineations were carried out by a single expert using a single method, which prevented us to compare the scale of inter-observer (or inter-segmentation method) variability with the differences between the delineated tumor and VOI features. The sensitivity of the results with respect to (moderate and reasonable) changes in the size or shape of the rough VOI was also not explored in the present work. It is expected to be obviously lower than the differences observed between the results obtained when exploiting features either from the rough VOI or from the accurately delineated tumor, which are already small. Finally, in order to fully automate the process for facilitating the radiomics workflow, the "rough" VOIs, which were manually created in the present work, should be reproduced by training a deep convolutional neural network (CNN) such as the U-Net, in a similar fashion, as we have recently demonstrated the feasibility regarding accurate tumor delineation (29). This way, the "rough" VOI could be obtained in a fully automated manner from the input PET/CT images without the need for a user intervention.

Several expansions of this work will be considered, such as a thorough comparison with deep learning-based feature extraction ("deep features") and a validation of our findings in our extended prospective cohort: about 150 patients prospectively recruited in the PRINCE project should be available for this analysis once the follow-up duration of at least 1 year will be reached for all patients. Further validation of these findings will also be carried out in external datasets.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and

institutional requirements. Indeed the patients/participants provided their written informed consent that their data could be collected for future research purposes as part of their routine management. No additional data was specifically collected for the present study.

AUTHOR CONTRIBUTIONS

CC, MH, and DV conceptualized the study. SS, OT, AI, and MH helped with the methodology. SS helped with software. SS and OT contributed to the formal analysis. SS, OT, and CC curated the data. SS and MH prepared the original draft. SS, OT, AI, CC, DV, and MH reviewed and edited the paper. CC and MH helped with funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the French Ministry of Research (PhD grant), the INCa and DGOS through a grant in the PRINCE project (PRTK-2015, #R16063NN), the PREDICT Innovative Training Network funded by the Marie Sklodowska-Curie Actions, part of the EU's Horizon 2020 Programme (grant agreement no. 766276), and the POPEYE project under the frame of ANR (Agence National de la Recherche; ANR-19-PERM-0007 POPEYE) and of ERA PerMed (POPEYE T11EPA4-00055).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.726865/full#supplementary-material

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68 (6):394–424. doi: 10.3322/caac.21492
2. Coche E, Lonneux M, Geets X. Lung Cancer: Morphological and Functional Approach to Screening, Staging and Treatment Planning. *Future Oncol Lond Engl* (2010) 6(3):367–80. doi: 10.2217/fon.10.7
3. Sauter AW, Schwenzer N, Divine MR, Pichler BJ, Pfannenberg C. Image-Derived Biomarkers and Multimodal Imaging Strategies for Lung Cancer Management. *Eur J Nucl Med Mol Imaging* (2015) 42(4):634–43. doi: 10.1007/s00259-014-2974-5
4. Hatt M, Le Rest CC, Tixier F, Badic B, Schick U, Visvikis D. Radiomics: Data Are Also Images. *J Nucl Med Off Publ Soc Nucl Med sept* (2019) 60(Suppl 2):38S–44S. doi: 10.2967/jnumed.118.220582
5. Hatt M, Cheze Le Rest C, Antonorsi N, Tixier F, Tankyevych O, Jaouen V, et al. Radiomics in PET/CT: Current Status and Future AI-Based Evolutions. *Semin Nucl Med* (2021) 51(2):126–33. doi: 10.1053/j.semnuclmed.2020.09.002
6. Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of Intratumour ^{18}F-FDG PET Uptake Heterogeneity Quantification for Therapy Response Prediction in Oesophageal Carcinoma. *Eur J Nucl Med Mol Imaging* (2013) 40(11):1662–71. doi: 10.1007/s00259-013-2486-8
7. Leijenaar RTH, Carvalho S, Velazquez ER, van Elmpt WJC, Parmar C, Hoekstra OS, et al. Stability of FDG-PET Radiomics Features: An Integrated Analysis of Test-Retest and Inter-Observer Variability. *Acta Oncol Stockh Swed* (2013) 52(7):1391–7. doi: 10.3109/0284186X.2013.812798
8. van Velden FHP, Kramer GM, Frings V, Nissen IA, Mulder ER, de Langen AJ, et al. Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [(18)F]FDG-PET/CT Studies: Impact of Reconstruction and Delineation. *Mol Imaging Biol* (2016) 18(5):788–95. doi: 10.1007/s11307-016-0940-2
9. Pfaehler E, Beukinga RJ, de Jong JR, Slart RHJA, Slump CH, Dierckx RAJO, et al. Repeatability of 18 F-FDG PET Radiomic Features: A Phantom Study to Explore Sensitivity to Image Reconstruction Settings, Noise, and Delineation Method. *Med Phys* (2019) 46(2):665–78. doi: 10.1002/mp.13322
10. Tixier F, Um H, Young RJ, Veeraraghavan H. Reliability of Tumor Segmentation in Glioblastoma: Impact on the Robustness of MRI-Radiomic Features. *Med Phys* (2019) 46(8):3582–91. doi: 10.1002/mp.13624
11. Qiu Q, Duan J, Duan Z, Meng X, Ma C, Zhu J, et al. Reproducibility and Non-Redundancy of Radiomic Features Extracted From Arterial Phase CT Scans in Hepatocellular Carcinoma Patients: Impact of Tumor Segmentation Variability. *Quant Imaging Med Surg* (2019) 9(3):453–64. doi: 10.21037/qims.2019.03.02
12. Hatt M, Laurent B, Fayad H, Jaouen V, Visvikis D, Le Rest CC. Tumour Functional Sphericity From PET Images: Prognostic Value in NSCLC and Impact of Delineation Method. *Eur J Nucl Med Mol Imaging* (2018) 45 (4):630–41. doi: 10.1007/s00259-017-3865-3
13. Leseur J, Roman-Jimenez G, Devillers A, Ospina-Arango JD, Williaume D, Castelli J, et al. Pre- and Per-Treatment 18F-FDG PET/CT Parameters to Predict Recurrence and Survival in Cervical Cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2016) 120(3):512–8. doi: 10.1016/j.radonc.2016.08.008
14. Xie H, Zhang X, Ma S, Liu Y, Wang X. Preoperative Differentiation of Uterine Sarcoma From Leiomyoma: Comparison of Three Models Based on Different Segmentation Volumes Using Radiomics. *Mol Imaging Biol* (2019) 21(6):1157–64. doi: 10.1007/s11307-019-01332-7
15. Kocak B, Ates E, Durmaz ES, Ulusan MB, Kilickesmez O. Influence of Segmentation Margin on Machine Learning-Based High-Dimensional Quantitative CT Texture Analysis: A Reproducibility Study on Renal Clear Cell Carcinomas. *Eur Radiol* (2019) 29(9):4765–75. doi: 10.1007/s00330-019-6003-8
16. Desseroit M-C, Visvikis D, Tixier F, Majdoub M, Perdrisot R, Guillevin R, et al. Development of a Nomogram Combining Clinical Staging With (18)F-FDG PET/CT Image Features in Non-Small-Cell Lung Cancer Stage I-III. *Eur J Nucl Med Mol Imaging* (2016) 43(8):1477–85. doi: 10.1007/s00259-016-3325-5
17. Sepehri S, Tankyevych O, Upadhaya T, Visvikis D, Hatt M, Cheze Le Rest C. Comparison and Fusion of Machine Learning Algorithms for Prospective Validation of PET/CT Radiomic Features Prognostic Value in Stage II-III Non-Small Cell Lung Cancer. *Diagn Basel Switz* (2021) 11(4):675. doi: 10.3390/diagnostics11040675
18. Tankyevych O, Tixier F, Antonorsi N, Filali Razzouki A, Mondon R, Pinto-Leite T, et al. Can Alternative PET Reconstruction Schemes Improve the Prognostic Value of Radiomic Features in Non-Small Cell Lung Cancer? *Methods* (2020) 11(4):675. doi: 10.1016/j.ymeth.2020.11.002
19. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A Fuzzy Locally Adaptive Bayesian Segmentation Approach for Volume Determination in PET. *IEEE Trans Med Imaging* (2009) 28(6):881–93. doi: 10.1109/TMI.2008.2012036
20. Hatt M, Cheze le Rest C, Descourt P, Dekker A, De Ruysscher D, Oellers M, et al. Accurate Automatic Delineation of Heterogeneous Functional Volumes in Positron Emission Tomography for Oncology Applications. *Int J Radiat Oncol Biol Phys* (2010) 77(1):301–8. doi: 10.1016/j.ijrobp.2009.08.018
21. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust Radiomics Feature Quantification Using Semiautomatic Volumetric Segmentation. *PloS One* (2014) 9(7):e102107. doi: 10.1371/journal.pone.0102107
22. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145

23. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, et al. Quality of Science and Reporting of Radiomics in Oncologic Studies: Room for Improvement According to Radiomics Quality Score and TRIPOD Statement. *Eur Radiol* (2020) 30(1):523–36. doi: 10.1007/s00330-019-06360-z
24. Chetan MR, Gleeson FV. Radiomics in Predicting Treatment Response in Non-Small-Cell Lung Cancer: Current Status, Challenges and Future Perspectives. *Eur Radiol* (2021) 31(2):1049–58. doi: 10.1007/s00330-020-07141-9
25. Abdurixiti M, Nijiati M, Shen R, Ya Q, Abuduxiku N, Nijiati M. Current Progress and Quality of Radiomic Studies for Predicting EGFR Mutation in Patients With Non-Small Cell Lung Cancer Using PET/CT Images: A Systematic Review. *Br J Radiol* (2021) 94(1122):20201272. doi: 10.1259/bjr.20201272
26. Hatt M, Lucia F, Schick U, Visvikis D. Multicentric Validation of Radiomics Findings: Challenges and Opportunities. *EBioMedicine* (2019) 47:20–1. doi: 10.1016/j.ebiom.2019.08.054
27. Da-Ano R, Visvikis D, Hatt M. Harmonization Strategies for Multicenter Radiomics Investigations. *Phys Med Biol* (2020) 65(24):24TR02. doi: 10.1088/1361-6560/aba798
28. Da-ano R, Masson I, Lucia F, Doré M, Robin P, Alfieri J, et al. Performance Comparison of Modified ComBat for Harmonization of Radiomic Features for Multicenter Studies. *Sci Rep* (2020) 10(1):10248. doi: 10.1038/s41598-020-66110-w
29. Iantsen A, Ferreira M, Lucia F, Jaouen V, Reinhold C, Bonaffini P, et al. Convolutional Neural Networks for PET Functional Volume Fully Automatic Segmentation: Development and Validation in a Multi-Center Setting. *Eur J Nucl Med Mol Imaging* (2021) 48(11):3444–56. doi: 10.1007/s00259-021-05244-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership