Functional and quantitative imaging of the lung

Edited by
Mark O. Wielpütz, Jim M. Wild and Edwin J. R. van Beek

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

Functional and quantitative imaging of the lung

Topic editors

Mark O. Wielpütz — Heidelberg University, Germany
Jim M. Wild — The University of Sheffield, United Kingdom
Edwin J. R. van Beek — University of Edinburgh, United Kingdom

Table of
contents

Frontiers | Frontiers in Medicine

Editorial: Functional and quantitative imaging of the lung

Mark O. Wielpütz[1,2,3]*, Jim M. Wild[4,5] and Edwin J. R. van Beek[6]

[1]Subdivision of Pulmonary Imaging, Department of Diagnostic and Interventional Radiology, University Hospital Heidelberg, Heidelberg, Germany, [2]Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), Heidelberg, Germany, [3]Department of Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany, [4]Pulmonary, Lung and Respiratory Imaging Sheffield (POLARIS), Division of Clinical Medicine, Faculty of Health, School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom, [5]Insigneo Institute, University of Sheffield, Sheffield, United Kingdom, [6]Edinburgh Imaging, Queens Medical Research Institute, University of Edinburgh, Edinburgh, United Kingdom

Editorial on the Research Topic
Functional and quantitative imaging of the lung

Pulmonary functional imaging provides important functional and quantitative metrics for a wide range of pathological conditions. Functional lung imaging research is based on two main paradigms: (1) in the lungs, structural alterations cause inevitable loss in lung function, ultimately leading to perfusion and ventilation abnormalities and subsequent reduction of blood oxygenation. Several different structural compartments, such as the airways, conducting pulmonary arterial and bronchial arterial as well as venous vessels, the capillaries, and the tissue barrier, allowing diffusion of oxygen from the alveoli to the red blood cells, can be affected in various combinations. (2) Traditional pulmonary function testing such as spirometry and full body plethysmography is insensitive to early lung disease since healthy areas may compensate for inhomogeneously distributed function loss. Also, spirometry is a global total lung and airways assessment, which cannot differentiate a regional pattern of tissue destruction, and has limited ability to differentiate the type of compartments affected.

The present Research Topic aimed to present the current plethora of techniques in pulmonary functional research based on radiologic imaging, which are scaled from small animal models to clinical imaging in patients with lung disease. The chain of quantification ranges from inherently quantitative techniques such as T1 mapping with magnetic resonance imaging (MRI) to delicate post-processing encompassing segmentation of lungs, vessels, airways and pulmonary abnormalities with heuristic and, more recently, artificial intelligence-based methods, and very often combinations of these. At the same time, the original works on display in this Research Topic demonstrate the strong translational focus of our field of research.

The works by Almeida et al., Ji et al., and Konietzke et al. sought to capture structural and functional abnormalities in the chronic obstructive pulmonary disease (COPD) lung by quantitative post processing. Ji et al. separated COPD patients into three groups based on the visual impression of interstitial lung abnormalities (ILA) accompanying the typical imaging features of COPD (1). The authors could show that in the group of COPD patients with visual presence of ILA, airway wall thickness was increased. Konietzke et al. performed a thorough analysis of airway wall and parenchymal metrics using the GOLD stage, demonstrating

that emphysema severity incrementally increases with every GOLD stage whereas airway wall thickening stagnated from GOLD II (2). Almeida et al. complemented these metrics with a novel artificial intelligence-based self-supervising network to detect anomalies in lung structure in COPD, adding to the traditional measurement of lung density to detect emphysema. The authors demonstrated good correlations of this AI-based abnormality score with traditional CT metrics and lung function decline. A novel but increasingly studied CT technique based on phase contrast (rather than absorption alone) was applied by Dullin et al. to a mouse model of allergic airway disease employing a synchrotron-derived x-ray source. The study by Yoshida et al. used a very large x-ray dataset in conjunction with a deep learning network to successfully predict spirometry values. Though the actual mechanism through which AI can derive such information remains obscure, it is a strong demonstration of how well imaging data is connected to lung function.

Magnetic resonance imaging (MRI) holds great potential for repeat and functional assessments of the lung, making it an ideal research tool. But it has also been introduced into clinical routine for a number of specific indications such as cystic fibrosis, primary ciliary dyskinesia and Pancoast tumor (3, 4). The works by Doellinger et al., Mummy et al., Ringwald et al., and Triphan et al. dive deep into this dynamic research field of functional lung MRI, using three different technological approaches. Mummy et al. explored the impact of anthropomorphic details such as age, sex and body mass index on MRI-derived measures of pulmonary gas exchange with hyperpolarized ^{129}Xe-MRI. Since this technique holds great potential to regionally assess impairments in the air-to-blood diffusion barrier, such work is of great importance for implementing this advanced technique in clinical routine at specialized centers (5). Doellinger et al. used a completely contrast agent-free approach to assess pulmonary ventilation and perfusion solely from time-resolved acquisitions in combination with registration and a matric-pencil decomposition, which separates cyclic changes of MRI signal intensity into contributions of pulsatile blood inflow and respiration (6). The authors could demonstrate a good correlation with the more widely established contrast-enhanced 4D perfusion technique and score, opening a perspective for contrast-free assessments of cystic fibrosis lung disease. Ringwald et al. used a large MRI dataset incl. 4D perfusion data to train a convolutional neural network to segment the lungs in children with cystic fibrosis. This marks progress in the effort to make MRI evaluation less user-dependent. Triphan et al. used another approach by directly quantifying the T1-relaxation times in an echo-time dependent manner in COPD patients. Apparently, shorter T1 at ultra-short echo times is more correlated with tissue abnormalities, whereas at longer echo times it is more correlated with perfusion (7). Thus, this technique allows for a sub-resolution assessment of tissue composition.

Finally, the systematic review by Hofmann et al. on the effects of vaping on the lungs is a nice summary of the different techniques used (CT, MRI and FDG-PET/CT) to study the structural and functional impairment caused by inhalative toxins.

Author contributions

MW: Conceptualization, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JW: Conceptualization, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. EB: Conceptualization, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

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.

References

1. Lynch DA, Austin JHM, Hogg JC, Grenier PA, Kauczor H-U, Bankier AA, et al. CT-Definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner Society. *Radiology*. (2015) 277:141579–141579. doi: 10.1148/radiol.2015141579

2. Galban CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, et al. Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. *Nat Med*. (2012) 18:1711–5. doi: 10.1038/nm.2971

3. Wielpütz M, Kauczor HU. MRI of the lung: state of the art. *Diagn Interv Radiol*. (2012) 18:344–53. doi: 10.4261/1305-3825.DIR.5365-11.0

4. Woods JC, Wild JM, Wielpütz MO, Clancy JP, Hatabu H, Kauczor HU, et al. Current state of the art MRI for the longitudinal assessment of cystic fibrosis. *J Magn Reson Imaging*. (2020) 52:1306–20. doi: 10.1002/jmri.27030

5. Mussell GT, Marshall H, Smith LJ, Biancardi AM, Hughes PJC, Capener DJ, et al. Xenon ventilation MRI in difficult asthma: initial experience in a clinical setting. *ERJ Open Res*. (2021) 7:3. doi: 10.1183/23120541.00785-2020

6. Bauman G, Bieri O. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. *Magn Reson Med*. (2017) 77:336–42. doi: 10.1002/mrm.26096

7. Triphan SMF, Weinheimer O, Gutberlet M, Heussel CP, Vogel-Claussen J, Herth F, et al. Echo time-dependent observed lung T1 in patients with chronic obstructive pulmonary disease in correlation with quantitative imaging and clinical indices. *J Magn Reson Imaging*. (2021) 54:1562–71. doi: 10.1002/jmri.27746

GOLD stage-specific phenotyping of emphysema and airway disease using quantitative computed tomography

Philip Konietzke[1,2,3]*, Christian Brunner[1,2,3], Marilisa Konietzke[1,2], Willi Linus Wagner[1,2,3], Oliver Weinheimer[1,2,3], Claus Peter Heußel[1,2,3], Felix J. F. Herth[2,4], Franziska Trudzinski[2,4], Hans-Ulrich Kauczor[1,2,3] and Mark Oliver Wielpütz[1,2,3]

[1]Department of Diagnostic and Interventional Radiology, University Hospital of Heidelberg, Heidelberg, Germany, [2]Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), University of Heidelberg, Heidelberg, Germany, [3]Department of Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik at University of Heidelberg, Heidelberg, Germany, [4]Department of Pulmonology, Thoraxklinik at University of Heidelberg, Heidelberg, Germany

Background: In chronic obstructive pulmonary disease (COPD) abnormal lung function is related to emphysema and airway obstruction, but their relative contribution in each GOLD-stage is not fully understood. In this study, we used quantitative computed tomography (QCT) parameters for phenotyping of emphysema and airway abnormalities, and to investigate the relative contribution of QCT emphysema and airway parameters to airflow limitation specifically in each GOLD stage.

Methods: Non-contrast computed tomography (CT) of 492 patients with COPD former GOLD 0 COPD and COPD stages GOLD 1–4 were evaluated using fully automated software for quantitative CT. Total lung volume (TLV), emphysema index (EI), mean lung density (MLD), and airway wall thickness (WT), total diameter (TD), lumen area (LA), and wall percentage (WP) were calculated for the entire lung, as well as for all lung lobes separately. Results from the 3rd-8th airway generation were aggregated (WT_{3-8}, TD_{3-8}, LA_{3-8}, WP_{3-8}). All subjects underwent whole-body plethysmography (FEV1%pred, VC, RV, TLC).

Results: EI was higher with increasing GOLD stages with $1.0\pm1.8\%$ in GOLD 0, $4.5\pm9.9\%$ in GOLD 1, $19.4\pm15.8\%$ in GOLD 2, $32.7\pm13.4\%$ in GOLD 3 and $41.4\pm10.0\%$ in GOLD 4 subjects ($p <0.001$). WP_{3-8} showed no essential differences between GOLD 0 and GOLD 1, tended to be higher in GOLD 2 with $52.4\pm7.2\%$, and was lower in GOLD 4 with $50.6\pm5.9\%$ ($p =0.010 - p =0.960$). In the upper lobes WP_{3-8} showed no significant differences between the GOLD stages ($p =0.824$), while in the lower lobes the lowest WP_{3-8} was found in GOLD 0/1 with $49.9\pm6.5\%$, while higher values were detected in GOLD 2 with $51.9\pm6.4\%$ and in GOLD 3/4 with $51.0\pm6.0\%$ ($p <0.05$). In a multilinear regression analysis, the dependent variable FEV1%pred can be predicted by a combination of both the independent variables EI ($p <0.001$) and WP_{3-8} ($p <0.001$).

Conclusion: QCT parameters showed a significant increase of emphysema from GOLD 0–4 COPD. Airway changes showed a different spatial pattern with higher values of relative wall thickness in the lower lobes until GOLD 2 and subsequent lower values in GOLD3/4, whereas there were no significant differences in the upper lobes. Both, EI and WP_{5-8} are independently correlated with lung function decline.

KEYWORDS

chronic obstructive pulmonary disease (COPD), quantitative CT (QCT), GOLD stages, airway disease, lung emphysema

Highlights

1. QCT showed significant differences between GOLD 0–4.
2. Airway parameters indicate a transient inflammatory response in GOLD 2 leading to airway destruction in GOLD 4.
3. QCT could detect spatial differences between upper and lower lung lobes.
4. FEV1%pred appears to be predicted by a linear combination of the independent variables EI and WP_{3-8}.

Introduction

Chronic obstructive pulmonary disease (COPD) is the fourth leading cause of death worldwide and typically results from prolonged inhalation of noxious particles (1, 2). The diagnosis is made by symptoms and pulmonary function testing (PFT), and severity is commonly classified according to the GOLD criteria (3). Early detection of the disease is of great interest because airway disease is potentially reversible with smoking cessation or appropriate treatment, thereby delaying irreversible disease progression (4, 5). PFT play a central role in COPD diagnosis, yet its role in early diagnosis and reproducibility are limited. FEV1 is often normal until more than 30% of lung tissue are damaged or more than 75% small airways are obstructed (6, 7) due to compensation mechanisms of the lungs. Koo et al. showed that airway changes such as thickened bronchiolar walls, decreased number of bronchioles, and narrowed bronchiolar lumen occur earlier than the presence of emphysema in patients with mild GOLD COPD (8).

The use of computed tomography (CT) has led to better characterization of disease heterogeneity across patients at risk of COPD (former GOLD 0) to late stages with advanced emphysema. CT imaging allows deeper insights into different disease phenotypes and more accurate assessment of disease severity and distribution (9–11). Moreover, software-based post-processing (QCT) may detect and quantify the presence and type of emphysema by analysis of lung density (12–14). The assessment of airway disease is more challenging and less validated (12). QCT of small airways disease (SAD) is particularly challenging because CT measurements of airways are accurate and reproducible up to an diameter of approximately 2 mm, but those airways may serve as a surrogate for the smaller airways (15, 16). All these parameters are heavily influenced by technical factors such as scanner type, contrast material, slice thickness, kernel, and post-processing software. However, larger airways that can be visualized on CT allow conclusions about the status of smaller airways (16, 17). At present, larger studies on the quantitative contribution of emphysema and airways disease to airflow limitation are sparse.

Thus, we hypothesized that emphysema and airway abnormalities are independent factors determining airflow limitation in COPD, and that both have different systematic impact in different GOLD stages. In this work-up, we used QCT for GOLD-stage specific differential characterization of emphysema and airway abnormalities in a relatively large group of 522 patients with former GOLD 0 COPD and GOLD stages COPD 1–4 were examined with a comparable protocol on the same scanner.

Materials and methods

Patients

This retrospective study was approved by the institutional ethics committee (S-646/2016). The patient cohort was retrospectively recruited from the institutional imaging database. All adult patients who had an inspiratory CT scan between 08/2016 and 01/2020 at our chest-hospital were included if whole-body plethysmography was available within 45 days. The clinical exclusion criteria were pulmonary infections, lung tumors >1 cm, prior lung surgery, or volume-reducing procedures, and technical exclusion criteria were use of contrast media, image artifacts, missing or other reconstructions as I40f\3 and I70f\3, slice thickness of 1.25 mm, or errors in data export. All examinations were visually inspected for the absence of significant motion artifacts and inclusion of all parts of the chest by a senior chest radiologist.

All patients were diagnosed with COPD according to GOLD 2020 (18). In addition to subjects of the GOLD 1–4 categories, smokers and former smokers with no assignable GOLD category including the "former GOLD 0" were enrolled. The "former GOLD 0" group includes subjects with normal PFT in terms of FEV1/FVC (ratio between forced expiratory volume in 1 s and forced vital capacity) but with COPD-specific symptoms. Based on these criteria, 5 study groups (GOLD 0–4) were defined.

Whole-body plethysmography

All patients had whole-body plethysmography 0–45 days within the CT scan using reference values according to the Global Lung Initiative (19). In this study, vital capacity (VC), forced vital capacity (FVC), residual volume (RC), total lung capacity (TLC), forced expiratory volume in 1 s (FEV1), FEV1 predicted (FEV1%) and the ratio FEV1/FVC were used.

CT acquisition and reconstruction

Non-contrast CT (Somatom Definition AS64, Siemens Healthineers AG) was performed in supine position, as recommended for COPD (20). All patients were instructed and carefully monitored for a stable full automatically instructed inspiratory breath-hold before scanning. The CT scanner was routinely calibrated every 3 months for water and daily for air. Scans were performed in caudocranial direction with a dose-modulated protocol (Caredose4D, Siemens Healthineers AG) using a reference of 120 kV and 70 mA or 100 kV

and 117 mA (120/70//100/117 [kV/mA]; GOLD 0 = 28/95, GOLD 1 = 9/24, GOLD 2 = 44/71, GOLD 3 = 83/98, GOLD 4 = 16/24) at a collimation of 64 × 0.6 mm and a pitch of 1.45. The reconstructed slice thickness was 1.00 mm with 0.825 mm increment. For each patient, a set of two reconstructions was available, namely a medium soft reconstruction algorithm (I40f\3) and the edge-enhancing reconstruction algorithm (I70f\3).

Quantitative post-processing

The in-house software YACTA, a non-commercial scientific software, segmented the airway tree and lung lobes fully automated, using the I40f\3 kernel for parenchymal analysis and the I70f\3 kernel for airway analysis as previously published (21–26). The total lung volume (TLV), emphysema index (EI) and mean lung density (MLD) were calculated for the total lung and for all lung lobes separately (right upper (RUL), middle (RML), and lower (RLL) lobe, as well as left upper lobe (LUL), lingula (LLi) and left lower lobe (LLL)). The airway parameters wall thickness (WT), total diameter (TD), lumen area (LA), and wall percentage (WP) were analysed generation-based in the trachea (G_1), right and left main stem (G_2), lobar (G_3), segmental (G_4), and the subsegmental bronchi (G_{5-8}). Airway results were simplified by consolidating the generation-based results for the 3rd to 8th generation (G_{3-8}).

Statistical analysis

Statistical analyses were performed using R (R 3.3.2, Foundation for Statistical Computing) and SigmaPlot (Systat Software GmbH). Data are presented as mean ± standard deviation One-way analysis of variance (ANOVA), Tukey multiple pairwise-comparisons (Tukey Test), Spearman linear and multiple linear regression were used. Correlation coefficients were interpreted as follows: 0.00–0.10 (negligible), 0.10–0.39 (weak), 0.40–0.69 (moderate), 0.70–0.89 (strong), and 0.90–1.00 (very strong) (27). A value of $p < 0.05$ was considered statistically significant.

Results

Patient cohort and demographics

The final cohort included 522 patients out of 15,631. 10,358 patients were excluded because of contrast media application, 4,503 patients because of clinical exclusion criteria. 275 patients were excluded due to technical exclusion criteria significant motion artifacts (74), incomplete inclusion of all parts of the chest (28), invalid export from PACS system (12), slice thickness 1.25 mm (27), missing I70f\3 reconstruction (77), other reconstructions as I40f\3 (conventional filtered backprojection) (21), B40f (11) and B40s (18) kernel. The final cohort consisted of 492 patients, of whom 29 patients were diagnosed with GOLD 0, 34 with GOLD 1 COPD, 123 with GOLD 2 COPD, 196 with GOLD 3 COPD, and 40 with GOLD 4 COPD.

Patients with GOLD 0 COPD had the highest vital capacity (VC) with 3.6 ± 1.2 L and the lowest residual volume (RV) with 2.6 ± 0.8 L as

FIGURE 1
Patient recruitment flowchart.
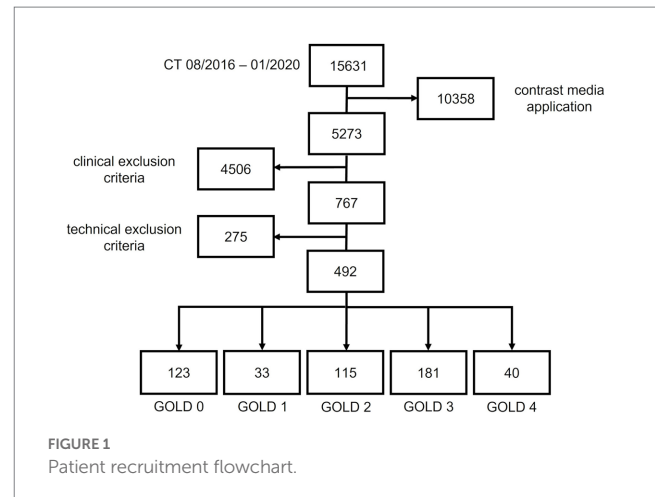
well as the lowest total lung capacity (TLC) with 6.2 ± 1.4 L. RV and TLC were higher in each GOLD stage form GOLD 0 to 3, being slightly lower again at GOLD 4 (all $p < 0.001$). FEV1%pred was lower at each GOLD stage as per definition ($p < 0.001$) (Figure 1; Table 1).

GOLD stage-specific quantification of emphysema and airway disease

TLV was 5198 ± 1349 cm³ in GOLD 0, higher with 7405 ± 1366 cm³ in GOLD 3 ($p < 0.001$) and lower in GOLD 4 with 6892 ± 1292 cm³ ($p = 0.193$). Emphysema was measured with an EI of 1.17 ± 1.8% in GOLD 0 and of 4.5 ± 9.9% in GOLD 1 ($p = 0.564$). EI was higher with increasing GOLD stages with 19.4 ± 15.8% in GOLD 2 ($p < 0.05$), 32.7 ± 13.4% in GOLD 3 ($p < 0.05$), and 41.4 ± 10.0% in GOLD 4 ($p < 0.05$). Accordingly, MLD showed a comparable pattern, with −802 ± 38 HU in GOLD 0 and progressively lower values in GOLD 1 with −820 ± 32 HU ($p = 0.044$), in GOLD 2 with −850 ± 30 HU ($p < 0.05$), and in GOLD 3 with −872 ± 19 HU ($p < 0.05$), while the MLD was again slightly higher in GOLD 4 with −876 ± 14 HU ($p = 0.971$) (Figures 2, 3; Table 2).

WT_{3-8} showed no essential differences between GOLD 0 and GOLD 1 with 1.17 ± 0.21 mm and 1.14 ± 0.16 mm ($p = 0.960$), respectively. In GOLD 2, it tended to be higher with 1.24 ± 0.24 mm ($p = 0.101$), while it was lower with 1.17 ± 0.19 mm in GOLD 3 ($p < 0.05$) and with 1.15 ± 0.14 mm in GOLD 4 ($p = 0.833$). WP_{3-8} showed a similar trend as there were no significant differences between GOLD 0 and GOLD 1, the highest values were found in GOLD 2, while the values in GOLD 3 and GOLD 4 were again lower. TD_{3-8} and LA_{3-8} showed no significant differences between the GOLD stages (Figures 2, 4; Table 2).

Lobe-based quantification of emphysema and airway disease

A lobe-based analysis was performed, whereby only lobes with a complete set of results for the airway generations 3rd-8th were included, which was the case for 913 out of the theoretical set of 2952 lobes. The lingula and middle lobe were excluded entirely due to a high number of missing results for individual generations. The lobe-based results

TABLE 1 Patient demograpics and body plerhysmography parameters.

	GOLD 0	GOLD 1	GOLD 2	GOLD 3	GOLD 4	pANOVA
N	123	33	115	181	40	-
Sex [f/m]	59/64	16/17	53/62	74/74	26/14	-
Age [y]	55 ± 15	63 ± 12*	63 ± 11	61 ± 10	64 ± 8	0.006
Height [cm]	1.71 ± 0.10	1.68 ± 0.11	1.68 ± 0.10	1.69 ± 0.08	1.66 ± 0.07	0.023
Weight [kg]	83 ± 19	79 ± 16	75 ± 18	72 ± 18	64 ± 12*	0.037
BMI	28 ± 5	28 ± 5	26 ± 5	25 ± 6	23 ± 4	0.001
TLC	6.2 ± 1.4	6.4 ± 1.7	7.8 ± 5.5	8.8 ± 4.5	8.1 ± 1.2	0.001
RV	2.6 ± 0.8	3.2 ± 1.5	4.8 ± 1.3*	5.8 ± 1.4*	5.4 ± 0.5	0.001
VC	3.6 ± 1.2	3.3 ± 1.2	2.4 ± 1.1*	2.7 ± 0.8	2.2 ± 0.8	0.001
FVC	3.5 ± 1.2	3.1 ± 1.2	2.2 ± 1.1*	2.3 ± 0.8	2.2 ± 0.9	0.001
FEV1 [L]	2.8 ± 0.9	2.2 ± 0.8	1.1 ± 0.6	0.9 ± 0.4	0.69 ± 0.23	<0.001
FEV1%pred	101.3 ± 6.4	88.3 ± 6.6	61.5 ± 9.7	41.7 ± 6.1	27.7 ± 2.3	<0.001
FEV1/FVC	0.8 ± 0.01	0.7 ± 0.1	0.5 ± 0.1	0.4 ± 0.1	0.28 ± 0.05	<0.001

Patient demographics (sex, age, height, weight, and BMI) and body plethysmography parameters total lung capacity (TLC), residual volume (RC), (vital capacity (VC) forced vital capacity (FVC), forced expiratory volume in 1 s (FEV1), FEV1 predicted (FEV1%), and FEV1/FVC ratio) are shown. All data are given as mean ± standard deviation. *$p < 0.05$ vs. previous GOLD stage.

FIGURE 2
Boxplots for total lung volume (TLV), emphysema index (EI), wall thickness (WT_{3-8}), total diameter (TD_{3-8}), lumen area (LA_{3-8}) and wall percentage (WP_{3-8}) for the COPD stages 0−4. Box mark 25-75th percentile, whiskers indicate 5th and 95th percentile, and individual outliers are given by black filled circles. Individual p-Values vs. previous GOLD stage are indicated.
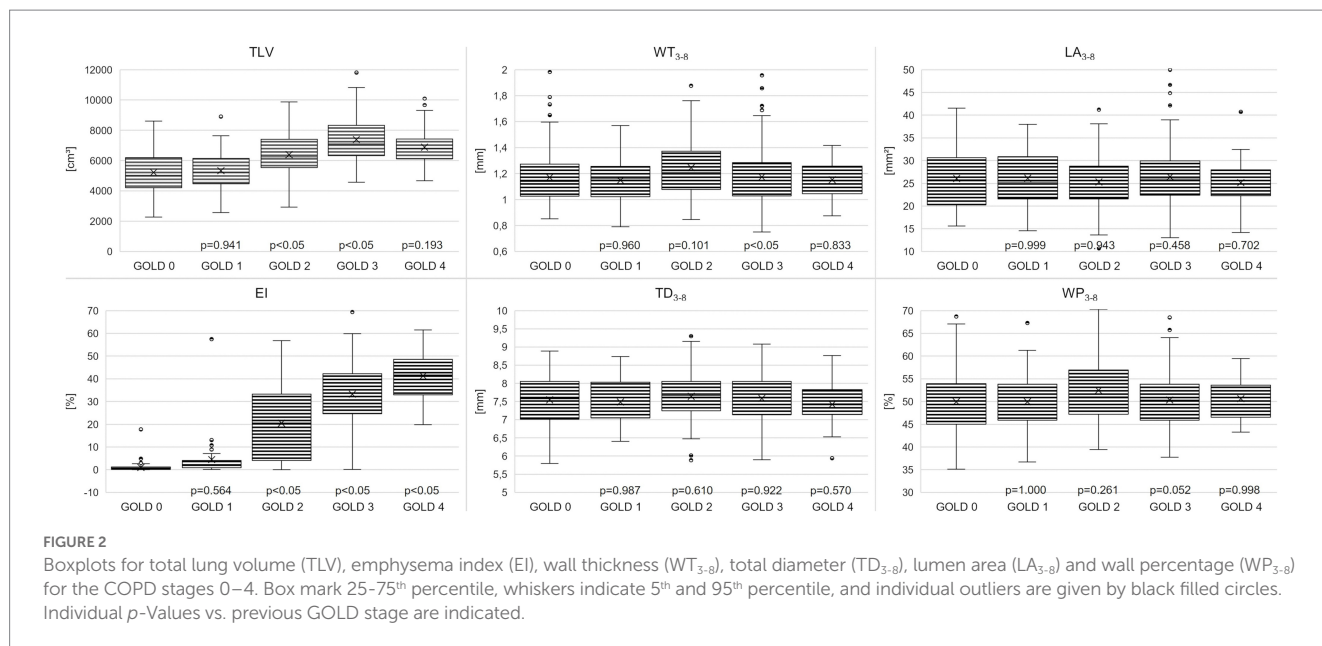
for GOLD 0 and 1, and GOLD 3 and 4 were combined into GOLD0/1 and GOLD3/4, respectively, to compensate for limited complete datasets. The lobe volume difference between GOLD 0/1 and GOLD 3/4 was higher in the upper lobes with a difference of 288 cm³ (+20%), than in the lower lobes with a difference of 162 cm³ (+10%). The EI showed no significant differences between upper and lower lobes in GOLD 1, while EI was higher in the upper lobes in GOLD 2 and GOLD3/4. The values for WP_{3-8} showed different trends for the upper and lower lobes. In the upper lobes, WP_{3-8} was overall lower than in the lower lobes and tended to be higher with increasing GOLD stages ($p = 0.887$, $p = 0.928$). In the lower lobes, WP_{3-8} was 49.9 ± 6.5% in GOLD 0/1, higher in GOLD 2 with 51.9 ± 6.4% ($p < 0.05$) and tended to be lower in GOLD 3/4 with 51.0 ± 6.0%. In the upper lobes, TD_{3-8} and LA_{3-8} tended to be higher with higher GOLD stages ($p = 0.201$, $p = 0.307$). In the lower lobes LA_{3-8} showed no significant differences,

while TD_{3-8} had the highest values at GOLD 2 ($p = 0.216$, $p = 0.866$) (Figure 5; Table 3).

Emphysema and airway disease are independent factors of airflow limitation

TLC showed strong correlations with TLV for all GOLD stages ($r = 0.79$–0.86), low to moderate correlations with EI ($r = 0.29$–0.64), and weak to moderate correlations with WP_{3-8} ($r = −0.10 - -0.51$). RV showed strong to correlations with TLV and EI for GOLD 0–4 ($r = 0.71$, $r = 0.77$), while the correlations with WP_{3-8} were negligible to moderate ($r = −0.04 - -0.54$). The correlations between VC and TLV, EI and WP_{3-8} were overall weak for GOLD 0–4 ($r = −0.13$–0.18). The correlation of FEV1%pred for GOLD 0–4 showed moderate

FIGURE 3
(Top line) Original CT images with increasing emphysema from GOLD 0 to GOLD 4. (Bottom line) Emphysema index (EI) parameter maps (emphysema=yellow) for the 5 different patients with GOLD 0 and GOLD 1 to 4 GOLD COPD.

TABLE 2 QCT parameters for GOLD stages 0−4.

	GOLD 0	GOLD 1	GOLD 2	GOLD 3	GOLD 4	pANOVA
TLV [cm³]	5198 ± 1349	5401 ± 1286	6353 ± 1368	7405 ± 1366	6892 ± 1292	<0.001
EI [%]	1.0 ± 1.8	4.5 ± 9.9	19.4 ± 15.8	32.7 ± 13.4	41.4 ± 10.0*	<0.001
MLD [HU]	−802 ± 38	−820 ± 32*	−850 ± 30	−872 ± 19	−876 ± 14	<0.001
WT_{3-8} [mm]	1.17 ± 0.21	1.14 ± 0.16	1.24 ± 0.24	1.17 ± 0.19	1.15 ± 0.14	<0.05
TD_{3-8} [mm]	7.54 ± 0.68	7.47 ± 0.55	7.63 ± 0.64	7.30 ± 0.66	7.42 ± 0.54	0.328
LA_{3-8} [mm²]	26.05 ± 6.62	26.11 ± 6.07	25.47 ± 5.67	26.60 ± 6.05	25.16 ± 4.86	0.626
WP_{3-8} [%]	50.0 ± 7.2	49.8 ± 6.2	52.4 ± 7.2	50.3 ± 5.9	50.6 ± 4.6	<0.05

Total lung volume (TLV), emphysema index (EI) and mean lung density (MLD) as well as wall thickness (WT_{3-8}), total diameter (TD_{3-8}), lumen area (LA_{3-8}) and wall percentage (WP_{3-8}) were calculated. Generation-based results for WT, TD, LA, and WP were consolidated into one value (G_{3-8}). All data are given as mean ± standard deviation. *$p < 0.05$ vs. previous GOLD stage.

correlations with TLV ($r = −0.59$, $p = 0.001$) and strong correlations with EI ($r = 0.78$, $p = 0.002$), while the correlations for the individual GOLD stages were low to moderate ($r = −0.10–0.47$). The correlation between FEV1%pred and WP_{3-8} was negligible to moderate ($r = 0.01–0.39$) for GOLD 0–4 and all individual GOLD stages (Table 4; Supplementary Table 1).

Importantly, in a multilinear regression analysis the dependent variable FEV1%pred can be predicted by a combination of both the independent variables EI ($p < 0.001$) and WP_{3-8} ($p < 0.001$), whereas the independent variables biological sex ($p = 0.893$) and age ($p = 0.598$) could not (Figure 6; Table 4).

Discussion

In this work, we demonstrate that emphysema and airway disease independently contribute to airflow limitation in COPD, in both high-risk patients and patients with COPD at GOLD stages 1–4, using QCT in 522 individuals studied with the same CT scanner and protocol. Second, we show that individual lobes show distinct differences of emphysema and airway disease at different GOLD stages with a

relatively higher emphysema values in the upper lobes compared to the lower lobes. Thus, we found typical values of EI and WP for each GOLD-stages.

Total lung volume (TLV) was higher in GOLD 3 than in GOLD 0, which is within expectations as emphysema progression is associated with hyperinflation. However, TLV was again lower in GOLD 4, which may represent scarring in advanced destructive emphysema, but could also be related to increased mortality with a selection bias. The EI increased across all GOLD stages, but only significantly between GOLD 1 and GOLD 2, and GOLD 2 and GOLD 3, suggesting that the severity of the emphysema progression may not be linear. This contradicts the assumption that emphysema progression might be accelerated in advanced COPD, since locally severely altered alveolar micromechanics within an injured lung might itself 'become an independent trigger of lung injury progression' (29, 30). The explanation may be that in GOLD 4, emphysema is nevertheless higher in relation to the remaining normal lung tissue, although the absolute difference between GOLD 3 and GOLD 4 is lower. Overall, our results are consistent with the literature in which QCT has been successfully used to detect the progression of emphysema (31, 32). MLD increased also significantly from GOLD 0 to GOLD 4, as MLD
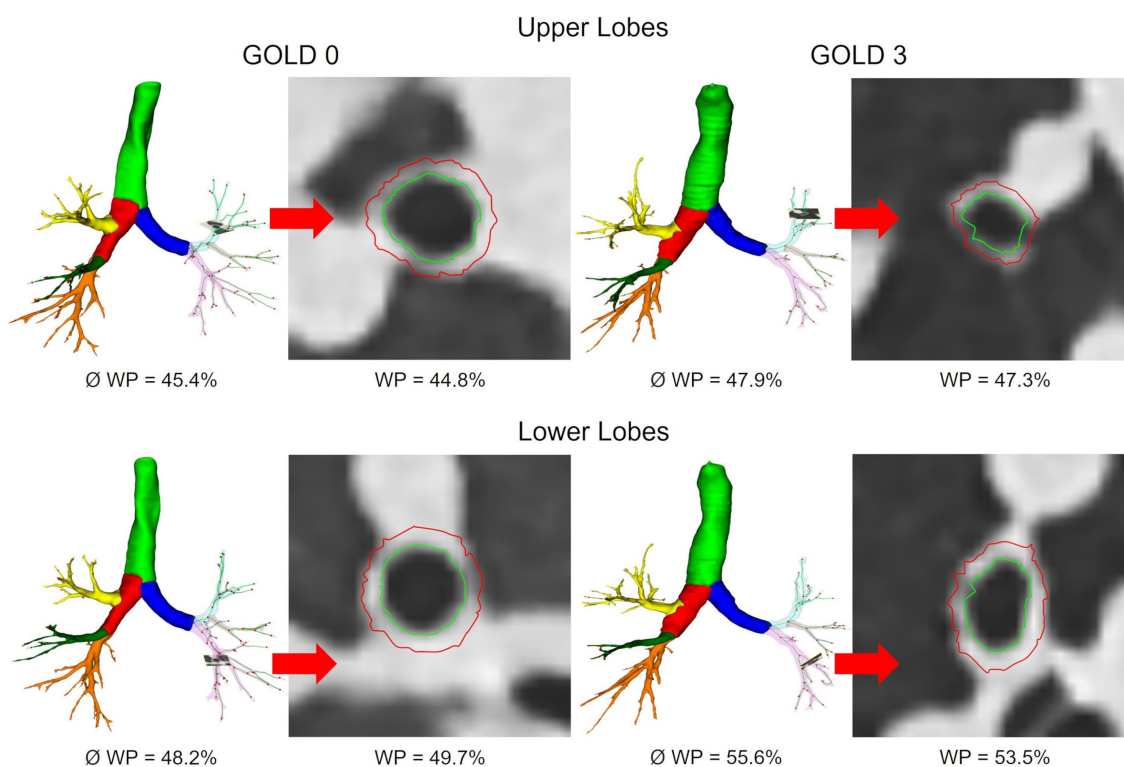
FIGURE 4

Segmented airway tree of the upper and lower lobes in GOLD 0 and GOLD 3. Average WP for the entire airway tree (Ø) and representative slices for individual bronchi are shown.
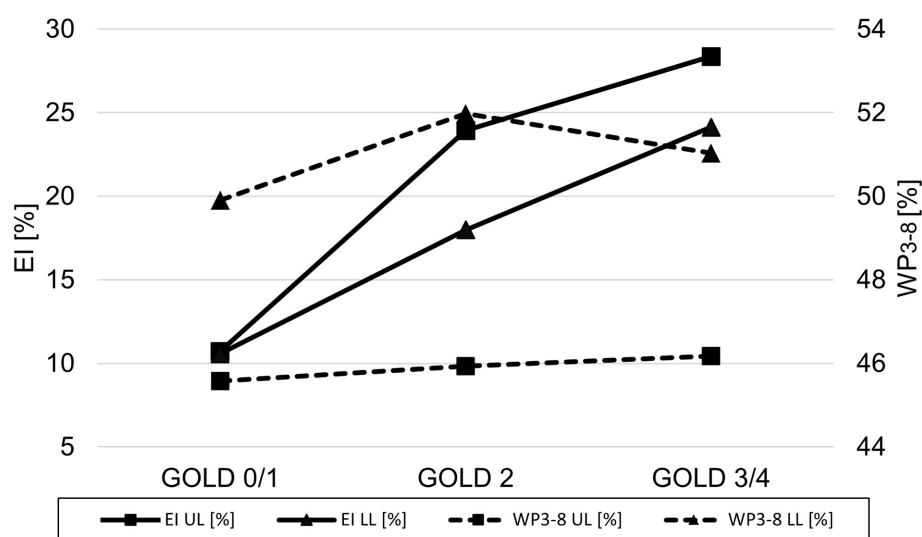
FIGURE 5

Line plot of emphysema index (■ EI) and wall percentage (— — - WP_{3-8}) for GOLD 0/1, GOLD 2, and GOLD 3/4 for the upper lobes (UL=square) and the lower lobes (LL=triangle).

and EI are strongly correlated. In this context, it should be mentioned that MLD in advanced COPD is influenced by both emphysema and lung fibrotic changes, the latter might explain the slightly higher MLD values in GOLD 4 when compared with GOLD 3. The use of the parameter Perc15 can be helpful to overcome this problem.

WT_{3-8} and WP_{3-8} showed no relevant differences between GOLD 0 and GOLD 1, most likely because the disease-related changes may be too subtle for detection. WT_{3-8} and WP_{3-8} showed the highest values at GOLD 2, accompanied by higher TD_{3-8} and a slightly lower LA_{3-8}, which could be explained by inflammatory swelling of the

TABLE 3 QCT parameters for the combined upper and lower lobes at different GOLD stages.

	GOLD 0/1	GOLD 2	GOLD 3/4	pANOVA
Upper lobes				
N	124	74	120	–
ULV [cm³]	1140 ± 387	1355 ± 437*	1428 ± 546	<0.001
EI [%]	11.0 ± 17.8	24.5 ± 17.6*	27.7 ± 22.5	<0.001
MLD [HU]	−833 ± 41	−864 ± 29*	−864 ± 39	<0.001
WT_{3-8} [mm]	1.10 ± 0.17	1.14 ± 0.21	1.15 ± 0.20	0.014
TD_{3-8} [mm]	7.86 ± 0.95	8.07 ± 1.07	8.25 ± 1.00	<0.05
LA_{3-8} [mm²]	32.01 ± 10.68	33.78 ± 10.58	36.40 ± 13.41	<0.05
WP_{3-8} [%]	45.6 ± 5.7	46.0 ± 5.0	46.0 ± 5.8	0.824
Lower lobes				
N	211	134	250	–
LLV [cm³]	1358 ± 422	1392 ± 476	1520 ± 482*	<0.001
EI [%]	11.0 ± 18.1	18.1 ± 17.2*	24.0 ± 19.1*	<0.001
MLD [HU]	−816 ± 51	−837 ± 41*	−845 ± 46	<0.001
WT_{3-8} [mm]	1.27 ± 0.20	1.35 ± 0.24	1.30 ± 0.23	<0.05
TD_{3-8} [mm]	8.33 ± 0.94	8.48 ± 0.95	8.41 ± 0.86	0.373
LA_{3-8} [mm²]	31.93 ± 9.98	31.79 ± 9.02	31.79 ± 8.14	0.924
WP_{3-8} [%]	49.9 ± 6.5	51.9 ± 6.4	51.0 ± 6.0	<0.05

Lobe volumes for the upper (ULV) and lower lobes (LLV), emphysema index (EI) and mean lung density (MLD), wall thickness (WT_{3-8}), total diameter (TD_{3-8}), lumen area (LA_{3-8}), and wall percentage (WP_{3-8}) are shown. All data are given as mean ± standard deviation. *$p < 0.05$ vs. previous GOLD stage.

airway walls in the process of bronchitis. This observation is supported by the literature, which recognizes small airway disease as a central feature of COPD, with increasing narrowing and destruction of the small airways representing a mixture of chronic inflammation (28, 33, 34). In this context, Koo et al. demonstrated significant loss of terminal and transitional bronchioles in patients with GOLD 1 and GOLD 2, as the remaining small airways had thickened walls and narrowed lumens that were also present in regions with no emphysema (8).

However, in contrast to parts of the literature, WT_{3-8}, WP_{3-8} and TD_{3-8} tended to be lower in GOLD 3 and GOLD 4, while LA_{3-8} was higher in GOLD 3 and again lower in GOLD 4. These changes may be interpreted as a possible transition from reversible airway inflammation to irreversible airway damage with airway wall degradation. This theory is supported by Smith et al., who observed significantly reduced airway wall thickness in most areas of the central tracheobronchial tree in COPD patients (35). They hypothesize that possible mechanisms include airway smooth muscle regression, apoptosis or replacement fibrosis resulting from chronic airway inflammation, and decreased bronchial vascular volume (36, 37). In addition, increasing emphysematous destruction of the lung parenchyma appears to have a significant impact on airway dimensions (38). In detail, emphysema leads to destruction of the lung parenchyma, resulting in loss of lung attachments that stabilize the airways and prevent them from collapse. Therefore, an increase in emphysema should result in partial airway collapse and thus a decrease in LA and TD, whereas WT and WP, unless the airway wall mass itself has not changed, should increase. On the other hand, hyperinflation leads to an increase in lung volume, stretching the airways and most likely increasing LA and TD and decreasing WT

and WP, having a partially opposing effects on the airways. In our cohort, the reduction of LA_{3-8} in GOLD 4 might contribute to the destabilizing effects of emphysema.

We paid attention to potential regional differences within the lung by analysing individual lung lobes. EI was significantly higher in both upper lobes, resulting in greater volume increase due to emphysematous hyperinflation, consistent with the literature. (39). The spatial differences in airway involvement have been investigated in only a few studies. The COPDGene Investigators reported that there were no differences in airway wall thickness or Pi10 between patients with predominant lower and predominant upper lobe emphysema, although the former group had greater airflow limitation and more air-trapping (40). Park et al. showed that patients with predominant lower lobe emphysema showed greater airway involvement than those with predominant upper lobe emphysema, possibly leading to more frequent exacerbations and poorer response to therapy (41). In accordance with Park et al., our results also showed significant higher WP_{3-8} in the lower lobes. However, looking at the pooled results for GOLD 0/1, GOLD 2, and GOLD 3/4, we see that WP_{3-8} was higher in the upper lobes with increasing GOLD stages, while in the lower lobes it was the highest in GOLD 2 and again lower in GOLD 3/4. Interestingly, we would have expected a lower WP_{3-8} in higher GOLD stages in the upper lobes, as advanced emphysema should be associated with airway degeneration and collapse. One possible explanation may be the pooling of GOLD 3 and GOLD 4 data, which was necessary because airway segmentation becomes more difficult as disease progresses. Another reason could be the initial differences in WP_{3-8} between upper and lower lobes, which may influence the changes of airway dimensions in relation to the GOLD stage.

TABLE 4 Spearman rank order correlation coefficient for lung function parameters and QCT.

	GOLD 0–4	GOLD 0	GOLD 1	GOLD 2	GOLD 3	GOLD 4
TLC						
TLV [cm³]	0.86 (0.001)	0.83 (0.001)	0.79 (0.001)	0.78 (0.001)	0.81 (0.001)	0.83 (0.001)
EI [%]	0.56 (0.001)	0.29 (0.001)	0.64 (0.001)	0.26 (0.001)	0.27 (0.001)	0.17 (0.399)
WP$_{3-8}$ [%]	−0.12 (0.012)	−0.11 (0.245)	−0.51 (0.002)	−0.17 (0.078)	−0.10 (0.165)	−0.25 (0.211)
RV						
TLV [cm³]	0.71 (0.001)	0.38 (0.001)	0.48 (0.001)	0.56 (0.001)	0.65 (0.001)	0.43 (0.297)
EI [%]	0.77 (0.001)	0.10 (0.271)	0.69 (0.001)	0.56 (0.001)	0.35 (0.001)	0.49 (0.286)
WP$_{3-8}$ [%]	−0.04 (0.386)	−0.07 (0.392)	−0.28 (0.118)	−0.16 (0.089)	−0.08 (0.252)	−0.54 (0.297)
VC						
TLV [cm³]	0.18 (0.001)	0.73 (0.001)	0.63 (0.001)	0.33 (0.001)	0.81 (0.001)	0.47 (0.001)
EI [%]	−0.38 (0.001)	0.25 (0.001)	0.1 (0.577)	−0.28 (0.001)	0.27 (0.001)	0.16 (0.331)
WP$_{3-8}$ [%]	−0.13 (0.001)	−0.09 (0.329)	−0.29 (0.101)	−0.08 (0.411)	−0.11 (0.165)	−0.05 (0.725)
FEV1%pred						
TLV [cm³]	−0.59 (0.001)	−0.1 (0.271)	−0.02 (0.926)	−0.31 (0.001)	−0.15 (0.042)	0.26 (0.658)
EI [%]	−0.78 (0.001)	−0.16 (0.069)	−0.18 (0.311)	−0.43 (0.001)	−0.13 (0.076)	−0.71 (0.136)
WP$_{3-8}$ [%]	0.01 (0.883)	0.01 (0.933)	0.39 (0.025)	0.04 (0.665)	0.02 (0.754)	0.31 (0.564)

Spearman rank order correlation coefficient were calculated for total lung volume (TLV), emphysema index (EI) and wall percentage (WP) with total lung capacity (TLC), residual volume (RV), vital capacity (VC) and FEV1%predicted (FEV1%pred).

In our study cohort, strong correlations between parenchymal QCT and whole-body plethysmography parameters were observed. The QCT parameter TLV, which increased from GOLD stage 0 to 3, showed strong correlations with TLC determined by whole-body plethysmography, which has already been described (42). TLV showed also low to moderate correlations with EI as emphysema progression is associated with hyperinflation. The residual volume (RV) showed strong correlations with TLV and EI for GOLD 0–4, as the main volume that increases with COPD severity is RV (43). The correlations between vital capacity (VC) and the CT parameters TLV, EI were weak for GOLD 0–4. The correlation between FEV1%pred and TLV as well as EI was strong since its connection is well established in the literature (44). In comparison whole-body plethysmography parameters TLC, RC and VC showed at best weak correlations with QCT airway parameters. However, WP$_{3-8}$ in GOLD 1 and EI in GOLD 4 appeared to correlate better with FEV1%pred, suggesting that FEV1 impairment may be more attributable to airway disease in the lower GOLD stages and more to emphysema in the higher GOLD stages. This may also be reflected in previous observations in the COPDGene study, demonstrating that small airways disease increases only from GOLD 0 to GOLD 4, and declines thereafter, whereas emphysema increases with every GOLD stage. Of note, results from larger airway analyses as in our present study can be regarded as surrogates for processes in the small airways not visible at CT (45). Restrictively, this observation may also be to selection bias since the patient numbers in GOLD 1 and GOLD 4 were relatively low. However, this observation also highlights the limitations of FEV1, as it is influenced by at least two very different pathophysiological mechanisms, namely airway wall thickening for inflammation (also in vessels) and attachment disruption (and vascular disruption) for protease activity in emphysema. In this study, this was also shown with multilinear regression, where the dependent variable FEV1%pred can be predicted by a linear combination of the two independent variables EI and WP$_{3-8}$. There was no association with biological sex and age,

which can be expected since FEV1% predicted is already normalized by the age, sex and other factors. However, emphysema, airway wall thickness (WT), total diameter (TD), and lumen area (LA) are dependent on sex and age (46, 47). In this context, it should be emphasized that other lung function parameters such as DCLO or FRC may be better suited to describe the different contributions of emphysema and airway disease, as already shown in the literature (48).

Our study has some limitations. The interpretation of quantitative parameters should be done carefully, as subtle changes may be due to noise or measurement errors. We tried to reduce technical confounders by using the same scanner, slice thickness, and reconstruction kernels. In our study cohort a dose-modulated protocol was used with a reference of 120 kV/70 mA (n = 180) or 100 kV/117 mA (n = 312), which may influence the results. However, two phantom based studies analysed the influence on the acquisition parameters current time product (mA) and tube potential (kV) on YACTAs QCT parameters (17, 49). The acquisition parameters used in the phantom study were not identical with 120 kV/60 mA and 80 kV/120 mA, but somehow comparable in their variance to each other. The comparison of the two protocols showed no significant differences, for example for MLD with −923.4 ± 6.4 HU vs. -923.2 ± 6.4 HU or for WP% with 43.3 ± 6.9% vs. 43.6 ± 7.2% (17, 49). However, the exact influence of the different protocols is difficult to quantify, and possible confounders cannot be completely excluded, although the use of both protocols in each GOLD group should weaken any possible effect. Other important confounders are the variation of lung volume and cigarette smoking status (9). The variation of lung volume was reduced as best as possible by instructing and monitoring all patients for a stable full inspiratory position. Smoking status was not considered because consistent smoking history was not available for all patients. This can be considered a substantial weakness, since cigarette smoking status influences the
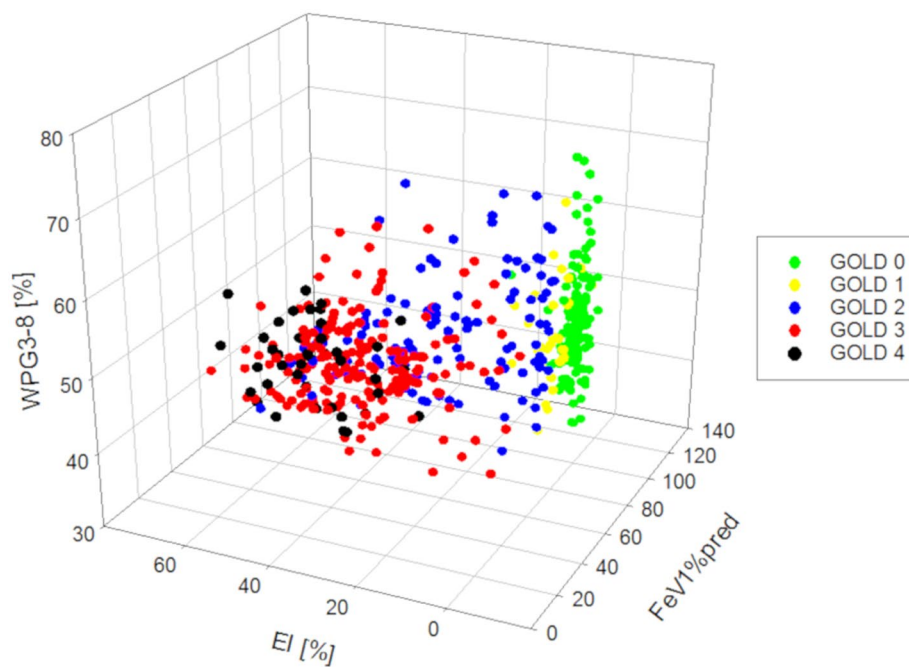
FIGURE 6
3D-residual scatterplots for the dependent variable FEV1%pred and the independent variables EI and WP$_{3-8}$.

quantification of emphysema and airway disease (4, 50). Several authors have shown that current smokers have higher lung density, presumably due to a smoking-related increase in inflammatory cells in the lungs of current smokers, whereas other studies have shown that bronchial dimensions depend on current smoking status (51–54). However, in a large study from Rotterdam of approximately 2000 COPD patients, 41% were current smokers and 57.6% were former smokers or never smokers, indicating a relatively even distribution in an unselected population. Assuming a similar distribution in our study cohort, the impact of smoking on QCT parameters should be limited within each GOLD group. Although we agree that smoking status would have increased data quality, we still consider the results robust in the context of a descriptive cross-sectional study.

This work focused on QCT for GOLD stage-specific quantification of emphysema and airway disease. QCT parameters showed significant differences for GOLD 0–4 COPD, while lung lobe analysis revealed significant differences in the changes of airway dimensions between upper and lower lobes. The observed changes in airway dimensions may indicate airway inflammation in GOLD 2, which may lead to irreversible airway destruction in GOLD 3 and GOLD 4. Emphysema and airway disease both contribute independently to lung function decline. However, COPD is a very heterogeneous disease, which means that different stages of airway disease may coexist in the same lung, emphasizing the need to assess the spatial distribution within the lung.

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 of the University of Heidelberg. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Funding

Acknowledgments

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2023.1184784/full#supplementary-material

References

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. (2006) 3:e442. doi: 10.1371/journal.pmed.0030442

2. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: Gold executive summary. *Am J Respir Crit Care Med*. (2007) 176:532–55. doi: 10.1164/rccm.200703-456SO

3. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. Gold executive summary. *Am J Respir Crit Care Med*. (2017) 195:557–82. doi: 10.1164/rccm.201701-0218PP

4. Jobst BJ, Weinheimer O, Buschulte T, Trauth M, Tremper J, Delorme S, et al. Longitudinal airway remodeling in active and past smokers in a lung Cancer screening population. *Eur Radiol*. (2019) 29:2968–80. doi: 10.1007/s00330-018-5890-4

5. Takayanagi S, Kawata N, Tada Y, Ikari J, Matsuura Y, Matsuoka S, et al. Longitudinal changes in structural abnormalities using Mdct in Copd: do the Ct measurements of Airway Wall thickness and small pulmonary vessels change in parallel with emphysematous progression? *Int J Chron Obstruct Pulmon Dis*. (2017) 12:551–60. doi: 10.2147/copd.S121405

6. Regan EA, Lynch DA, Curran-Everett D, Curtis JL, Austin JH, Grenier PA, et al. Clinical and radiologic disease in smokers with Normal spirometry. *JAMA Intern Med*. (2015) 175:1539–49. doi: 10.1001/jamainternmed.2015.2735

7. Jimborean G, Ianoși ES, Postolache P, Arghir O. The role of quantitative computed tomography in the diagnosis of chronic obstructive pulmonary disease. *Pneumologia*. (2016) 65:184–8.

8. Koo HK, Vasilescu DM, Booth S, Hsieh A, Katsamenis OL, Fishbane N, et al. Small airways disease in mild and moderate chronic obstructive pulmonary disease: a cross-sectional study. *Lancet Respir Med*. (2018) 6:591–602. doi: 10.1016/s2213-2600(18)30196-6

9. Lynch DA, Al-Qaisi MA. Quantitative computed tomography in chronic obstructive pulmonary disease. *J Thorac Imaging*. (2013) 28:284–90. doi: 10.1097/RTI.0b013e318298733c

10. Coxson HO, Leipsic J, Parraga G, Sin DD. Using pulmonary imaging to move chronic obstructive pulmonary disease beyond Fev1. *Am J Respir Crit Care Med*. (2014) 190:135–44. doi: 10.1164/rccm.201402-0256PP

11. Labaki WW, Martinez CH, Martinez FJ, Galbán CJ, Ross BD, Washko GR, et al. The role of chest computed tomography in the evaluation and Management of the Patient with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2017) 196:1372–9. doi: 10.1164/rccm.201703-0451PP

12. Lynch DA, Austin JH, Hogg JC, Grenier PA, Kauczor HU, Bankier AA, et al. Ct-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner society. *Radiology*. (2015) 277:141579:192–205. doi: 10.1148/radiol.2015141579

13. Gevenois PA, De Vuyst P, de Maertelaer V, Zanen J, Jacobovitz D, Cosio MG, et al. Comparison of computed density and microscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med*. (1996) 154:187–92. doi: 10.1164/ajrccm.154.1.8680679

14. Madani A, Zanen J, Maertelaer V, Gevenois PA. Pulmonary emphysema: objective quantification at multi–detector row Ct—comparison with macroscopic and microscopic morphometry. *Radiology*. (2006) 238:1036–43. doi: 10.1148/radiol.2382042196

15. Hackx M, Bankier AA, Gevenois PA. Chronic obstructive pulmonary disease: Ct quantification of airways disease. *Radiology*. (2012) 265:34–48. doi: 10.1148/radiol.12111270

16. Nakano Y, Wong JC, de Jong PA, Buzatu L, Nagao T, Coxson HO, et al. The prediction of small airway dimensions using computed tomography. *Am J Respir Crit Care Med*. (2005) 171:142–6. doi: 10.1164/rccm.200407-874OC

17. Leutz-Schmidt P, Wielpütz MO, Skornitzke S, Weinheimer O, Kauczor HU, Puderbach MU, et al. Influence of acquisition settings and radiation exposure on Ct lung densitometry-an anthropomorphic ex vivo phantom study. *PLoS One*. (2020) 15:e0237434. doi: 10.1371/journal.pone.0237434

18. Halpin DMG, Criner GJ, Papi A, Singh D, Anzueto A, Martinez FJ, et al. Global initiative for the diagnosis, management, and prevention of chronic obstructive lung disease. The 2020 Gold science committee report on Covid-19 and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2021) 203:24–36. doi: 10.1164/rccm.202009-3533SO

19. Cooper BG, Stocks J, Hall GL, Culver B, Steenbruggen I, Carter KW, et al. The global lung function initiative (Gli) network: bringing the World's respiratory reference values together. *Breathe*. (2017) 13:e56:–e64. doi: 10.1183/20734735.012717

20. Kauczor HU, Wielputz MO, Owsijewitsch M, Ley-Zaporozhan J. Computed tomographic imaging of the Airways in Copd and Asthma. *J Thorac Imaging*. (2011) 26:290–300. doi: 10.1097/RTI.0b013e3182277113

21. Wielpütz MO, Weinheimer O, Eichinger M, Wiebel M, Biederer J, Kauczor HU, et al. Pulmonary emphysema in cystic fibrosis detected by densitometry on chest multidetector computed tomography. *PLoS One*. (2013) 8:e73142. doi: 10.1371/journal.pone.0073142

22. Weinheimer O, Wielpütz MO, Konietzke P, Heussel CP, Kauczor HU, Brochhausen C, et al. Fully automated lobe-based airway taper index calculation in a low dose Mdct Cf study over 4 time-points. *Image Proc*. (2017) 10133:242–50. doi: 10.1117/12.2254387

23. Konietzke P, Weinheimer O, Wielputz MO, Savage D, Ziyeh T, Tu C, et al. Validation of automated lobe segmentation on paired inspiratory-expiratory chest Ct in 8-14 year-old children with cystic fibrosis. *PLoS One*. (2018) 13:e0194557. doi: 10.1371/journal.pone.0194557

24. Wielputz MO, Eichinger M, Weinheimer O, Ley S, Mall MA, Wiebel M, et al. Automatic airway analysis on multidetector computed tomography in cystic fibrosis: correlation with pulmonary function testing. *J Thorac Imaging*. (2013) 28:104–13. doi: 10.1097/RTI.0b013e3182765785

25. Weinheimer O, Achenbach T, Düber C. Fully automated extraction of airways from Ct scans based on self-adapting region growing. *Computerized Tomography*. (2009) 27:64–74.

26. Weinheimer O, Achenbach T, Bletz C, Duber C, Kauczor HU, Heussel CP. About objective 3-D analysis of airway geometry in computerized tomography. *IEEE Trans Med Imaging*. (2008) 27:64–74. doi: 10.1109/TMI.2007.902798

27. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg*. (2018) 126:1763–8. doi: 10.1213/ANE.0000000000002864

28. Nakano Y, Muro S, Sakai H, Hirai T, Chin K, Tsukino M, et al. Computed tomographic measurements of airway dimensions and emphysema in smokers. Correlation with lung function. *Am J Respir Crit Care Med*. (2000) 162:1102–8. doi: 10.1164/ajrccm.162.3.9907120

29. Knudsen L, Ochs M. The micromechanics of lung alveoli: structure and function of surfactant and tissue components. *Histochem Cell Biol*. (2018) 150:661–76. doi: 10.1007/s00418-018-1747-9

30. Konietzke P, Wielpütz MO, Wagner WL, Wuennemann F, Kauczor HU, Heussel CP, et al. Quantitative Ct detects progression in Copd patients with severe emphysema in a 3-month interval. *Eur Radiol*. (2020) 30:2502–12. doi: 10.1007/s00330-019-06577-y

31. Cavigli E, Camiciottoli G, Diciotti S, Orlandi I, Spinelli C, Meoni E, et al. Whole-lung densitometry versus visual assessment of emphysema. *Eur Radiol*. (2009) 19:1686–92. doi: 10.1007/s00330-009-1320-y

32. Stolk J, Putter H, Bakker EM, Shaker SB, Parr DG, Piitulainen E, et al. Progression parameters for emphysema: a clinical investigation. *Respir Med*. (2007) 101:1924–30. doi: 10.1016/j.rmed.2007.04.016

33. Hasegawa M, Nasuhara Y, Onodera Y, Makita H, Nagai K, Fuke S, et al. Airflow limitation and airway dimensions in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2006) 173:1309–15. doi: 10.1164/rccm.200601-037OC

34. Hogg JC, Macklem PT, Thurlbeck WM. Site and nature of airway obstruction in chronic obstructive lung disease. *N Engl J Med*. (1968) 278:1355–60. doi: 10.1056/NEJM196806202782501

35. Smith BM, Hoffman EA, Rabinowitz D, Bleecker E, Christenson S, Couper D, et al. Comparison of spatially matched airways reveals thinner airway walls in Copd. The multi-ethnic study of atherosclerosis (Mesa) Copd study and the subpopulations and intermediate outcomes in Copd study (Spiromics). *Thorax*. (2014) 69:987–96. doi: 10.1136/thoraxjnl-2014-205160

36. Cosio M, Ghezzo H, Hogg JC, Corbin R, Loveland M, Dosman J, et al. The relations between structural changes in small airways and pulmonary-function tests. *N Engl J Med*. (1978) 298:1277–81. doi: 10.1056/nejm197806082982303

37. Thurlbeck WM, Pun R, Toth J, Frazer RG. Bronchial cartilage in chronic obstructive lung disease. *Am Rev Respir Dis*. (1974) 109:73–80. doi: 10.1164/arrd.1974.109.1.73

38. Diaz AA, Estépar RSJ, Washko GR. Computed tomographic airway morphology in chronic obstructive pulmonary disease. Remodeling or innate anatomy? *Ann Am Thorac Soc*. (2016) 13:4–9. doi: 10.1513/AnnalsATS.201506-371PP

39. Takahashi M, Fukuoka J, Nitta N, Takazakura R, Nagatani Y, Murakami Y, et al. Imaging of pulmonary emphysema: a pictorial review. *Int J Chron Obstruct Pulmon Dis*. (2008) 3:193–204. doi: 10.2147/copd.s2639

40. Boueiz A, Chang Y, Cho MH, Washko GR, San José Estépar R, Bowler RP, et al. Lobar emphysema distribution is associated with 5-year radiological disease progression. *Chest*. (2018) 153:65–76. doi: 10.1016/j.chest.2017.09.022

41. Park J, Kim EK, Lee SH, Kim MA, Kim JH, Lee SM, et al. Phenotyping Copd patients with emphysema distribution using quantitative Ct measurement; more severe airway involvement in lower dominant emphysema. *Int J Chron Obstruct Pulmon Dis*. (2022) 17:2013–25. doi: 10.2147/copd.S362906

42. Mets OM, de Jong PA, van Ginneken B, Gietema HA, Lammers JW. Quantitative computed tomography in Copd: possibilities and limitations. *Lung*. (2012) 190:133–45. doi: 10.1007/s00408-011-9353-9

43. Cooper CB. The connection between chronic obstructive pulmonary disease symptoms and hyperinflation and its impact on exercise and function. *Am J Med*. (2006) 119:21–31. doi: 10.1016/j.amjmed.2006.08.004

44. Oh AS, Baraghoshi D, Lynch DA, Ash SY, Crapo JD, Humphries SM. Emphysema progression at Ct by deep learning predicts functional impairment and mortality: results from the Copdgene study. *Radiology*. (2022) 304:672–9. doi: 10.1148/radiol.213054

45. Bhatt SP, Soler X, Wang X, Murray S, Anzueto AR, Beaty TH, et al. Association between functional small airway disease and Fev1 decline in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2016) 194:178–84. doi: 10.1164/rccm.201511-2219OC

46. Kim SS, Jin GY, Li YZ, Lee JE, Shin HS. Ct quantification of lungs and airways in normal Korean subjects. *Korean J Radiol*. (2017) 18:739–48. doi: 10.3348/kjr.2017.18.4.739

47. Copley SJ, Giannarou S, Schmid VJ, Hansell DM, Wells AU, Yang GZ. Effect of aging on lung structure in vivo: assessment with Densitometric and fractal analysis of high-resolution computed tomography data. *J Thorac Imaging*. (2012) 27:366–71. doi: 10.1097/RTI.0b013e31825148c9

48. Barjaktarevic I, Springmeyer S, Gonzalez X, Sirokman W, Coxson HO, Cooper CB. Diffusing capacity for carbon monoxide correlates best with tissue volume from quantitative Ct scanning analysis. *Chest*. (2015) 147:1485–93. doi: 10.1378/chest.14-1693

49. Leutz-Schmidt P, Weinheimer O, Jobst BJ, Dinkel J, Biederer J, Kauczor HU, et al. Influence of exposure parameters and iterative reconstruction on automatic airway segmentation and analysis on Mdct-an ex vivo phantom study. *PLoS One*. (2017) 12:e0182268. doi: 10.1371/journal.pone.0182268

50. Jobst BJ, Triphan SM, Sedlaczek O, Anjorin A, Kauczor HU, Biederer J, et al. Functional lung Mri in chronic obstructive pulmonary disease: comparison of T1 mapping, oxygen-enhanced T1 mapping and dynamic contrast enhanced perfusion. *PLoS One*. (2015) 10:e0121520. doi: 10.1371/journal.pone.0121520

51. Camiciottoli G, Cavigli E, Grassi L, Diciotti S, Orlandi I, Zappa M, et al. Prevalence and correlates of pulmonary emphysema in smokers and former smokers. A Densitometric study of participants in the Italung trial. *Eur Radiol*. (2009) 19:58–66. doi: 10.1007/s00330-008-1131-6

52. Grydeland TB, Dirksen A, Coxson HO, Pillai SG, Sharma S, Eide GE, et al. Quantitative computed tomography: emphysema and Airway Wall thickness by sex, age and smoking. *Eur Respir J*. (2009) 34:858–65. doi: 10.1183/09031936.00167908

53. Ashraf H, Lo P, Shaker SB, de Bruijne M, Dirksen A, Tønnesen P, et al. Short-term effect of changes in smoking behaviour on emphysema quantification by Ct. *Thorax*. (2011) 66:55–60. doi: 10.1136/thx.2009.132688

54. Jobst BJ, Weinheimer O, Trauth M, Becker N, Motsch E, Groß ML, et al. Effect of smoking cessation on quantitative computed tomography in smokers at risk in a lung Cancer screening population. *Eur Radiol*. (2018) 28:807–15. doi: 10.1007/s00330-017-5030-6

Glossary

ANOVA	One-way analysis of variance
COPD	Chronic obstructive pulmonary disease
CT	omputed tomography
EI	mphysema index
FEV1	Forced Expiratory Pressure in 1 Second
FEV1%pre	FEV1 predicted
GOLD	Global Initiative for Obstructive Lung Disease
HU	ounsfield units
LA	Lumen area
LLi	Lingula
LLL	Left lower lobe
LUL	Left upper lobe
RML	Middle lobe
RLL	Right lower lobe
RUL	Right upper lobe
MLD	Mean lung density
PFT	Pulmonary function test
QCT	Quantitative computed tomography
RV	Residual volume
SAD	Small airway disease
TD	Total diameter
TLC	Total lung capacity
TLV	Total lung volume
VC	Vital capacity
WP	Wall percentage
WT	Wall thickness

Check for updates

Quantitative assessment of airway wall thickness in COPD patients with interstitial lung abnormalities

Yingying Ji[1,2], Leqing Chen[1,2], Jinrong Yang[1,2], Xiangying Yang[1,2] and Fan Yang[1,2]*

[1]Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China, [2]Hubei Province Key Laboratory of Molecular Imaging, Wuhan, China

Background: Whether the airway is involved in the pathogenesis of interstitial lung abnormalities (ILA) is not well understood. Also the impact of ILA on lung function in COPD patients remains controversial. We aimed to assess the quantitative CT measurements of airway wall thickness (AWT) and lung function according to ILA status in COPD patients.

Methods: 157 COPD patients discharged from our hospital from August 1, 2019 through August 31, 2022 who underwent chest CT imagings and pulmonary function tests were retrospectively enrolled. Linear regression analysis and multiple models were used to analyze associations between quantitative assessment of airway wall changes and the presence of ILA.

Results: In 157 COPD patients, 23 patients (14.6%) had equivocal ILA, 42 patients (26.8%) had definite ILA. The definite ILA group had the highest measurements of Pi10 (square root of theoretical airway wall area with a lumen perimeter of 10mm), segmental AWT and segmental WA% (percentage of wall area), whereas the no ILA group had the lowest measurements of Pi10, segmental AWT and segmental WA%. In the adjusted analyses (adjusted by age, sex, body mass index, smoking intensity, COPD GOLD stage, lung function, slice thickness and scanner type), compared to COPD patients without ILA, the measurements of Pi10, segmental AWT and segmental WA% were higher in definite ILA group with differences of 0.225mm (p=0.012), 0.152mm (p<0.001), 4.8% (p<0.001) respectively. COPD patients with definite ILA tended to have higher FEV1% predicted, FVC% predicted and lower MMEF$_{75/25}$% predicted, but there were no statistically differences among the three groups.

Conclusion: Our study demonstrates the higher AWT measures in COPD patients with ILA compared to the patients without ILA. These findings suggest that the airway may be involved in the pathogenesis of ILA.

KEYWORDS

interstitial lung abnormalities, idiopathic pulmonary fibrosis, COPD, quantitative CT, lung function

1 Introduction

Interstitial lung abnormalities (ILA) are asymptomatic radiological abnormalities detected by chest CT, shown as high-attenuation areas in the lung field. ILA are usually associated with interstitial lung disease (ILD) and idiopathic pulmonary fibrosis (IPF), and can be considered as the subclinical phase of ILD (1, 2). It is becoming increasingly clear that certain forms of

ILD progress from an asymptomatic or "subclinical" stage to a final clinical diagnosis. IPF is a progressive, destructive disease with unknown causes and few treatment options, and its incidence has gradually increased in recent years (3). The molecular mechanisms of ILD and IPF have garnered increased attention due to the historical lack of understanding of their etiology and natural history, as well as the absence of effective treatments (4). Previous studies have shown that the undiagnosed patients with ILA share several similar characteristics with clinically significant ILD patients, including reduced lung volume, limited function, increased lung symptoms, histopathological changes, and molecular characteristics, as well as similar but milder syndromes and genetic and genomic similarities with IPF patients, and ILA may have a common pathogenesis with ILD and IPF (5). Therefore, a full understanding of ILA can improve the understanding of the natural process of ILD and IPF, and make it possible for the early management and timely intervention of these diseases.

Several studies have shown that IPF may be caused by abnormal behavior of alveolar epithelial cells combined with excessive fibroblast activation (3, 6). It has been gradually observed that bronchiolar epithelial cells and broncho-alveolar junctions are related targets of cell damage in IPF, thus the correlation between airway epithelial biology and IPF has been gradually realized (4). Therefore, this study hypothesized that the airway might be involved in the pathogenesis of ILA and tested this hypothesis by measuring airway wall thickness (AWT) in ILA patients using quantitative CT.

To our knowledge, studies using quantitative CT measurements of AWT have focused on COPD, while few studies on ILA (7, 8). And because ILA are relatively common in COPD patients, we also want to investigate whether the prevalence of ILA exacerbates the lung function in COPD patients, as some studies have demonstrated (2, 9). Thus, we hypothesized that the COPD patients with equivocal ILA and definite ILA had higher AWT and lower lung function compared to those without ILA. We assessed the quantitative CT measurements of AWT and lung function in COPD patients to evaluate the hypothesis.

2 Materials and methods

2.1 Study populations

The Medical Ethics Committee of our hospital approved this study (NO.0271–01), and formal consent was not required as this was a retrospective study. Patients with COPD discharged from our hospital from August 1, 2019 through August 31, 2022 were retrospectively enrolled. The inclusion criteria for this study were as follows: (1) COPD patients diagnosed with spirometric criteria (post-bronchodilator FEV1/FVC < 0.7), (2) patients who had completed chest CT scans and pulmonary functional tests (PFTs) within one week. The exclusion criteria included: (1) other chronic respiratory diseases such as asthma, pulmonary tuberculosis, or lung cancer, (2) a history of pulmonary surgery or thoracic deformity, (3) infectious diseases affecting radiological changes in the lung parenchyma. A total of 157 patients were ultimately enrolled. The detailed flow chart was shown as Figure 1.

2.2 CT scanning

All participants underwent imagings with IQon Spectral CT (Philips Healthcare, Best, The Netherlands) or SOMATOM Definition AS+ (Siemens Healthineers. Forchheim, Germany) at full inspiration in the supine position. The instrument settings were as follows: detector collimation width of 64×0.6 mm or 128×0.6 mm; tube voltage of 120 kV; tube current of 40–80 mA, the tube current was regulated by automatic exposure control system (CARE Dose 4D; Siemens Healthineers) or (DoseRight, Philips Healthcare); slice thicknesses of 1.25 mm or 2.00 mm and internals of 1.25 mm or 2.00 mm, slice thicknesses of 2.00 mm accounts for 7/92 (7.6%) in the no ILA group, 2/23 (8.7%) in the equivocal ILA group, 3/42 (7.1%) in the definite ILA group; All images were reconstructed with lung and soft tissue kernels for quantitative analysis, reconstruction kernel of B70f kernel and a mediastinal B30f kernel.

2.3 Pulmonary functional tests

All PFTs were performed using Master Lab equipment (CareFusion, Hoechberg, Germany), Spirometry for FEV1 and FVC was performed before and 15 min after inhalation of 200 mg salbutamol following the American Thoracic Society Committee guidelines (10). The following values were recorded: FEV1/FVC, FEV1, FVC and maximum mid-expiratory flow (MMEF75/25). All values were expressed as a percentage of the predicted value. All PFTs and CT scans were completed within one week.

2.4 Imaging review

Three radiologists with 5, 7, and 15 years of experience visually assessed chest CT scans for ILA, blinded to any patient-specific information. Findings definite for ILA were defined as nondependent changes that affected more than 5% of any lung zone, including nondependent ground-glass opacity (GGO) or reticular abnormality, diffuse centrilobular nodularity, honeycombing, traction bronchiectasis, nonemphysematous cysts, or architectural distortion. Focal or unilateral ground-glass attenuation, focal or unilateral reticulation, and patchy ground-glass abnormalities that affected less than 5% of the lung were considered equivocal findings (2). Discrepancies in CT diagnosis or ILA characterization were resolved by consensus.

In addition, the extent of ILA were graded for every lobe using the following five-point scale: score of 0, absent; score of 1, 1 to 25% involvement; score of 2, 26 to 50% involvement; score of 3, 51 to 75% involvement; and score of 4, 76 to 100% involvement. Total ILA score ranges from 0 to 24 with the lingular segment counted separately (2).

2.5 Quantitative CT measurements

Quantitative CT measurements were evaluated using the COPD Analysis package (Philips IntelliSpace Portal). The extent of emphysema was quantified by evaluating the areas of the lung with attenuation less than −950 Hounsfield units (HU) at full inspiration,
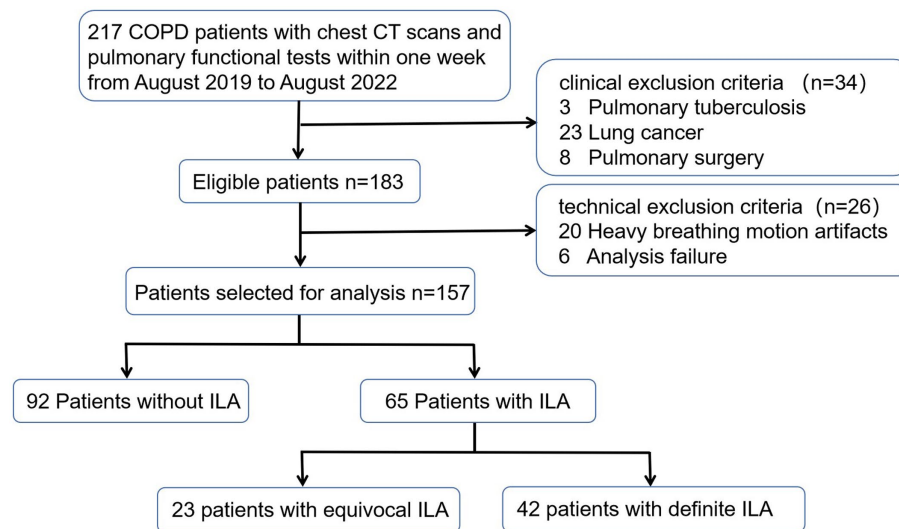
FIGURE 1
Flow diagram of the study population. ILA, interestitial lung abnormality.
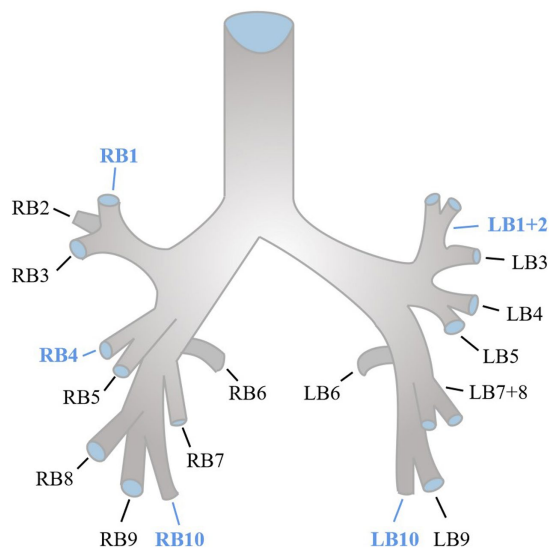
FIGURE 2
Anatomy of segmental bronchi of normal segmented airway tree.

commonly referred to as LAA-950 (10). We used an integral-based method for airway measurements (11). To calculate the segmental airway measurements, the average of five segmental airways (RB1, RB4, RB10, LB1+2, and LB10) along the middle third of each airway was taken (10, 12), the anatomy of normal segmental bronchi was shown as Figure 2. The segmental wall area percentage (WA%) was determined by taking the average of the five segmental airways and calculating (WA/total cross-sectional area) × 100. Pi10, which was assessed from the regression line calculated by plotting the square root of the wall area and the lumen perimeter of the airway, was defined as the square root of the wall area of a theoretical airway with a lumen perimeter of 10 mm (13).

2.6 Statistical analysis

The study reported continuous variables as mean ± SD and categorical data as percentages. Differences in continuous variables between groups were tested using Kruskal-Wallis or ANOVA tests depending on the normality of distribution, and the LSD-t method was used for post-hoc comparisons. Chi-square tests were used to compare differences in categorical data between groups. Linear regression analysis was conducted to evaluate the associations between the quantitative assessment of airway changes (Pi10, segmental AWT, segmental WA%) and the presence of ILA. Multivariate models were used to adjust for potential confounding factors such as age, sex, BMI, smoking intensity, classification of Chronic Obstructive Pulmonary Disease (GOLD), lung function (FEV1/FVC, FEV1 percent predicted, FVC percent predicted, MMEF75/25 percent predicted), slice thickness and scanner type. Logistic regression analysis was conducted to evaluate the associations between age and the presence of ILA, potential confounding factors including sex, BMI, smoking intensity, COPD GOLD Stage were used for multivariate models. Statistical analysis was performed using SPSS software (version 27.0), and the significance level was set at $p < 0.05$ (two-tailed).

3 Results

3.1 Baseline characteristics

After assessing 217 COPD patients, 157 eligible patients completed both chest CT scanning and PTFs within one week and were included in the study (Figure 1). Table 1 presents the baseline characteristics of the study population. Of these COPD patients, 92 (58.6%) patients showed no evidence of ILA, 23 (14.6%) patients showed equivocal ILA, and 42 (26.8%) patients had definite ILA. Most of study populations were men (90%) with the mean age of 63.2 years. Similar to previous studies (2, 7), the COPD patients with ILA tended

TABLE 1 Comparison of baseline demographics, lung function, quantitative CT measurements by ILA.

	No ILA ($n = 92$)	Equivocal ILA ($n = 23$)	Definite ILA ($n = 42$)	P0	P1	P2	P3
Age (year)	61.5 ± 9.0	64.0 ± 7.1	66.5 ± 8.8	0.008	0.610	0.837	0.007
Male sex (%)	85 (92.4)	21 (91.3)	35 (83.3)	0.266	0.572	0.312	0.102
BMI (kg/m^2)	23.0 ± 3.2	23.5 ± 3.4	23.3 ± 3.0	0.721	1.000	1.000	1.000
Smoking intensity (pack-y)	16.1 ± 21.6	15.0 ± 23.5	15.3 ± 19.1	0.962	1.000	1.000	1.000
Baseline lung function	…	…	…	…	…	…	…
FEV1/FVC	61.2 ± 10.4	62.8 ± 8.6	61.9 ± 8.6	0.764	1.000	1.000	1.000
FEV1 (%$_{pred}$)	84.1 ± 21.9	86.0 ± 18.2	86.0 ± 23.8	0.877	1.000	1.000	1.000
FVC (%$_{pred}$)	107.7 ± 19.1	107.6 ± 15.3	108.9 ± 20.6	0.940	1.000	1.000	1.000
MMEF$_{75/25}$ (%$_{pred}$)	35.0 ± 15.3	33.9 ± 14.5	33.8 ± 12.7	0.903	1.000	1.000	1.000
Imaging	…	…	…	…	…	…	…
LAA$_{-950}$ (%)	…	…	…	…	…	…	…
Whole lung (%)	5.94 ± 7.13	3.60 ± 4.95	4.85 ± 6.01	0.276	0.389	1.000	1.000
Upper lobes (%)	7.91 ± 10.43	3.28 ± 3.98	7.06 ± 9.59	0.118	0.003	0.087	0.954
Lower lobes (%)	4.88 ± 6.3	3.73 ± 6.21	3.46 ± 4.4	0.371	0.397	0.862	0.194
Pi10(mm)	5.15 ± 0.37	5.27 ± 0.57	5.39 ± 0.55	0.014	0.714	0.768	0.030
Segmental AWT (mm)	1.20 ± 0.16	1.31 ± 0.20	1.36 ± 0.20	<0.001	0.031	0.778	<0.001
Segmental LD (mm)	5.04 ± 0.71	5.05 ± 0.77	4.86 ± 0.89	0.419	1.000	0.988	0.633
Segmental LA (mm^2)	21.21 ± 5.91	21.39 ± 6.18	20.03 ± 7.21	0.563	0.903	0.408	0.317
Segmental WA%	54.37 ± 6.18	56.62 ± 6.38	59.19 ± 7.09	<0.001	0.411	0.382	<0.001
TLCCT (L)	5.50 ± 1.08	4.90 ± 0.96	4.64 ± 1.23	<0.001	0.061	1.000	<0.001

Data are presented as No. (%) or mean ± SD unless otherwise indicated. BMI, Body mass index; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; MMEF$_{75/25}$, maximum mid-expiratory flow between 25–75%; LAA$_{-950}$, Percentage area with CT attenuation values less than −950 HU at inspiration; Pi10, square root of the wall area of a theoretical airway with internal perimeter of 10 mm; AWT, airway wall thickness; LD, lumen diameter; LA, lumen area; WA%, wall area percent; TLCCT, total lung capacity on CT. Segmental AWT, segmental LD and segmental WA% were calculated as the average of five segmental airway (RB1, RB4, RB10, LB1+2,and LB10). Lingula lobe was excluded from the calculation of the upper lobes. P0 is the p value of the three groups. P1 is the p value of no ILA group and the equivocal ILA group. P2 is the p value of the equivocal ILA group and definite ILA group. P3 is the p value of no ILA group and the definite ILA group.

to be older, and increased age (OR adjusted = 1.069, 95%CI = 1.024–1.117, $p = 0.003$) was associated with definite ILA (Supplementary Table S1). For lung function, higher FEV1% predicted, FVC% predicted and lower MMEF$_{75/25}$% predicted were noticed in the definite ILA patients, but all of the included lung function and the smoking intensity showed no statistically differences between these groups.

Representative imaging findings of COPD patients without ILA and with equivocal/definite ILA are shown in Figure 3. Table 2 presents radiologic findings of chest CT scans in patients with equivocal and definite ILA. The most common radiologic findings of equivocal and definite ILA patients were focal or unilateral reticulation (73.9%) and nondependent ground-glass opacity (59.5%), respectively. The ILA scores of definite ILA tended to be higher than the patients with equivocal ILA (1.22 ± 0.52 vs. 2.93 ± 1.72).

3.2 Quantitative CT measurements

Quantitative CT measurements are presented in Table 1. Consistent with previous studies (14), COPD patients with ILA had lower measurements of LAA-950, but no significant differences were observed between the three groups. In the spatial distribution of emphysema, similar to previous study (15), in COPD patients without

ILA, the upper lobes had a higher proportion of LAA-950 than the lower lobes, such phenomenon can also be observed in COPD patients with definite ILA, and the difference was more obvious in the latter group; The proportion of LAA-950 in the upper lobes and lower lobes was similar in the those with equivocal ILA. The measurements of Pi10, segmental AWT, and segmental WA% exhibited highest in the definite ILA group, whereas lowest in the no ILA group. In Pi10 and segmental WA% measurements, significant differences were only observed between COPD patients with no ILA and those with definite ILA ($p = 0.030$ and $p < 0.001$, respectively). COPD patients with no ILA had higher measurements of TLCCT than those with ILA ($p < 0.001$). Meanwhile, the measurements of segmental AWT were significantly different between COPD patients with no ILA and those with equivocal ILA ($p = 0.031$) and between COPD patients with no ILA and those with definite ILA ($p < 0.001$). The segmental LD and segmental LA results showed no differences among the three groups.

3.3 AWT measurements according to ILA status

Quantitative CT measurements of AWT (including Pi10, segmental AWT, segmental WA%) according to ILA status were analyzed (Table 3). The measures of Pi10 and segmental WA% in
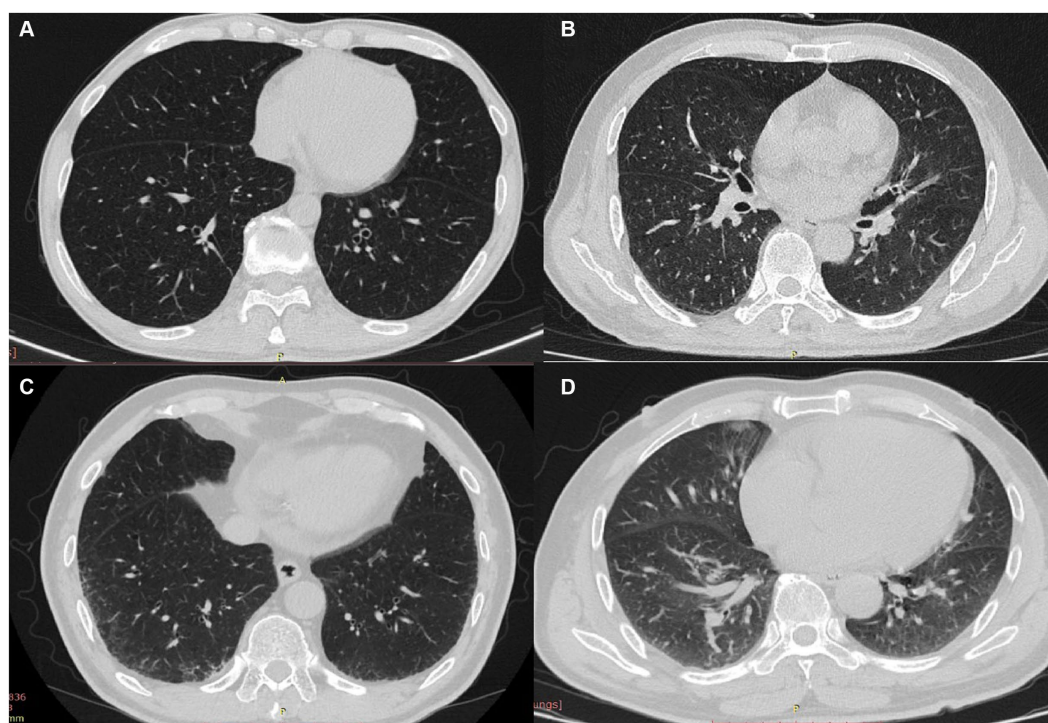
FIGURE 3
(A–D) CT findings of COPD patients without ILA and with equivocal/definite ILA; patients without ILA **(A)** and characteristic images of equivocal/definite ILA **(B–D)**, including unilateral reticulation **(B)**, bilateral reticular opacities **(C)** and nondependent group-glass opacity **(D)**.

TABLE 2 Radiologic findings in patients with equivocal and definite ILA.

Radiologic finding	Equivocal ILA ($n = 23$)	Definite ILA ($n = 42$)
Equivocal ILA		
Focal or unilateral GGO	5 (21.7)	1 (2.4)
Focal or unilateral reticulation	17 (73.9)	1 (2.4)
Patchy GGO	1 (4.3)	/
Definite ILA		
Nondependent ground-glass opacity	/	25 (59.5)
Nondependent reticular abnormality	/	16 (38.1)
Diffuse centrilobular nodularity	/	2 (4.8)
Traction bronchiectasis	/	1 (2.4)
Honeycombing	/	2 (4.8)
Architectural distortion	/	1 (2.4)
ILA score	1.22 ± 0.52	2.93 ± 1.72

Data are presented as No. (%) or mean ± SD. GGO, ground-glass opacity.

COPD patients with definite ILA were significantly higher than those with no ILA. In addition, the segmental AWT measures in COPD patients with equivocal and definite ILA were significantly higher compared with those with no ILA. These results were shown in both

univariate analyses and multivariate analyses adjusted by age, sex, BMI, smoking intensity, classification of COPD GOLD, lung function, slice thickness and scanner type. In the adjusted analyses, compared to COPD patients with no ILA, the measurements of Pi10, segmental AWT and segmental WA% were higher in definite ILA group with differences of 0.225 mm ($p = 0.012$), 0.152 mm ($p < 0.001$), 4.8% ($p < 0.001$) respectively.

Furthermore, the findings of this association could be due to the ILA lesions only changing the adjacent airways, we excluded this possibility by selecting patients with ILA whose lesions were confined to the lower lobes of both lungs, and quantitative CT measurements of these patients (including AWT of RB1, RB4, LB1+2, the average of AWT of RB1, RB4, LB1+2, the average WA% of RB1, RB4, LB1+2) compared to patients with no ILA were analyzed (Table 4; Figure 4), We did not include Pi10 as the airway measure because it is a composite indicator. Similar higher measurements of AWT were noted in COPD patients whose ILA lesions were confined to the lower lobes, the results were observed in both univariate and multivariate analyses, which were adjusted for important covariates. The measures of AWT of RB1, the average AWT and WA% of these three segmental airways in COPD patients with ILA were significantly higher than those with no ILA. In addition, similar higher AWT measurements of RB4 and LB1+2 were also noted in COPD patients with ILA compared with those with no ILA, but no significant differences were observed between the two groups. In the adjusted analyses, compared to COPD patients without ILA, the average AWT and WA% measures of these three segmental airways were 0.189 mm ($p = 0.001$) and 3.897% ($p = 0.013$) higher in those with ILA, and higher AWT measures of RB1 with the difference of 0.287 mm ($p < 0.001$) in those with ILA.

TABLE 3 AWT measurements of equivocal and definite ILA patients compared to patients without ILA.

Variables	Unadjusted analysis		Adjusted analysis	
	β ± SE	P value	β ± SE	p value
Pi10 (mm)				
Equivocal ILA	0.121 ± 0.048	0.257	0.05 ± 0.109	0.651
Definite ILA	0.249 ± 0.085	0.004	0.225 ± 0.088	0.012
Segmental AWT (mm)				
Equivocal ILA	0.107 ± 0.041	0.010	0.087 ± 0.043	0.042
Definite ILA	0.159 ± 0.033	<0.001	0.152 ± 0.034	<0.001
Segmental WA%				
Equivocal ILA	2.3 ± 1.5	0.137	2.0 ± 1.5	0.186
Definite ILA	4.8 ± 1.2	<0.001	4.8 ± 1.2	<0.001

Data are presented as No. (%) or mean ± SD. See Table 1 legend for expansion of abbreviations. Adjusted analyses are adjusted for age, sex, BMI, smoking intensity, COPD GOLD Stage and lung function (FEV1/FVC, FEV1 percent predicted, FVC percent predicted, MMEF75/25 percent predicted, slice thickness and scanner type).

TABLE 4 AWT measurements between patients whose ILA lesions were confined to the lower lobes of both lungs and patients without ILA.

Variables	Unadjusted analysis		Adjusted analysis	
	β ± SE	P value	β ± SE	P value
Segmental AWT*(mm)	0.152 ± 0.048	0.002	0.189 ± 0.053	0.001
Segmental WA%*	2.776 ± 1.333	0.040	3.897 ± 1.484	0.010
RB1 AWT (mm)	0.242 ± 0.064	<0.001	0.287 ± 0.068	<0.001
RB4 AWT (mm)	0.126 ± 0.084	0.138	0.178 ± 0.096	0.066
LB1+2 AWT (mm)	0.089 ± 0.057	0.120	0.101 ± 0.065	0.124

Data are presented as No. (%) or mean ± SD. See Table 1 legend for expansion of abbreviations. *Segmental AWT and segmental WA% were calculated as the average of three segmental airway (RB1, RB4, and LB1+2). Adjusted analyses are adjusted for age, sex, BMI, smoking intensity, COPD GOLD Stage and lung function (FEV1/FVC, FEV1 percent predicted, FVC percent predicted, MMEF75/25 percent predicted, slice thickness and scanner type).

4 Discussion

In this study, we demonstrate the quantitative measurements of AWT (including Pi10, segmental AWT, segmental WA%) are higher in COPD patients with equivocal ILA and definite ILA. And even in cases where ILA lesions were confined to the lower lobes of both lungs, thicker AWT were observed in the upper lobes of both lungs and in the right middle lobe of the lungs in COPD patients with ILA than those without ILA. But all of the lung function (including FEV1/FVC, FEV1% predicted, FVC% predicted, MMEF$_{75/25}$% predicted) we enrolled showed no differences between these groups.

As far as we know, there have been limited studies employing quantitative CT to investigate AWT in COPD patients with ILA. Similar to our findings, Miller et al. (7) observed the higher Pi10 measurements in ILA and IPF patients in three separate centers (COPDGene, ECLIPSE, Framingham Heart Study). However, in a

small sample study by Li et al. (8) (54 ILA patients vs. 18 healthy non-smokers), no difference in Pi10 measurements was observed between the two groups. In addition, as Pi10 and segmental WA% are comprehensive measurement index, their increases may be due to the increase of AWT or factors that may cause airway stenosis (such as mucus in airway cavity) (7). Therefore, in our study, the actual AWT, LD and LA of the five segments of airways were also measured. Even though the technical limitations made it impossible to measure the diameter of the airway less than 2 mm, numerous studies have demonstrated that larger airway wall thickness measurements can serve as a reliable surrogate for smaller airway measurements, and that the inflammatory response observed in smaller airways can also be detected in larger airways (16–18). Our study observed higher Pi10 and actual AWT measures in patients with ILA compared to those without ILA, while no significant differences in segmental LD and LA measures were observed between patients with equivocal/definite ILA and those without ILA, this ruled out the possibility those factors which result in narrowing of the airway lumen had effect on segmental WA% measures, nevertheless, the effect on Pi10 measurements cannot be excluded because the calculation of Pi10 also includes other airways besides the segmental airways. These findings suggest that airway wall thickening occurs in ILA patients, and given the genetic and genomic similarities between ILA and IPF, airway involvement may play a role in the pathogenesis of these diseases.

The data on the prevalence of ILA in asymptomatic smokers are increasing (mainly due to ongoing CT-based trials for lung cancer screening and other diagnostic purposes) (19, 20), in our study, the incidence of ILA was 41.4%, and definite ILA and equivocal ILA accounted for 26.8 and 14.6% respectively, which is higher than the proportion of 13.5–22.2% reported among smoking cohorts (21, 22), but similar to the prevalence of 48% in Framingham Heart studies and 43% in COPD Gene (23). As previous study have shown, increasing age and tobacco smoke exposure are clear risk factors for ILA (19, 24), we did found the COPD patients with definite ILA tended to be older, with each 1-year increase in age, the risk of definite ILA increased by 6.9%. Because of the close relationship between ILA and aging, the reticular abnormalities are sometimes considered to be part of the normal spectrum of senescent lung (25). However, based on the present studies, ILA is a risk factor for acute exacerbation and increased risk of mortality independent of age, highlighting its clinical importance (2, 26). Inconsistent with previous studies, the smoking intensity showed no difference among these groups in our study.

While the clinical significance of ILA is becoming more recognized, the etiology and natural history of ILA, ILD, and IPF remain unclear resulting in a lack of effective therapeutic techniques. Some studies have proposed the correlation between airway epithelial biology and IPF (4). Since ILA may have common pathogenesis with ILD and IPF (6), this study hypothesized that the airway may be involved in the pathogenesis of ILA, and evaluated the AWT of ILA patients from the perspective of imaging to verify this hypothesis. This provides additional support for the possible involvement of the airway in the pathogenesis of ILA and may improve understanding of the natural course of ILD and IPF. Considering previous studies have showed ILA was a risk factor for acute exacerbation and ultimately may affect COPD mortality, moreover, ILA was proved to be associated with an increased rate of dying from pulmonary fibrosis (2, 9, 23), ILA may represent a precursor of pulmonary fibrosis development in COPD patients.
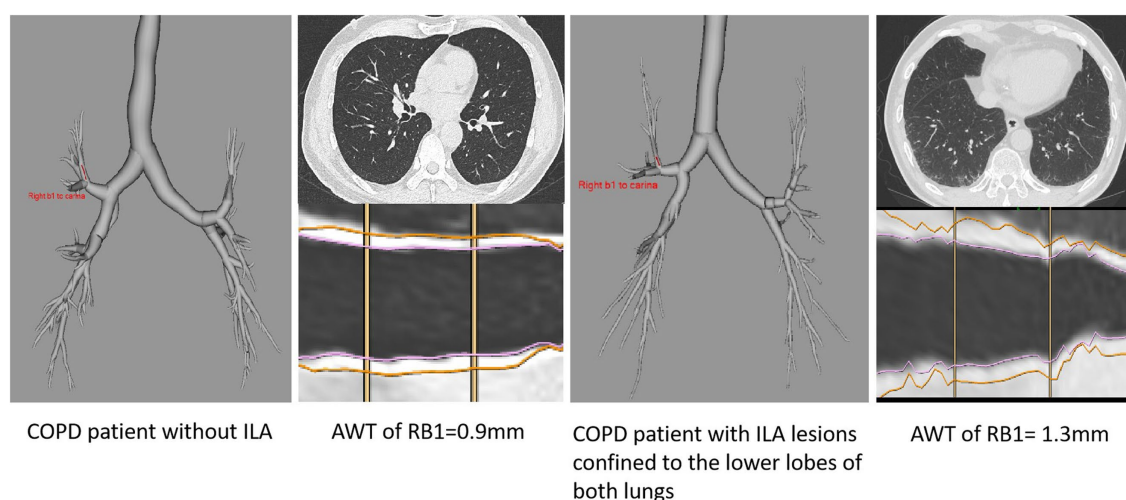
FIGURE 4
Entire airway tree slice of RB1 bronchi of COPD patients with and without ILA.

Studies utilizing quantitative CT to evaluate the correlation between AWT and clinical course have primarily focused on COPD populations. Gietema et al. (16) reported that an increase in AWT was linked to a decrease in quality of life in patients with COPD, while Johannessen et al. (18) observed that an increase in AWT was linked to respiratory mortality in patients with severe emphysema. In addition, several studies have demonstrated that thicker AWT in COPD patients is associated with poorer lung function, higher risks of exacerbation of acute COPD, and higher scores on the St. George breathing questionnaire and body index (BMI, airflow obstruction, dyspnea, and exercise capacity) (27–29). All these studies provide evidence that airway wall thickening was associated with worse clinical outcomes.

Because ILA and COPD share similar risk factors (such as smoking, male gender, and old age), the two diseases often co-exist in patients, previous studies have reported a higher incidence of ILA in COPD patients than in smokers (22). However, few studies have investigated the relationship between ILA and COPD (19, 30). COPD is a heterogeneous disease, emphysema and airway disease can contribute independently to lung function decline (15). Considering the AWT measures increased in the ILA group, we speculated that COPD patients with ILA would exhibit poorer lung function than COPD patients without ILA. However, the statistical results were not consistent with our hypothesis, and all of the included lung function (FEV1/FVC, FEV1% predicted, FVC% predicted, MMEF75/25% predicted) were not different between the patients with equivocal/definite ILA and the patients without ILA. This result may be related to the fact that the majority of our patients with COPD were GOLD I stage and GOLD II stage (144/157, 91.7%). The relationship between ILA and spirometry results has also remained controversial in several studies evaluating the relationship between ILA and COPD (20). In some cohorts, the additional presence of ILA features was associated with a higher predicted percentage of FEV1 and a lower percent predicted diffusing capacity of the lung for carbon monoxide (DLCO) (9, 22, 31). similar to our study, some cohort studies showed there was no significant difference in FEV1 and FVC between patients with and without ILA (32), and in ECLIPSE cohort, significant higher FEV1

and FEV1/FVC were observed in COPD patients with ILA compared to no ILA (23). Interestingly, a recent study (33) based on smokers revealed the combination of ILA and emphysema may lead to a relative preservation of FVC, FEV1 and FEV1/FVC, but not in DLCO. Their findings demonstrated ILA can result in a relative preservation of measures of spirometry in patients with emphysema, that may lead clinicians to false assumptions about disease severity when spirometry is used alone, this study might further help us explain why the lung function (including FEV1% predicted, FVC% predicted) exhibited slightly higher in the COPD patients with definite ILA than those without ILA. Besides, note that restrictive and obstructive pulmonary diseases can have opposite effects on lung volume, which makes individual spirometry less sensitive in detecting clinically relevant changes, highlighting the role of CT in assessing this disease (5, 9), also Kahnert et al. (34) suggested Pi10 can be seen as an imaging biomarker for the course of COPD and response to treatment. In addition to lung function studies, previous studies have reported that ILA patients with COPD are associated with a higher risk of all-cause mortality, more respiratory symptoms, more severe clinical disease severity, reduced exercise capacity, and higher mortality (9, 20, 23, 35, 36).

There are several limitations to our study. Firstly, the sample size was small, and the number of ILA patients was relatively low. Additionally, it was a retrospective study conducted in a single center, which means that the evidence level was moderate. To address these limitations, future studies should be prospective, multi-center, and have a larger sample size. Second, because CT imaging was only performed in supine position, the high subpleural attenuation area in the lower lobe of both lungs caused by some pulmonary blood deposition effect was also classified into ILA population, which resulted in a slightly higher incidence of ILA patients in this study. Third, the generalization of the results is limited. The patients' inspiratory status, different scanners, slice thickness and reconstruction kernel may affect the measurement results. In order to reduce these interference factors, we tried to guide the patients to maintain a stable full inspiratory position and made different scanners and slice thickness were distributed equally

between groups. Fourth, the available lung function indicators in this study are not perfect, and lack of indicators that can reflect patients' restrictive ventilation disorder and lung diffusion function (TLC% predicted, DLCO% predicted). Fifth, the proportion of males in the study population is quite high, and due to the physiological differences in bronchial wall thickness between males and females, this may limit the generalization of the study results. Sixth, in our study, the measurements of AWT and lung function were measured at only one point in time, therefore, future longitudinal assessments of AWT would be needed.

5 Conclusion

In conclusion, our study shows higher AWT measures in COPD patients with equivocal and definite ILA compared to patients without ILA. These findings suggest that the airway may be involved in the pathogenesis of ILA from an imaging perspective. However, our study did not observe a worse pulmonary function in COPD patients due to the prevalence of ILA.

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.

Ethics statement

The studies involving humans were approved by the Ethics Commission of Wuhan Union Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/ next of kin in accordance with the national legislation and institutional requirements.

Author contributions

YJ: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. LC: Conceptualization, Data curation, Methodology, Writing – review & editing. JY: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. XY: Data curation, Formal analysis, Writing – review & editing. FY: Supervision, Validation, Visualization, Writing – review & editing.

Funding

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2023.1280651/full#supplementary-material

References

1. Araki T, Putman RK, Hatabu H, Gao W, Dupuis J, Latourelle JC, et al. Development and progression of interstitial lung abnormalities in the Framingham heart study. *Am J Respir Crit Care Med*. (2016) 194:1514–22. doi: 10.1164/rccm.201512-2523OC

2. Lee TS, Jin KN, Lee HW, Yoon SY, Park TY, Heo EY, et al. Interstitial lung abnormalities and the clinical course in patients with COPD. *Chest*. (2021) 159:128–37. doi: 10.1016/j.chest.2020.08.017

3. King TE Jr, Pardo A, Selman M. Idiopathic pulmonary fibrosis. *Lancet*. (2011) 378:1949–61. doi: 10.1016/S0140-6736(11)60052-4

4. Chilosi M, Poletti V, Zamò A, Lestani M, Montagna L, Piccoli P, et al. Aberrant Wnt/beta-catenin pathway activation in idiopathic pulmonary fibrosis. *Am J Pathol*. (2003) 162:1495–502. doi: 10.1016/S0002-9440(10)64282-4

5. Doyle TJ, Hunninghake GM, Rosas IO. Subclinical interstitial lung disease: why you should care. *Am J Respir Crit Care Med*. (2012) 185:1147–53. doi: 10.1164/rccm.201108-1420PP

6. Ryu C, Homer RJ, Herzog EL. The airway in idiopathic pulmonary fibrosis: protecting the lung or promoting disease? *Am J Respir Crit Care Med*. (2016) 193:1081–2. doi: 10.1164/rccm.201601-0055ED

7. Miller ER, Putman RK, Diaz AA, Xu H, San José Estépar R, Araki T, et al. Increased Airway Wall thickness in interstitial lung abnormalities and idiopathic pulmonary fibrosis. *Ann Am Thorac Soc*. (2019) 16:447–54. doi: 10.1513/AnnalsATS.201806-424OC

8. Li YZ, Jin GY, Chae KJ, Han YM. Quantitative assessment of airway changes in fibrotic interstitial lung abnormality patients by chest CT according to cumulative cigarette smoking. *Tomography*. (2022) 8:1024–32. doi: 10.3390/tomography8020082

9. Ash SY, Harmouche R, Ross JC, Diaz AA, Rahaghi FN, Vegas Sanchez-Ferrero G, et al. Interstitial features at chest CT enhance the deleterious effects of emphysema in the COPDGene cohort. *Radiology*. (2018) 288:600–9. doi: 10.1148/radiol.2018172688

10. Kirby M, Tanabe N, Tan WC, Zhou G, Obeidat M, Hague CJ, et al. Total airway count on computed tomography and the risk of chronic obstructive pulmonary disease progression. Findings from a population-based study. *Am J Respir Crit Care Med*. (2018) 197:56–65. doi: 10.1164/rccm.201704-0692OC

11. Achenbach T, Weinheimer O, Brochhausen C, Hollemann D, Baumbach B, Scholz A, et al. Accuracy of automatic airway morphometry in computed tomography-correlation of radiological-pathological findings. *Eur J Radiol*. (2012) 81:183–8. doi: 10.1016/j.ejrad.2010.09.012

12. Oelsner EC, Ortega VE, Smith BM, Nguyen JN, Manichaikul AW, Hoffman EA, et al. A genetic risk score associated with chronic obstructive pulmonary disease susceptibility and lung structure on computed tomography. *Am J Respir Crit Care Med*. (2019) 200:721–31. doi: 10.1164/rccm.201812-2355OC

13. Bhatt SP, Bodduluri S, Nakhmani A, Kim YI, Reinhardt JM, Hoffman EA, et al. Sex differences in Airways at Chest CT: results from the COPDGene cohort. *Radiology*. (2022) 305:699–708. doi: 10.1148/radiol.212985

14. Washko GR, Hunninghake GM, Fernandez IE, Nishino M, Okajima Y, Yamashiro T, et al. Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med*. (2011) 364:897–906. doi: 10.1056/NEJMoa1007285

15. Konietzke P, Brunner C, Konietzke M, Wagner WL, Weinheimer O, Heußel CP, et al. GOLD stage-specific phenotyping of emphysema and airway disease using quantitative computed tomography. *Front Med (Lausanne)*. (2023) 10:1184784. doi: 10.3389/fmed.2023.1184784

16. Gietema HA, Edwards LD, Coxson HO, Bakke PS, Investigators ECLIPSE. Impact of emphysema and airway wall thickness on quality of life in smoking-related COPD. *Respir Med*. (2013) 107:1201–9. doi: 10.1016/j.rmed.2013.04.016

17. Nakano Y, Wong JC, de Jong PA, Buzatu L, Nagao T, Coxson HO, et al. The prediction of small airway dimensions using computed tomography. *Am J Respir Crit Care Med*. (2005) 171:142–6. doi: 10.1164/rccm.200407-874OC

18. Johannessen A, Skorge TD, Bottai M, Grydeland TB, Nilsen RM, Coxson H, et al. Mortality by level of emphysema and airway wall thickness. *Am J Respir Crit Care Med*. (2013) 187:602–8. doi: 10.1164/rccm.201209-1722OC

19. Bozzetti F, Paladini I, Rabaiotti E, Franceschini A, Alfieri V, Chetta A, et al. Are interstitial lung abnormalities associated with COPD? A nested case-control study. *Int J Chron Obstruct Pulmon Dis*. (2016) 11:1087–96. Published 2016 May 26. doi: 10.2147/COPD.S103256

20. Hata A, Schiebler ML, Lynch DA, Hatabu H. Interstitial lung abnormalities: state of the art. *Radiology*. (2021) 301:19–34. doi: 10.1148/radiol.2021204367

21. Jin GY, Lynch D, Chawla A, Garg K, Tammemagi MC, Sahin H, et al. Interstitial lung abnormalities in a CT lung cancer screening population: prevalence and progression rate. *Radiology*. (2013) 268:563–71. doi: 10.1148/radiol.13120816

22. Choi JY, Song JW, Rhee CK. Chronic obstructive pulmonary disease combined with interstitial lung disease. *Tuberc Respir Dis (Seoul)*. (2022) 85:122–36. doi: 10.4046/trd.2021.0141

23. Putman RK, Hatabu H, Araki T, Gudmundsson G, Gao W, Nishino M, et al. Association between interstitial lung abnormalities and all-cause mortality. *JAMA*. (2016) 315:672–81. doi: 10.1001/jama.2016.0518

24. Sanders JL, Putman RK, Dupuis J, Xu H, Murabito JM, Araki T, et al. The Association of Aging Biomarkers, interstitial lung abnormalities, and mortality [published correction appears in am J Respir Crit care med. 2022 Sep 1; 206(5): 653]. *Am J Respir Crit Care Med*. (2021) 203:1149–57. doi: 10.1164/rccm.202007-2993OC

25. Copley SJ, Wells AU, Hawtin KE, Gibson DJ, Hodson JM, Jacques AET, et al. Lung morphology in the elderly: comparative CT study of subjects over 75 years old versus those under 55 years old. *Radiology*. (2009) 251:566–73. doi: 10.1148/radiol.2512081242

26. Putman RK, Gudmundsson G, Axelsson GT, Hida T, Honda O, Araki T, et al. Imaging patterns are associated with interstitial lung abnormality progression and mortality. *Am J Respir Crit Care Med*. (2019) 200:175–83. doi: 10.1164/rccm.201809-1652OC

27. Patel BD, Coxson HO, Pillai SG, Agustí AG, Calverley PM, Donner CF, et al. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2008) 178:500–5. doi: 10.1164/rccm.200801-059OC

28. Aziz ZA, Wells AU, Desai SR, Ellis SM, Walker AE, MacDonald S, et al. Functional impairment in emphysema: contribution of airway abnormalities and distribution of parenchymal disease. *AJR Am J Roentgenol*. (2005) 185:1509–15. doi: 10.2214/AJR.04.1578

29. Mohamed Hoesein FA, de Jong PA, Lammers JW, Mali WPTM, Schmidt M, de Koning HJ, et al. Airway wall thickness associated with forced expiratory volume in 1 second decline and development of airflow limitation. *Eur Respir J*. (2015) 45:644–51. doi: 10.1183/09031936.00020714

30. Grenier PA. Relationship between interstitial lung abnormalities and emphysema in smokers with and those without COPD. *Radiology*. (2018) 288:610–1. doi: 10.1148/radiol.2018180270

31. Ohgiya M, Matsui H, Tamura A, Kato T, Akagawa S, Ohta K. The evaluation of interstitial abnormalities in group B of the 2011 global initiative for chronic obstructive lung disease (GOLD) classification of chronic obstructive pulmonary disease (COPD). *Intern Med*. (2017) 56:2711–7. doi: 10.2169/internalmedicine.8406-16

32. Ono M, Kobayashi S, Hanagama M, Ishida M, Sato H, Makiguchi T, et al. Clinical characteristics of Japanese patients with chronic obstructive pulmonary disease (COPD) with comorbid interstitial lung abnormalities: a cross-sectional study. *PLoS One*. (2020) 15:e0239764. doi: 10.1371/journal.pone.0239764

33. Menon AA, Putman RK, Sanders JL, Hino T, Hata A, Nishino M, et al. Interstitial lung abnormalities, emphysema, and spirometry in smokers. *Chest*. (2022) 161:999–1010. doi: 10.1016/j.chest.2021.10.034

34. Kahnert K, Jörres RA, Kauczor HU, Alter P, Trudzinski FC, Herth F, et al. Standardized airway wall thickness Pi10 from routine CT scans of COPD patients as imaging biomarker for disease severity, lung function decline, and mortality. *Ther Adv Respir Dis*. (2023) 17:175346662211486. doi: 10.1177/17534666221148663

35. Doyle TJ, Washko GR, Fernandez IE, Nishino M, Okajima Y, Yamashiro T, et al. Interstitial lung abnormalities and reduced exercise capacity. *Am J Respir Crit Care Med*. (2012) 185:756–62. doi: 10.1164/rccm.201109-1618OC

36. Hatabu H, Hunninghake GM, Richeldi L, Brown KK, Wells AU, Remy-Jardin M, et al. Interstitial lung abnormalities detected incidentally on CT: a position paper from the Fleischner society. *Lancet Respir Med*. (2020) 8:726–37. doi: 10.1016/S2213-2600(20)30168-5

Echo time-dependent observed T1 and quantitative perfusion in chronic obstructive pulmonary disease using magnetic resonance imaging

Simon M. F. Triphan[1,2]*, Marilisa Konietzke[1,2,3], Jürgen Biederer[1,2,4,5], Monika Eichinger[1,2,6], Claus F. Vogelmeier[7], Rudolf A. Jörres[8], Hans-Ulrich Kauczor[1,2,6], Claus P. Heußel[1,6], Bertram J. Jobst[1,2,6] and Mark O. Wielpütz[1,2,6], on behalf of the COSYCONET study group

[1]Department of Diagnostic and Interventional Radiology, Subdivision of Pulmonary Imaging, University Hospital of Heidelberg, Heidelberg, Germany, [2]Translational Lung Research Center Heidelberg (TLRC), Member of the German Center for Lung Research (DZL), Heidelberg, Germany, [3]Boehringer Ingelheim Pharma GmbH and Co. KG, Biberach an der Riß, Germany, [4]Faculty of Medicine, University of Latvia, Riga, Latvia, [5]Faculty of Medicine, Christian-Albrechts-Universität zu Kiel, Kiel, Germany, [6]Department of Diagnostic and Interventional Radiology With Nuclear Medicine, Thoraxklinik at the University Hospital of Heidelberg, Heidelberg, Germany, [7]Department of Medicine, Pulmonary and Critical Care Medicine, Philipps-University of Marburg (UMR), Member of the German Center for Lung Research (DZL), Marburg, Germany, [8]Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), University Hospital, Ludwig Maximilians University (LMU) Munich, Munich, Germany

Introduction: Due to hypoxic vasoconstriction, perfusion is interesting in the lungs. Magnetic Resonance Imaging (MRI) perfusion imaging based on Dynamic Contrast Enhancement (DCE) has been demonstrated in patients with Chronic Obstructive Pulmonary Diseases (COPD) using visual scores, and quantification methods were recently developed further. Inter-patient correlations of echo time-dependent observed T_1 [T_1(TE)] have been shown with perfusion scores, pulmonary function testing, and quantitative computed tomography. Here, we examined T_1(TE) quantification and quantitative perfusion MRI together and investigated both inter-patient and local correlations between T_1(TE) and quantitative perfusion.

Methods: 22 patients (age 68.0 ± 6.2) with COPD were examined using morphological MRI, inversion recovery multi-echo 2D ultra-short TE (UTE) in 1–2 slices for T_1(TE) mapping, and 4D Time-resolved angiography With Stochastic Trajectories (TWIST) for DCE. T_1(TE) maps were calculated from 2D UTE at five TEs from 70 to 2,300 μs. Pulmonary Blood Flow (PBF) and perfusion defect (QDP) maps were produced from DCE measurements. Lungs were automatically segmented on UTE images and morphological MRI and these segmentations registered to DCE images. DCE images were separately registered to UTE in corresponding slices and divided into corresponding subdivisions. Spearman's correlation coefficients were calculated for inter-patient correlations using the entire segmented slices and for local correlations separately using registered images and subdivisions

for each TE. Median $T_1(TE)$ in normal and defect areas according to QDP maps were compared.

Results: Inter-patient correlations were strongest on average at $TE_2 = 500\,\mu s$, reaching up to $|\rho| = 0.64$ for T_1 with PBF and $|\rho| = 0.76$ with QDP. Generally, local correlations of T_1 with PBF were weaker at TE_2 than at TE_1 or TE_3 and with maximum values of $|\rho| = 0.66$ (from registration) and $|\rho| = 0.69$ (from subdivision). In 18 patients, T_1 was shorter in defect areas than in normal areas, with the relative difference smallest at TE_2.

Discussion: The inter-patient correlations of T_1 with PBF and QDP found show similar strength and TE-dependence as those previously reported for visual perfusion scores and quantitative computed tomography. The local correlations and median T_1 suggest that not only base T_1 but also the TE-dependence of observed T_1 in normal areas is closer to that found previously in healthy volunteers than in defect areas.

1 Introduction

Monitoring disease progression and therapy response in chronic obstructive pulmonary disease (COPD) is commonly based on pulmonary function testing (PFT), which provides global measures for lung function impairment ([1]). Computed tomography (CT) is used for the regional assessment of lung structure and distinction of COPD phenotypes ([2]). Magnetic resonance imaging (MRI) of the lungs is particularly challenging due to the low proton density in lung parenchyma, respiratory motion and especially the extremely short T_2^* relaxation time caused by the structure of the lungs ([3]). MRI was suggested as a radiation-free, non-invasive alternative to PFT and CT and has the additional benefit of techniques for the regional assessment of functional impairment. MRI methods based on image registration and the oscillation of signal intensity allow for the quantification of pulmonary perfusion and ventilation ([4], [5]). However, arguably the most practical and established approach for functional lung MRI in clinical routine is dynamic contrast enhanced (DCE) MRI. In the lungs, local perfusion is regulated by hypoxic pulmonary vasoconstriction which links perfusion directly to ventilation. Accordingly, the detection of perfusion defects can thus be used to identify abnormal ventilation as long as ventilation/perfusion mismatches can be disregarded. Reduction of the pulmonary vascular bed in emphysema also results in perfusion defects. A visual scoring system for DCE-MRI by radiologist readers has been demonstrated to provide a measure of functional lung abnormality that correlates well with other clinical metrics, including lung function tests in COPD and cystic fibrosis ([6]–[10]).

The determination of this visual score requires trained radiologists and is simultaneously limited in spatial resolution to the level of lung lobes. It is also severely limited in numerical resolution since scores cannot practically have fine graduations and it also essentially disregards the difference in volume between lung lobes. Quantification of perfusion from DCE-MRI data has been established in other organs and explored in the lungs as well ([11], [12]). For perfusion quantification, both the calculation of perfusion metrics like pulmonary blood flow (PBF), pulmonary blood volume (PBV) and mean transit time (MTT) as well as the determination of the extent of defects have been shown ([13]). While these have been examined mainly as global metrics for the entire lung, especially considering the fraction of defects as perfusion defects in percent (QDP), they are also available locally for each image voxel ([14]).

T_1 quantification in the lungs was originally pursued in the context of oxygen-enhanced imaging ([15]). In these studies, significantly lower baseline lung T_1 was found in patients with COPD even without oxygen supplementation, and median T_1 was found to correlate with the visual MRI perfusion score ([8]). In subsequent work, it was further demonstrated that observed lung T_1 [$T_1(TE)$] is dependent on the echo time (TE) of the measurement. It was subsequently assumed that $T_1(TE)$ reflects different compartments of lung tissue protons inside each imaging voxel (protons in blood and extravascular protons), since the inter-patient correlations between $T_1(TE)$ and the MRI morphology and MRI perfusion score as well as other metrics vary with TE in COPD and cystic fibrosis ([16]–[18]). It was hypothesized that $T_1(TE)$ has the potential to differentiate tissue abnormalities such as inflammation or emphysema from perfusion abnormalities on the level of image voxels without the need for contrast material. Previous studies have so far examined the relationship between $T_1(TE)$ and DCE-MRI perfusion using only visual MRI perfusion scores on the level of the entire lungs for median T_1, and on the level of lobes for visually scored T_1 ([8], [17], [18]). The present work was conducted to further investigate the relationship of $T_1(TE)$ with quantitative perfusion metrics on a local level including voxel-wise correlations.

2 Methods

2.1 Patient characteristics

The study included participants of a single study center within the German multicenter COPD cohort study COSYCONET ("COPD and SYstemic consequences-COmorbidities NETwork," ClinicalTrials.gov Identifier: NCT01245933) (19). The overarching COSYCONET study as well as the part of the imaging-based sub-study ["Image-based structural and functional phenotyping of the COSYCONET cohort using MRI and CT (MR-COPD)," ClinicalTrials.gov Identifier: NCT02629432] described here were approved by the responsible ethics committees of the coordinating centers [Institutional Review Board of the Medical Faculty of the University of University of Marburg (200/09) and of the University of Heidelberg (S-656/2012), Germany]. The participants of the present sub-study gave their written informed consent to undergo extensive clinical assessment including lung function testing, non-contrast CT, and morpho-functional MRI. The patient characteristics are summarized in Table 1. Some patients were included in separate previous reports, which did not assess the analyses made in the present study (13, 18).

2.2 Magnetic resonance imaging acquisition

All measurements were performed on the same 1.5 T scanner (Magnetom Aera, Siemens Healthineers, Erlangen, Germany). A standardized protocol was employed including T_1- and T_2-weighted morphological MRI sequences. Of those, the T_1-weighted morphological Volumetric Interpolated Breath-hold Examination (VIBE) measurements were used for segmentation. Imaging parameters for those were: Transverse VIBE: TR = 2.61 ms, TE = 1.69 ms, Field of View (FoV) $400 \times 300 \times 4$ mm³, matrix size $320 \times 240 \times 88$. Coronal VIBE: TR = 3.35 ms, TE = 1.63 ms, FoV $400 \times 400 \times 4$ mm³, matrix size $288 \times 288 \times 56$. A keyhole-based 3D gradient echo sequence [Time-resolved angiography With Stochastic Trajectories (TWIST), Siemens Healthineers, Erlangen, Germany], was used for DCE perfusion imaging, resulting in 1.7 s effective time resolution over 20 3D image timepoints. Imaging parameters were: up to 56 slices at 5 mm thickness, matrix size 208×256, FoV 366×450 mm², TE 0.76 ms, and flip angle 20°. DCE measurements utilized intravenous contrast injection of 2 mL Gadobutrol 1 mmoL/mL (Gadovist, Bayer Vital, Leverkusen, Germany) injected at a rate of

TABLE 1 Patient characteristics.

Total no. of patients	22
Sex	12 female, 10 male
Age (years)	68.0 ± 6.2 (51.0–77.0)
Weight (kg)	77.3 ± 16.2 (49.0–115.0)
Height (cm)	169.8 ± 6.8 (156.0–181.0)
FEV1% predicted	59.9 ± 15.2 (34.0–84.0)
GOLD stage	1.6 ± 1.3 (0–4)

Values are given as mean ± standard deviation with the range in parentheses.

4 mL/s, followed by a 30 mL 0.9% NaCl chaser. Both VIBE and TWIST were acquired in inspiratory breath-holds.

T_1(TE) mapping was based on an inversion recovery multi-echo 2D ultra-short TE sequence, applied directly before contrast injection (20). Echos at $TE_{1-5} = 70$, 500, 1,200, 1,650, and 2,300 μs were acquired split into two separate measurements with TE_1, TE_3, and TE_5 in the first and TE_2 and TE_4 in the second measurement. Imaging parameters were: One or two coronal slices at 15 mm thickness, matrix size 128×128, FoV 500×500 mm², TR = 5 ms, flip angle 6°, and total acquisition time 4 min (1 slice) or 8 min (2 slices). To cover as much lung parenchyma as possible, the first slice was placed through the descending aorta and the second slice 30 mm ventral. For each measurement, a total of 6,000 radial spokes in a golden angle distribution were acquired, divided into blocks starting with an adiabatic inversion pulse followed by 300 radial spokes and separated by 3 s delays. These were then sorted by inversion time (TI), split using a sliding window 120 spokes wide with 60 spokes step-width and reconstructed using a non-uniform Fourier transform implemented in MATLAB (Mathworks, Natick, United States). To achieve ultra-short echo times, pairs of half sinc-pulses with opposing slice-selection gradients, 1,250 μs pulse duration and time-bandwidth-product 2 were employed. For each TE, T_1 maps were calculated per voxel from the time-course along TI, based on the UTE images in each slice (21). UTE measurements were acquired during free breathing and gated to expiration using 50% of data, as described in Triphan et al. (20).

2.3 Magnetic resonance imaging analysis

Magnetic resonance images were segmented independently for UTE and DCE measurements: In UTE images, segmentation was performed on images at TE_3 using simple region growing (for details, see the online supplement). Since the DCE images are three-dimensional and designed to maximize contrast between tissue with and without contrast agent, these were not segmented directly. Instead, a region growing-based segmentation algorithm was applied to morphological images with a higher resolution and the resulting segmentation masks were then registered to the DCE images, as described previously (13, 22).

For perfusion quantification, an arterial input function was detected automatically and then used to calculate a residual function R(t) by applying a voxel-wise deconvolution with the time-dependent signal in the DCE images (23). Pulmonary blood flow (PBF) maps were derived from R(t) by determining the maximum along the time-course in each individual voxel (24). To calculate defect classification maps and QDP, the following steps were performed: The timepoint of maximum contrast enhancement in the entire lungs (t_{max}) was determined by taking the mean of R(t) in the entire lung for every timepoint. Otsu's method was then applied to the values of $R(t_{max})$ within the lungs to determine two thresholds (25). Defect classification maps were produced by classifying all voxels below the lower threshold as perfusion defects and the remainder as normal (13). The fraction of defect voxels within the segmented lungs on these maps relative to the total number of segmented voxels was defined as QDP. For all of the following steps, the three partitions of the 3D DCE images that correspond to the location of the UTE slice(s) were determined and averaged to produce a thicker projection image. Note that t_{max} and the

FIGURE 1
Example horizontal **(A)** and vertical **(B)** deformation field as determined by image registration, scaled in number of voxels. Segmentations with subdivided areas for a DCE image **(C)** and the corresponding 2D UTE image **(D)**. Areas that correspond to each other in the subdivision are drawn in identical colors.

classification thresholds were calculated based on the entire segmented lungs but QDP to be compared to $T_1(TE)$ only on these slices corresponding to UTE measurements.

To facilitate local (intra-patient) comparison of $T_1(TE)$ and quantitative perfusion maps, two approaches were investigated: (1) DCE images were averaged over their entire time-course and registered onto the corresponding UTE images using Advanced Normalization Tools (ANTs) (26) using the same parameters for all patients. The deformation fields produced by this registration were then applied to the PBF and defect classification maps as well as the lung segmentations derived from morphological imaging to allow for voxel-wise comparisons. An example of these deformation fields is given in Figures 1A,B. (2) On each slice, the left and right lungs were individually subdivided into areas with the following approach: Each lung was subdivided into 10 strips of equal volume vertically and then each of these strips was separated into 10 blocks of equal volume horizontally. By applying this subdivision algorithm on the segmentations of the DCE images and the UTE independently, this approach yields subdivided areas for both different resolutions and different respiratory states, which nevertheless correspond to the same physical volumes in the lungs and thus allow for correlations and comparisons without a registration step. An example of this subdivision is shown in Figures 1C,D.

2.4 Statistical analyses

Inter-patient correlations were calculated for median $T_1(TE)$ with median PBF (ρ_{IPBF}) and QDP (ρ_{IQDP}) within the segmented slices. For each patient, Spearman's correlation coefficients between $T_1(TE)$ and PBF were calculated separately for registered voxels (ρ_{regPBF}) and subdivided areas (ρ_{subPBF}), respectively. Registered voxels in UTE were classified as normal perfusion or perfusion defect according to the corresponding classification map from DCE, and the median $T_1(TE)$ in each fraction was compared by Wilcoxon rank-sum tests. For subdivided areas, the fraction of defect voxels (i.e., local QDP) in each subdivided area was correlated with median $T_1(TE)$ in the corresponding area, analogous to PBF, as the local coefficient ρ_{subQDP}. A value of $p < 0.05$ was considered statistically significant.

3 Results

Suitable UTE and DCE measurements were available from 22 patients. In four patients, one of the two measurements comprising the UTE acquisition was mismatched in one slice or insufficient time was available to acquire both slices. Figure 2 shows an example set of base images and parameter maps in one patient.

The coefficients of the inter-patient correlation of $T_1(TE)$ with median PBF for all corresponding lung slices were $\rho_{IPBF1-5} = 0.56$, 0.64, 0.59, 0.60, and 0.29 at TE_{1-5} and the respective correlations of $T_1(TE)$ with QDP were $\rho_{IQDP1-5} = -0.56$, -0.76, -0.63, -0.53, and -0.26. Of these, only the correlations at TE_{1-4} were significant (at $p < 0.02$). Table 2 shows the mean local correlation coefficients of $T_1(TE)$ with PBF based on both methods for associating UTE with DCE images, as well as the correlation of $T_1(TE)$ with QDP. The values of ρ_{regPBF} and ρ_{subPBF} for individual patients are also shown in Figure 3 to illustrate the inter-patient variation.

At $TE_1 = 70\,\mu s$, in 18 out of 22 patients, T_1 was shorter in voxels classified as perfusion defect when compared to the rest of the lung voxels classified as normal perfusion in the respective individual. In three patients, no significant difference was found and in one patient, T_1 was longer in the defect area. While, as shown in Figure 4A, T_1 in normal (non-defect) areas and ΔT_1, the relative difference, varied strongly among patients, median T_1 in defect areas was shorter at all TEs when considered over all patients. Notably, as shown in Figure 4B, mean ΔT_1 was smaller at TE_2 than at TE_1 and TE_3, with $\Delta T_1 = 6.2$, 4.0, 6.2, 9.1, and 16.6%, similar to the local correlations observed.

As demonstrated in Figure 2, visual inspection of parameter maps suggested that stronger local correlations appeared in patients where areas with clearly recognizable perfusion and longer T_1 remained. To investigate this observation, the inter-patient correlation between median $T_1(TE_1)$ and the local correlation coefficients of T_1 and PBF was calculated. At $TE_1 = 70\,\mu s$, this was $\rho_{regcorr} = 0.68$ when using registration and $\rho_{subcorr} = 0.32$ when using subdivided areas.

4 Discussion

In this work, inter-patient correlations of $T_1(TE)$ with quantitative perfusion measures showed the strongest correlation coefficient at $TE_2 = 500\,\mu s$. Local voxel-wise correlations instead showed the

FIGURE 2
(A) Original image from DCE MRI, averaged over all timepoints and three slices. **(B)** The same image, registered to the UTE image. **(C)** Image derived from inversion recovery UTE that was used as registration target. **(D)** PBF map after application of the distortion field produced by the image registration, with perfusion defects delineated in blue. **(E)** T_1 map acquired using UTE at $TE_1 = 70\,\mu s$. **(F)** T_1 map at $TE_2 = 500\,\mu s$.

TABLE 2 Mean local correlations.

TE	70 μs	500 μs	1,200 μs	1,650 μs	2,300 μs
ρ_{regPBF}	0.44 ± 0.13 (22)	0.39 ± 0.12 (22)	0.42 ± 0.13 (22)	0.39 ± 0.13 (22)	0.24 ± 0.13 (21)
ρ_{subPBF}	0.46 ± 0.15 (22)	0.39 ± 0.15 (22)	0.45 ± 0.16 (22)	0.43 ± 0.16 (22)	0.38 ± 0.16 (21)
ρ_{subQDP}	-0.33 ± 0.16 (18)	-0.21 ± 0.16 (12)	-0.23 ± 0.17 (15)	-0.22 ± 0.15 (16)	-0.17 ± 0.16 (15)

Mean local correlations of T_1(TE) with PBF, based on image registration for voxel-wise correlation (ρ_{regPBF}) and based on subdivisions (ρ_{subPBF}), as well as with QDP based on subdivisions (ρ_{subQDP}). Values are given as means with standard deviations and the number of patients in whom the correlation was significant in brackets.

FIGURE 3
Spearman's correlation coefficients for the correlation of T_1(TE) with PBF, **(A)** voxel-wise correlation based on image registration and **(B)** based on subdivisions, respectively. Each patient is shown with the same color in both graphs.

FIGURE 4
(A) Relative difference of T_1 at $TE_1 = 70\,\mu s$ between voxels classified as perfusion defect and voxels classified as normal. Each point represents one individual. **(B)** Mean of individual median T_1 over all patients at each TE, according to defect classification.

strongest correlations at $TE_1 = 70\,\mu s$ and in fact weaker correlations at TE_2 than at $TE_3 = 1{,}200\,\mu s$. Finally, at TE_1, in 18 out of 22 patients, T_1 was shorter in voxels classified as defect than in normal areas.

The visual score for perfusion MRI has been established in COPD and cystic fibrosis, and were shown to correlate with clinical metrics (6, 10) and with T_1 at conventional echo times (8). Subsequently, it was demonstrated that the inter-patient correlation of median $T_1(TE)$ in the entire lungs with this perfusion score and quantitative CT parameters depends on TE. Specifically, the strongest correlation was found at $TE = 500\,\mu s$ (18). This agrees with inter-patient correlations of $T_1(TE)$ with quantitative perfusion measures shown in this work which, also gave the strongest correlation at TE_2.

Perfusion defects in percent was developed to compensate for the vulnerabilities in the quantification of PBF related to non-linearity of Gadolinium-concentration with MRI signal, artifacts, limited temporal resolution, and low SNR in the lung parenchyma (which is particularly low in diseased areas of the lung parenchyma) (14, 27, 28). This is supported by stronger inter-patient correlations between QDP and the visual perfusion score than were found with visual perfusion scores previously (13). In our present work, the inter-patient correlation between $T_1(TE)$ and QDP was also stronger at TE_2 and stronger than with PBF. However, while the correlations of T_1 with PBF based on subdivision were slightly stronger than those based on image registration, only very weak correlations of T_1 with local QDP within subdivisions were found. This is most likely to due to many subdivided areas containing only defect or only normal voxels and thus not describing larger defect areas well. QDP derived from DCE using a different algorithm was found to locally match QDP derived from contrast-free PREFUL MRI (28). QDP derived from $T_1(TE)$ (rather than DCE-MRI) was shown to correlate similarly to $T_1(TE)$ itself in patients with *CF* (17), but not in patients in COPD (18). Thus, we did not re-investigate defect classification based on $T_1(TE)$ in the present work.

Unlike the inter-patient correlations, the local correlations found in the present study were highest at TE_1 and higher at TE_3 than at TE_2. This can be attributed to lung areas more affected by disease having not only shorter T_1 but also exhibiting a different dependence of T_1 on TE. This is supported by the behavior of T_1 in areas classified as perfusion defect or normal, when averaged over the patient collective: There, the average difference between these areas decreased from TE_1 to TE_2 and then

increased again at TE_3. This behavior becomes clearer when considering the data published previously comparing patients with COPD to healthy volunteers (18). There, the TE-dependence of T_1 was much more gradual and observable at longer TEs in healthy volunteers, while the observed T_1 in patients with COPD did not notably increase further after $TE = 500\,\mu s$, which also meant that the separation between these groups was smallest at $TE = 500\,\mu s$. Assuming that the TE-dependence of T_1 in areas classified as normal through perfusion is closer to that in healthy volunteers and in defect areas closer to that in the patient group, this is consistent with the results seen in this work. It should however be noted that in the present collective areas classified as normal are still quite abnormal in T_1 compared to values observed in healthy volunteers. Thus, our results show that both base T_1 and its TE-dependence differ in areas of normal and defective perfusion. This further supports the hypothesis that it may be possible to separate the effects of local tissue changes and perfusion on T_1 and investigate both with a single measurement. However, this is not yet possible with the current approach.

A large variation of local correlations was observed, with the strongest observed local correlations reaching the same magnitude as the inter-patient correlations. The overall visual impression of T_1 and PBF parameter maps was that strong correlations were mainly present in patients with some volume of remaining functional lung tissue. In those cases, a notable spread of perfusion values exists that can be correlated in the first place. As such, we assumed that low local correlations might indicate relatively uniformly severely damaged lungs and thus also short T_1. In order to confirm this assumption, we calculated the inter-patient correlation of median $T_1(TE_1)$ with local correlation coefficients in each patient. That this correlation was significant supports the assertion that weak local correlations were primarily found in patients with more damaged lungs. However, this correlation was much weaker when based on subdivided areas. Since regions with long T_1 are unlikely to coincide with the borders of subdivided areas, they can be assumed to describe these less well than the direct voxel-wise correlations. Similarly, in cases where a perfusion deficit was present all over the lung with corresponding short T_1, these borders may matter less. At the same time, subdivided areas were less vulnerable to outlier values produced by quantification at very low SNR due to reduced proton density in severe emphysema. In further research, it may be worthwhile to investigate this in milder or

presymptomatic cases of COPD, where a larger spread of values to correlate could be investigated. Additionally, MRI in general struggles in patients with severe emphysema simply due to lack of proton signal, though we might surmise that T_1 mapping at $TE_1 = 70\,\mu s$ should be better able to cope. However, in this work we were limited by the possibilities provided by DCE imaging.

Since in COPD patients, the TE-dependence of T_1 is shifted to shorter TEs in comparison to healthy subjects, a tight spacing of TEs was necessary to depict the effects of interest in this work (16, 18): The TE-dependence of local and inter-patient correlations discussed here was visible at $TE_1 = 70\,\mu s$ and $TE_2 = 500\,\mu s$, but it is not possible to measure T_1 at both TEs in one TR due to the time required for the read dephasing gradient in between. Accordingly, the acquisition had to be split into two sets of echoes, which were implemented in separate measurements. This had the additional problem that the dc-signal used for respiratory gating is not necessarily aligned between those two measurements in case the patient changes how deep they breathe between them. Additionally, the two measurements could be misaligned due to operator error, making the entire slice unusable. After this study was concluded, we redesigned the sequence such that different sets of TEs could be acquired in alternating successive TRs which has neither of these disadvantages while still providing the required tight spacing of TEs.

The 4D MRI used for perfusion quantification was a single measurement, which meant that if the contrast injection was administered too early or late or the patient holding their breath for unsufficient time, it was entirely unusable for quantification. Since the UTE measurement was not part of the primary study protocol, we could only plan two slices and did not have time to retry in case of mistakes. However, these were separately acquired and each split into two measurements, as stated above. In cases where one slices was unusable due to operator error in at least one such sub-measurement, we decided to nevertheless use the remaining slice since the total number of patient measurements was so small. Thus, the UTE images being based on 2D slices was a major limitation of this work: a reliable 3D implementation of the multi-echo inversion recovery UTE would allow for a better quality of registration and give better coverage of the lungs, which would also provide improved comparability with perfusion images.

We investigated the approach to derive local correlations from subdivided areas in addition to image registration to avoid reliance on the registration algorithm, as it is not practically possible to guarantee a successful registration: The quality of image registration limits the comparability and thus voxel-wise correlation of parameter maps. The approach using subdivided areas was designed as an alternative that avoids this problem by increasing the size of spatial units to potentially correct for some imprecision in registration but only with the limitation that it is based on the assumption that both lungs have equal volume. This was implemented in order to be able to use the same number and pattern of subdivisions in each patient. In future work, the accuracy of local correlations could be increased by generating a specific subdivision pattern for each patient, possibly in 3D, which employs different number of subdivisions for each lung. Furthermore, the two segmentations may be subtly different beyond the issue of comparing two- with three-dimensional data. Finally, a relevant limitation for both approaches was that DCE images were acquired in inspiratory breath-holds, while UTE images were acquired during free breathing and reconstructed to expiration retrospectively.

As noted above, the approach discussed here relies entirely on using local perfusion as an indicator of lung function. However,

even in early COPD stages, ventilation/perfusion (V/Q) mismatches have been reported, suggesting a blocked hypoxic pulmonary vasoconstriction mechanism to inhibit inflammation (29, 30). Nevertheless, there remains a significant association between ventilation and perfusion in COPD patients and correlation of the extent of V/Q mismatches with disease severity was relatively modest (31). To incorporate the effect of local V/Q mismatches, an additional direct measure of ventilation could be combined with perfusion imaging techniques. While non-contrast enhanced lung perfusion imaging methods have been introduced that also depict ventilation, it should be noted that the 2D UTE sequence used in this work was initially developed for oxygen-enhanced imaging since both T_1 and $T_2^{\star}$ (which was not considered in this work but can be quantified from the same measurement) depend on the presence of molecular oxygen (dissolved in tissue and as gas in the alveoli, respectively) (4, 20). An additional acquisition while supplying subjects with oxygen could thus be used to determine local ventilation directly and thus investigate V/Q mismatches.

In conclusion, we have shown inter-patient correlations of $T_1(TE)$ with quantitative perfusion measures that are comparable to those described previously for $T_1(TE)$ with semi-quantitative visual perfusion scores from perfusion MRI and quantitative CT parameters. Local (intra-patient) correlations were much weaker on average, but the strongest correlations found reached a similar scale as the inter-patient correlation coefficients. Importantly, the TE-dependency of the local correlation coefficients implies that not only base T_1 but also its TE-dependency in normal areas is closer to the behavior previously found in healthy volunteers than in defect areas. That the correlation coefficients themselves correlate with overall average T_1 implies that weak local correlations would be observed in patients with few remaining areas of well-functioning lung tissue and thus low contrasts between local observations. Accordingly, it may be interesting to investigate DCE perfusion and $T_1(TE)$ in a patient cohort with less severe, earlier disease, both individually and in combination.

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.

Ethics statement

The studies involving humans were approved by Institutional Review Board of the Medical Faculty of the University of Marburg (200/09) and of the University of Heidelberg (S-656/2012). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ST: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization,

Writing – original draft, Writing – review & editing. MK: Conceptualization, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. JB: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing – review & editing. ME: Conceptualization, Supervision, Writing – review & editing. CV: Conceptualization, Funding acquisition, Project administration, Resources, Writing – review & editing. RJ: Conceptualization, Funding acquisition, Project administration, Resources, Writing – review & editing. H-UK: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. CH: Conceptualization, Methodology, Project administration, Writing – review & editing. BJ: Conceptualization, Investigation, Methodology, Writing – review & editing. MW: Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

MK was an employee of Boehringer Ingelheim.

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.

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2023.1254003/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Relative difference of T1 at TE1 to TE5 between voxels classified as perfusion defect and voxels classified as normal. Each point represents one individual. This corresponds to Figure 4A at all TE.

References

1. Hashimoto N, Wakahara K, Sakamoto K. The importance of appropriate diagnosis in the practical Management of Chronic Obstructive Pulmonary Disease. *Diagnostics*. (2021) 11. doi: 10.3390/diagnostics11040618

2. Lynch DA, Austin JHM, Hogg JC, Grenier PA, Kauczor H-U, Bankier AA, et al. CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner society. *Radiology*. (2015) 277:192–205. doi: 10.1148/radiol.2015141579

3. Hatabu H, Alsop DC, Listerud J, Bonnet M, Gefter WB. T2* and proton density measurement of normal human lung parenchyma using submillisecond echo time gradient echo magnetic resonance imaging. *Eur J Radiol*. (1999) 29:245–52. doi: 10.1016/S0720-048X(98)00169-7

4. Bauman G, Puderbach M, Deimling M, Jellus V, Chefd'hotel C, Dinkel J, et al. Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of fourier decomposition in proton MRI. *Magn Reson Med*. (2009) 62:656–64. doi: 10.1002/mrm.22031

5. Bauman G, Bieri O. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. *Magn Reson Med*. (2017) 77:336–42. doi: 10.1002/mrm.26096

6. Eichinger M, Optazaite D-E, Kopp-Schneider A, Hintze C, Biederer J, Niemann A, et al. Morphologic and functional scoring of cystic fibrosis lung disease using MRI. *Eur J Radiol*. (2012) 81:1321–9. doi: 10.1016/j.ejrad.2011.02.045

7. Wielpütz MO, Puderbach M, Kopp-Schneider A, Stahl M, Fritzsching E, Sommerburg O, et al. Magnetic resonance imaging detects changes in structure and perfusion, and response to therapy in early cystic fibrosis lung disease. *Am J Respir Crit Care Med*. (2014) 189:956–65. doi: 10.1164/rccm.201309-1659OC

8. Jobst BJ, Triphan SMF, Sedlaczek O, Anjorin A, Kauczor HU, Biederer J, et al. Functional lung MRI in chronic obstructive pulmonary disease: comparison of T1 mapping, oxygen-enhanced T1 mapping and dynamic contrast enhanced perfusion. *PLoS One*. (2015) 10:e0121520. doi: 10.1371/journal.pone.0121520

9. Stahl M, Wielpütz MO, Graeber SY, Joachim MC, Sommerburg O, Kauczor H-U, et al. Comparison of lung clearance index and magnetic resonance imaging for assessment of lung disease in children with cystic fibrosis. *Am J Respir Crit Care Med*. (2017) 195:349–59. doi: 10.1164/rccm.201604-0893OC

10. Wielpütz MO, Eichinger M, Wege S, Eberhardt R, Mall MA, Kauczor H-U, et al. Mid-term reproducibility of chest MRI in adults with clinically stable cystic fibrosis and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2019) 200:103–7. doi: 10.1164/rccm.201812-2356LE

11. Hatabu H, Tadamura E, Levin DL, Chen Q, Li W, Kim D, et al. Quantitative assessment of pulmonary perfusion with dynamic contrast-enhanced MRI. *Magnet Reson Med*. (1999) 42:1033–8. doi: 10.1002/(sici)1522-2594(199912)42:6<1033:aid-mrm7>3.0.co;2-7

12. Ohno Y, Hatabu H, Murase K, Higashino T, Kawamitsu H, Watanabe H, et al. Quantitative assessment of regional pulmonary perfusion in the entire lung using three-dimensional ultrafast dynamic contrast-enhanced magnetic resonance imaging: preliminary experience in 40 subjects. *J Magn Reson Imaging*. (2004) 20:353–65. doi: 10.1002/jmri.20137

13. Schiwek M, Triphan SMF, Biederer J, Weinheimer O, Eichinger M, Vogelmeier CF, et al. Quantification of pulmonary perfusion abnormalities using DCE-MRI in COPD: comparison with quantitative CT and pulmonary function. *Eur Radiol*. (2022) 32:1879–90. doi: 10.1007/s00330-021-08229-6

14. Konietzke M, Triphan SMF, Eichinger M, Bossert S, Heller H, Wege S, et al. Unsupervised clustering algorithms improve the reproducibility of dynamic contrast-enhanced magnetic resonance imaging pulmonary perfusion quantification in muco-obstructive lung diseases. *Front Med*. (2022) 9:1022981. doi: 10.3389/fmed.2022.1022981

15. Jakob PM, Wang T, Schultz G, Hebestreit H, Hebestreit A, Hahn D. Assessment of human pulmonary function using oxygen-enhanced T1 imaging in patients with cystic fibrosis. *Magn Reson Med.* (2004) 51:1009–16. doi: 10.1002/mrm.20051

16. Triphan SMF, Jobst BJ, Breuer FA, Wielpütz MO, Kauczor H-U, Biederer J, et al. Echo time dependence of observed T1 in the human lung. *J Magn Reson Imaging.* (2015) 42:610–6. doi: 10.1002/jmri.24840

17. Triphan SMF, Stahl M, Jobst BJ, Sommerburg O, Kauczor H-U, Schenk J-P, et al. Echo time-dependence of observed lung T1 in patients with cystic fibrosis and correlation with clinical metrics. *J Magn Reson Imaging.* (2020) 52:1645–54. doi: 10.1002/jmri.27271

18. Triphan SMF, Weinheimer O, Gutberlet M, Heußel CP, Vogel-Claussen J, Herth F, et al. Echo time-dependent observed lung T1 in patients with chronic obstructive pulmonary disease in correlation with quantitative imaging and clinical indices. *J Magn Reson Imaging.* (2021) 54:1562–71. doi: 10.1002/jmri.27746

19. Karch A, Vogelmeier C, Welte T, Bals R, Kauczor H-U, Biederer J, et al. The German COPD cohort COSYCONET: aims, methods and descriptive analysis of the study population at baseline. *Respir Med.* (2016) 114:27–37. doi: 10.1016/j.rmed.2016.03.008

20. Triphan SMF, Breuer FA, Gensler D, Kauczor H-U, Jakob PM. Oxygen enhanced lung MRI by simultaneous measurement of T1 and T2* during free breathing using ultrashort TE. *J Magn Reson Imaging.* (2015) 41:1708–14. doi: 10.1002/jmri.24692

21. Deichmann R, Haase A. Quantification of T1 values by SNAPSHOT-FLASH NMR imaging. *J Magn Reson Imaging.* (1992) 96:608–12. doi: 10.1016/0022-2364(92)90347-A

22. Kohlmann P, Strehlow J, Jobst B, Krass S, Kuhnigk J-M, Anjorin A, et al. Automatic lung segmentation method for MRI-based lung perfusion studies of patients with chronic obstructive pulmonary disease. *Int J Comput Assist Radiol Surg.* (2015) 10:403–17. doi: 10.1007/s11548-014-1090-0

23. Kohlmann P, Laue H, Krass S, Peitgen H-O (2011). "Fully-automatic determination of the arterial input function for dynamic contrast-enhanced pulmonary MR imaging" in *MIUA.*

24. Sourbron S, Dujardin M, Makkat S, Luypaert R. Pixel-by-pixel deconvolution of bolus-tracking data: optimization and implementation. *Phys Med Biol.* (2007) 52:429–47. doi: 10.1088/0031-9155/52/2/009

25. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern.* (1979) 9:62–6. doi: 10.1109/TSMC.1979.4310076

26. Avants BB, Tustison N, Song G. Advanced normalization tools (ANTS). *Insight J.* (2009) 2:1–35. doi: 10.54294/uvnhin

27. Woodhouse N, Wild JM, Paley MNJ, Fichele S, Said Z, Swift AJ, et al. Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers. *J Magn Reson Imaging.* (2005) 21:365–9. doi: 10.1002/jmri.20290

28. Kaireit TF, Voskrebenzev A, Gutberlet M, Freise J, Jobst B, Kauczor H-U, et al. Comparison of quantitative regional perfusion-weighted phase resolved functional lung (PREFUL) MRI with dynamic gadolinium-enhanced regional pulmonary perfusion MRI in COPD patients. *J Magn Reson Imaging.* (2019) 49:1122–32. doi: 10.1002/jmri.26342

29. Agustí AG, Barberá JA, Roca J, Wagner PD, Guitart R, Rodriguez-Roisín R. Hypoxic pulmonary vasoconstriction and gas exchange during exercise in chronic obstructive pulmonary disease. *Chest.* (1990) 97:268–75. doi: 10.1378/chest.97.2.268

30. Hoffman EA, Chon D. Computed tomography studies of lung ventilation and perfusion. *Proc Am Thorac Soc.* (2005) 2:492–506. doi: 10.1513/pats.200509-099DS

31. Rodríguez-Roisin R, Drakulovic M, Rodríguez DA, Roca J, Barberà JA, Wagner PD. Ventilation-perfusion imbalance and chronic obstructive pulmonary disease staging severity. *J Appl Physiol.* (2009) 106:1902–8. doi: 10.1152/japplphysiol.00085.2009

Review of quantitative and functional lung imaging evidence of vaping-related lung injury

Joseph J. Hofmann[1†], Victoria C. Poulos[1†], Jiahai Zhou[1†], Maksym Sharma[1,2], Grace Parraga[1,2,3]* and Marrissa J. McIntosh[1,2]

[1]Robarts Research Institute, London, ON, Canada, [2]Department of Medical Biophysics, London, ON, Canada, [3]Department of Medical Imaging, Western University, London, ON, Canada

Introduction: The pulmonary effects of e-cigarette use (or vaping) became a healthcare concern in 2019, following the rapid increase of e-cigarette-related or vaping-associated lung injury (EVALI) in young people, which resulted in the critical care admission of thousands of teenagers and young adults. Pulmonary functional imaging is well-positioned to provide information about the acute and chronic effects of vaping. We generated a systematic review to retrieve relevant imaging studies that describe the acute and chronic imaging findings that underly vaping-related lung structure-function abnormalities.

Methods: A systematic review was undertaken on June 13th, 2023 using PubMed to search for published manuscripts using the following criteria: [("Vaping" OR "e-cigarette" OR "EVALI") AND ("MRI" OR "CT" OR "Imaging")]. We included only studies involving human participants, vaping/e-cigarette use, and MRI, CT and/ or PET.

Results: The search identified 445 manuscripts, of which 110 (668 unique participants) specifically mentioned MRI, PET or CT imaging in cases or retrospective case series of patients who vaped. This included 105 manuscripts specific to CT (626 participants), three manuscripts which mainly used MRI (23 participants), and two manuscripts which described PET findings (20 participants). Most studies were conducted in North America ($n = 90$), with the remaining studies conducted in Europe ($n = 15$), Asia ($n = 4$) and South America ($n = 1$). The vast majority of publications described case studies ($n = 93$) and a few described larger retrospective or prospective studies ($n = 17$). In e-cigarette users and patients with EVALI, key CT findings included ground-glass opacities, consolidations and subpleural sparing, MRI revealed abnormal ventilation, perfusion and ventilation/perfusion matching, while PET showed evidence of pulmonary inflammation.

Discussion and conclusion: Pulmonary structural and functional imaging abnormalities were common in patients with EVALI and in e-cigarette users with or without respiratory symptoms, which suggests that functional MRI may be helpful in the investigation of the pulmonary health effects associated with e-cigarette use.

KEYWORDS

e-cigarettes, MRI, CT, PET, EVALI

Introduction

In the summer of 2019, public health authorities in the US issued a public health alert and launched an investigation which resulted in the reporting of a large cluster of hospitalizations in young electronic cigarette (e-cigarette) users (1). The outbreak, first identified in July 2019 (2), stemmed from the hospital admission of five previously healthy teens who were recent e-cigarette users, which was then reported by the Centers for Disease Control (CDC) in the US (3). Because of this concerted public health approach, nearly 3,000 cases of idiopathic acute lung injury were reported in the fall of 2019, a few months prior to the beginning of the coronavirus disease 2019 (COVID-19) pandemic. Serious, life-threatening lung injury requiring critical care and mechanical ventilation and, in some cases, extracorporeal membrane oxygenation (ECMO), was reported primarily in previously healthy adolescents, which lead to the deaths of at least 68 adolescents and young adults (4), the youngest of whom was 13 years of age (5). Lung injury related to e-cigarette use was contemporaneously termed EC-related or vaping product use-associated lung injury (EVALI) (6). While the number and intensity of vaping related hospitalizations diminished during the pandemic years, many unanswered questions remain about the impact of vaping on lung structure and function in combustible cigarette users and in people who had never smoked combustible cigarettes prior. Such questions included: "what was the exact compound and pathophysiologic mechanisms responsible for the e-cigarette-related acute lung injury in 2019?"; "what are the acute and chronic effects of vaping on pulmonary health?"; "with the public health alert now passed, do unreported intensive care hospital admissions still occur in young adults and adolescents?; "how do the different e-liquid components impact the lungs?"; and, finally, "how does the apparent lung damage that stems from vaping or e-cigarette use directly compare to the acute and long term effects of combustible cigarette use?"

Unfortunately, there are very few studies designed to answer these questions, and even fewer studies which provide high quality evidence based on randomized, controlled prospective research designs. This systematic review was undertaken to uncover and summarize the pulmonary imaging evidence already published and explore some of these unanswered questions.

Historical context and technical developments

Historical development of vaping-related devices

The first e-cigarette was developed and patented in 1930 (U.S. Patent No. 1775947A) for individual use and a similar device was patented in 1963 (U.S. Patent No. 3200819A); neither of these devices were manufactured for commercialization. In 2003, a Chinese pharmacist, Hon Lik, initially developed the e-cigarette as an alternative to combustible cigarettes. It was commercialized in 2004 in Canada and China and was on the market in Europe and the United States in 2006 (7).

As shown in Figure 1, modern e-cigarette devices consist of a battery, mouthpiece, heating element, liquid solution reservoir and

FIGURE 1
Typical commercial electronic cigarette or vape device.
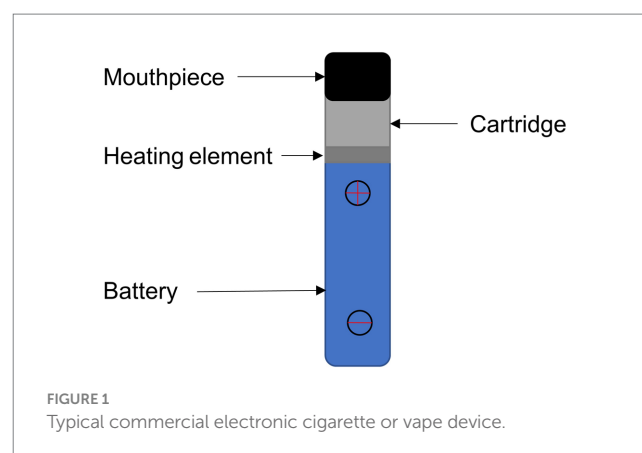
disposable cartridge or pod. The battery powers a heating element which, when applied to the liquid solution, rapidly increases the temperature enabling the transition of the e-liquid to a gaseous or aerosol state, which is subsequently inhaled via a mouthpiece into the lungs.

E-cigarette versus combustible cigarette risk

The long-term use of combustible cigarettes can cause various cardiopulmonary health risks such as chronic obstructive pulmonary disease, pulmonary hypertension and cancer. Despite some national advisory committees suggesting that the relative health risks of e-cigarette use are reduced as compared to combustible cigarettes (8), very few head-to-head comparison studies have been completed. Previous investigations suggest that e-cigarette use, either alone or in combination with combustible cigarettes, is associated with reduced overall health, breathing difficulties, and cardiovascular abnormalities (9, 10). Several systematic reviews have been conducted to examine the impact of e-cigarette use on respiratory illness (11) and human health (12); here, we summarize pulmonary imaging findings in e-cigarette users.

Composition of vaping e-liquids

e-liquids contain flavors, solvent carriers and active ingredients such as tetrahydrocannabinol (THC) or nicotine. The solvent carriers typically consist of propylene glycol, ethylene glycol, glycerol, tobacco-specific nitrosamines, volatile organic compounds, phenolic compounds, tobacco alkaloids, aldehydes, free radicals, reactive oxygen species, furans and metals (i.e., nickel, lead, chromium). Flavoring substances include menthol, ethyl maltol and diacetyl which are present in the most popular flavors of e-cigarettes (13), as shown in Figure 2.

Flavors and excipients

Table 1 lists common excipients and flavors present in e-liquids. Active ingredients and concentrated flavoring compounds are

FIGURE 2
Distribution of vape flavors by preference. Figure was created using data from Canadian Tobacco and Nicotine Survey (CTNS) (14).
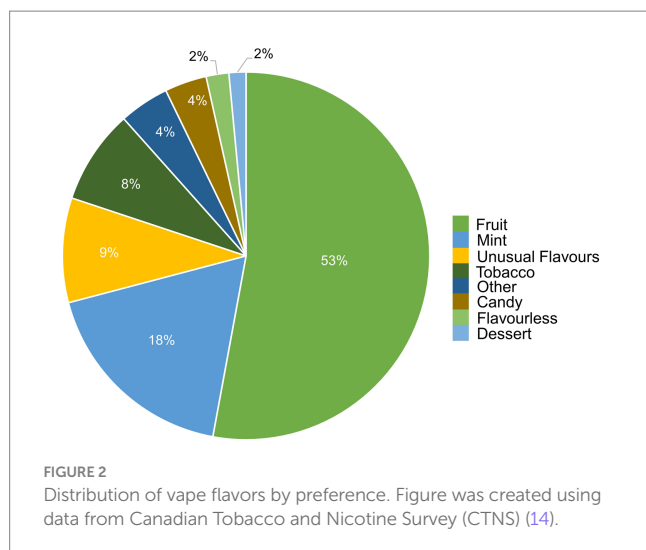
dissolved in inactive excipients, which are then delivered to users as an aerosol.

There are more than 7,000 vape liquid flavors and over 450 brands available on the market (27). Though many flavorings fall under the "generally recognized as safe" (GRAS) provision by the US Food and Drug Administration (FDA), the GRAS status only applies to the use of such flavorings in ingested foods and not for inhaled products (28). In fact, at least 65 individual flavoring ingredients in flavored e-liquids were observed to cause toxicity in the respiratory track by inducing cytotoxicity, generating reactive oxygen species and impairing clearance mechanisms (21). Furthermore, cinnamaldehyde, vanillin, menthol, ethyl vanillin, benzaldehyde, ethyl maltol and linalool were present in the flavors that caused the most toxicity in *in vitro* studies (21). Unfortunately, the variety of products available as well as the ongoing modifications of e-cigarette and vaping devices makes it difficult to comprehensively evaluate the biological risk of vaping and specific e-liquids (29).

As early as 2000 (15, 30), the excipient diacetyl, which provides a buttery flavor in e-liquids (31), was associated with the development of bronchiolitis obliterans or "popcorn lung" in workers at a microwave popcorn plant; inflammation and fibrosis in the lung can result in the partial or complete obstruction of the peripheral airway lumen, thus leading to bronchiolitis obliterans.

Previous work has demonstrated that propylene glycol, which acts as an emulsifier for active ingredients in e-liquids (32), may also damage peripheral airways by harming epithelial cells and reducing cell proliferation; e-cigarette users with underlying chronic obstructive pulmonary disease (COPD) were more susceptible to small airway epithelial damage than those without (33).

Vegetable glycerin serves as a delivery vehicle for the active contents in e-liquids (34). Vegetable glycerin e-cigarette aerosols disturb the human nasal cystic fibrosis transmembrane conductance regulator, resulting in consequent mucus hyperconcentration and potentially harming the airway by inducing inflammation and ion channel dysfunction (34).

Vitamin E acetate is commonly used as a thickening agent in vaping products that contain cannabis derivatives (31). The inhalation of vitamin E acetate may lead to impaired pulmonary function (35–37). Furthermore, when heated, vitamin E acetate thermally degrades

into ketene, alkene and benzene, all of which may contribute to epithelial lung injury (18). Additional studies observed that vitamin E acetate alters surfactant expansion and compression cycles, potentially compromising surfactant function (38). With compromised lung surfactant, the alveolar surface tension would increase and may cause an inflammatory cascade in lung tissue. The CDC previously observed vitamin E acetate in the bronchoalveolar lavage fluid of 94% of EVALI patients (19, 39), pointing to an association between vitamin E acetate and the development of EVALI.

Active ingredients

Commonly used active ingredients in e-liquids include THC and nicotine. THC is the psychoactive compound in cannabis and may be added to e-liquids (21). While THC can be degraded to toxins such as methacrolein and benzene when heated in vapes (18), other compounds in THC-based e-liquids may also have negative effects on respiratory health. THC-based e-liquids differ from nicotine-based liquids because THC (a highly hydrophobic chemical) requires a hydrophobic emulsifier to be incorporated into the vape liquid (18). Thus, vitamin E acetate, described in the previous section, is commonly used in THC-based e-liquids. While the vast majority of EVALI cases (80%–86%) are associated with vaping THC products, the overall health risks of cannabis vaping is largely unknown (18, 29, 39, 40).

e-cigarettes are marketed as a safe alternative to traditional combustible cigarettes as they do not contain carcinogenic incomplete combustion byproducts nor tobacco nitrosamines (41). However, previous investigations have demonstrated an association between nicotine and vaping-induced COPD pathologies (42, 43), such as emphysema, as well as an increased risk of lung cancer among nicotine e-cigarette users (44). Although vaping does not require tobacco or combustion, stimulant nicotine in e-liquids may still cause DNA damage (45) and may mediate tumor growth by promoting the self-renewal of stem-like cells in tumor initiation and metastasis (46).

Pathophysiology of vaping-induced lung injury
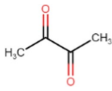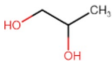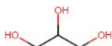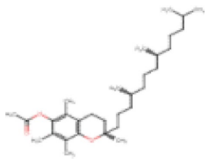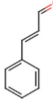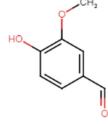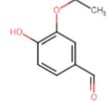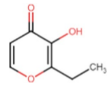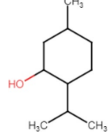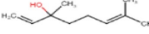
Symptoms

The majority of EVALI patients experience respiratory symptoms, including cough, chest pain, shortness of breath and hemoptysis (2, 39, 47), in addition to gastrointestinal symptoms, such as abdominal pain, nausea, vomiting and diarrhea; gastrointestinal symptoms may precede respiratory symptoms in some patients (47). Additional common symptoms include unexplained weight loss, headache and fatigue (2), and, more generally, EVALI patients are frequently admitted with tachycardia, tachypnea, fever and hypoxemia (39). Because lung biopsies not commonly obtained in EVALI patients, the pathological drivers of symptoms and disease progression remain poorly understood.

The respiratory system

The respiratory system consists of two main parts: the upper and lower respiratory tract. The upper tract includes the nose,

TABLE 1 Common excipients and flavors in e-liquids.

Chemical	Molecular structure	Function	Impact on lung health
Excipients			
Diacetyl		• Buttery flavor	• Associated with bronchiolitis obliterans (15)
Propylene glycol		• Emulsifies ingredients to increase solidity	• Generates pulmonary irritants and carcinogenic carbonyl compounds (16)
Vegetable glycerin (glycerol/glycerin)		• Delivery vehicle	• Increases mucin expression primary airway epithelia (17)
Vitamin E acetate (alpha-tocopherol acetate)		• Thicken or dilute e-liquids with THC derivatives	• Alters surfactant expansion and compression cycles (18) • Linked to EVALI (19)
Flavors			
Cinnamaldehyde		• Cinnamon	• Suppression of macrophage phagocytosis (20) • Alters cell morphology and motility (21) • DNA strand breakage due to oxidative burst (21)
Vanillin		• Vanilla	• Airway epithelial cell metabolic disruption (22) • Metabolic effect amino acids, fatty acids, lipids and mitochondrial function (22)
Ethyl vanillin		• Vanilla	• Decreased neutrophil oxidative burst (23)
Ethyl maltol		• Caramel	• Co-exposure with copper causes epithelial cell apoptosis and DNA damage (24)
Menthol		• Mint	• Decreases cell proliferation (25) • Increased oxidative stress (25) • Damages respiratory epithelium (25)
Benzaldehyde		• Cherry-almond	• Attenuates oxidative burst capacity of neutrophil (21)
Linalool		• Floral, sweet	• Increases cytotoxicity (21)

Molecular structure representations were generated with Marvin SJ (26). EVALI, e-cigarette-related or vaping-associated lung injury.

nasal cavity, throat and larynx. This tract is responsible for bringing in air from outside the body, through the nose and mouth.

As shown in Figure 3, the lower respiratory tract is divided into two zones, known as the conducting and respiratory zones. These zones encompass a total of 23 generations of airways. The conducting
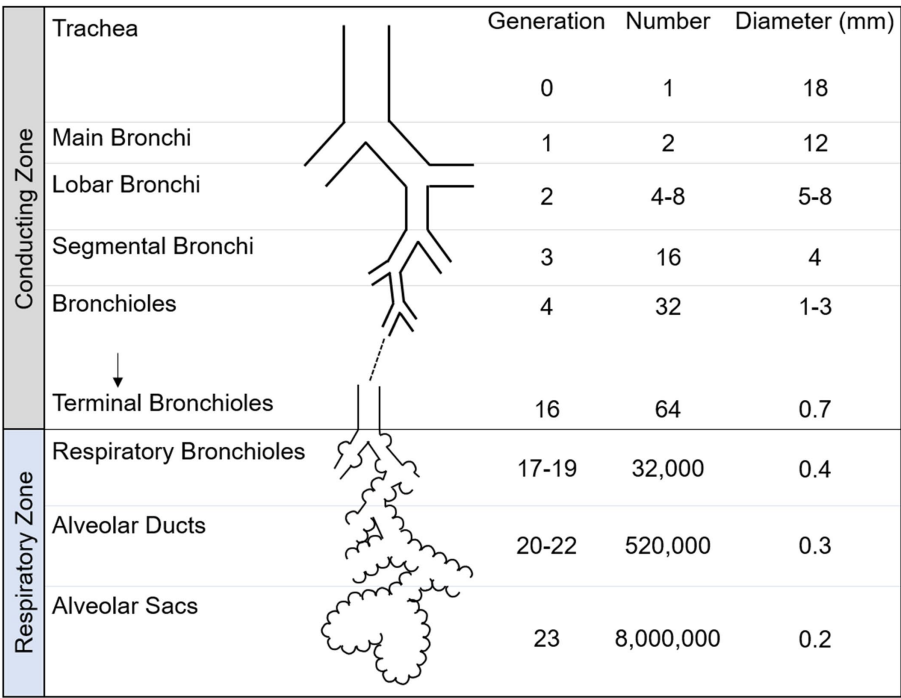
FIGURE 3
Schematic of airway tree conducting and respiratory zone. The human airway tree consists of the conducting zone and respiratory zone, with corresponding generation, number and diameter shown. Adapted from Nunn's Applied Respiratory Physiology, 8th edition.

zone, which includes airway generations 0–16, is responsible for guiding inhaled air towards the alveoli and humidifying it. The conducting zone begins with the trachea which then bifurcates across 16 generations of airways to the terminal bronchioles, which conclude the conducting zone. As the airways progress into further generations, their diameter gradually decreases. The airways within the conducting zone consist of thick walls of mucosa, smooth muscle and cartilage and are lined with cilia to remove dust and foreign particles from the lung.

The respiratory zone is responsible for facilitating gas exchange and contains airway generations 17–23. The respiratory zone contains respiratory bronchioles, alveolar ducts and alveolar sacs. Within the lungs, there are numerous microscopic sacs known as alveoli, which begin to bud along the walls of the respiratory bronchioles and become increasingly abundant with each subsequent airway generation. These alveoli are surrounded by a network of capillaries, whose primary function is to facilitate the exchange of inhaled oxygen with carbon dioxide. The alveolar sacs complete the respiratory zone and the airway tree.

Pathophysiology

Potential mechanisms of vaping-related lung injury are provided in schematic in Figure 4. When aerosolized e-liquids are inhaled, the particles come in direct contact with the entire respiratory system. Consequently, chemicals in e-cigarette aerosols, such as menthol and ethyl maltol may lead to inflammation, which is thought to be the dominant cause or contribution to vaping-related lung injury (18, 31, 48–50). It has also been postulated that the pro-inflammatory effects caused by vape aerosols are partially mediated by reactive oxygen species (ROS) (27), which may lead to cellular apoptosis through

ROS-mediated autophagy (51); this cellular death pathway is also mechanistically attributed to emphysema (27). In addition, pulmonary inflammation and fibrosis may result in bronchiolitis obliterans, which is characterized by hypertrophy of the bronchiolar smooth muscle, peribronchiolar inflammatory infiltrates, mucus accumulation in the bronchiolar lumen and bronchiolar scarring (52). This scarring is irreversible and bronchiolitis obliterans has no known cure. While healthy patients may achieve complete improvement, the pulmonary health of most patients progressively worsens and some may even require mechanical ventilation or lung transplants in severe cases (53).

Heavy metals have been detected in the vape aerosols produced from pod-type vapes (27), including chromium, nickel, copper, zinc, cadmium, tin, manganese and lead (54). The metallic components of e-cigarette devices, such as the filaments and coils, are comprised of such heavy metals and can degrade when exposed to oxidized acidic e-liquids (54). While metal exposure is a risk factor for multiple pulmonary diseases including respiratory inflammation, asthma, COPD and respiratory cancer (55–57), the health effects of metal exposure in e-cigarette users is still largely unknown (58).

Patients with EVALI or vaping-related lung injury consistently present with lipid laden alveolar macrophages or foam cells in the lung or in bronchoalveolar lavage fluid (59). These macrophages serve as the main phagocytes in the innate immune system, clearing the airspaces of potentially harmful particles (48). Recent work has observed lipid laden alveolar macrophages in the bronchoalveolar lavage fluid of 80% of EVALI patients, demonstrating its potential as a biomarker of vaping-related lung injury. Nevertheless, the presence of lipid laden alveolar macrophages is not unique to vaping, which has been reported in a number of other pulmonary conditions (18). Thus, they are a non-specific marker of vape product use, and do not provide a direct prognosis of vaping-related lung injury (60).

FIGURE 4
Potential mechanisms of pulmonary injury associated with vaping.
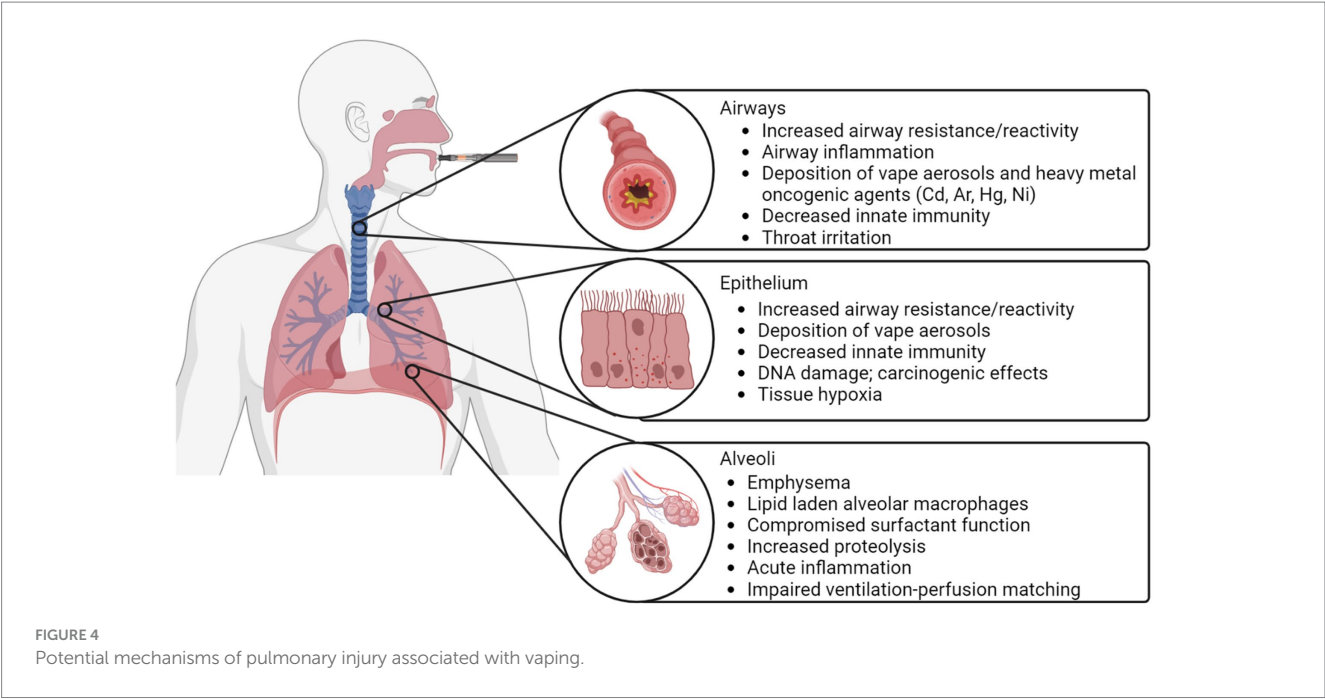
TABLE 2 Inclusion and exclusion criteria for systematic review.

Parameter	Inclusion	Exclusion
Imaging modalities	MRI, CT, PET	Other
Body system	Respiratory	Non-respiratory
Test subjects	Human	Non-human
Type of vape	Electronic cigarettes	Combustible cigarettes (primary use)
Type of article	Case study, case series, retrospective studies, prospective studies	Review articles, others

MRI, magnetic resonance imaging; CT, computed tomography; PET, positron emission tomography.

Knowledge and health care gaps

There are currently many unanswered questions associated with vaping. An unexhausted list of these questions includes: "what are the longitudinal effects of vaping on pulmonary health?"; "what is the safety profile of e-cigarette use relative combustible cigarettes?"; "what are the effects of each e-liquid component on respiratory health?"; and "what role can pulmonary imaging play in the diagnosis and monitoring of vaping-related lung injury?"

This review is designed to explore some of these unanswered questions in the context of pulmonary imaging.

Research questions

We aim to investigate the current understanding of vaping in literature, the conditions associated with it, and the common imaging modalities used to evaluate vaping-related lung injury, specifically magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET).

Methods

This review was conducted according to the preferred reporting items for systematic reviews and meta-analyses (PRISMA). We used PubMed to search for manuscripts related to vaping/e-cigarette use or EVALI and pulmonary imaging on June 13, 2023 using the terms [("Vaping" OR "e-cigarette" OR "EVALI") AND ("MRI" OR "CT" OR "Imaging")].

Inclusion and exclusion criteria

The inclusion and exclusion criteria are presented in Table 2. The inclusion criteria were: (a) imaging modalities: strictly MRI, CT and PET; (b) bodily systems scanned; strictly structures in the respiratory system: such as lungs, respiratory epithelium, and alveoli; (c) test subjects: living human subjects who had previously used e-cigarettes/ vapes. Exclusion criteria included: (1) any types of non-electric cigarette such as conventional cigarettes, joints, cigars, and hookahs; (2) any articles that were reviews or not directly treating a single/ group of patients.

Selection process and data collection

Four reviewers (JH, VP, JZ, and GP) independently screened the abstracts of the retrieved reports to evaluate whether they met the predetermined inclusion/exclusion criteria. A full text review of the remaining studies was performed to further evaluate the inclusion eligibility of the reports and any discrepancies were presented and

discussed between the four reviewers to ensure a consensus was made regarding the inclusion of each report.

Risk of bias assessment

Three reviewers (JH, VP, and JZ) independently appraised the methodological qualities of the included studies in accordance with two risk of bias tools, detailed further in Supplementary Figure S1. The JBI critical appraisal checklist for case reports was deemed most appropriate for one report [Eddy et al. (61)] while the Cochran risk of bias in non-randomized studies-of exposure tool (ROBINS-E) was used for the remaining studies. The JBI tool assesses bias arising from: the patient's demographics (D1), history (D2) and clinical condition on presentation (D3), the diagnostic tests or assessment methods (D4), the interventions or treatments used (D5), the post-intervention clinical condition (D6), the adverse effects on unanticipated events (D7), and the takeaway lessons (D8). The ROBINS-E tool assesses bias arising from: confounds (D1), measurement of the exposure (D2), selection of the participants in the study (D3), post-exposure interventions (D4), missing data (D5), measurement of the outcome (D6), and selection of the reported result (D7). The level of risk judgement for each domain was categorized into four categories: high risk, some concerns, low risk, or no information.Synthesis methods

The criteria used to determine which studies were eligible for each synthesis was that, for each patient, the study must report age, gender, location, and history of e-cigarette or vape use. We used a table to separate patient characteristics such as age, type of e-device vaped and symptoms. No data conversions were necessary, as we qualitatively analysed patient symptoms and characteristics. The method used to prepare missing summary statistics of age range was to report the mean age.

Evaluation methods

Computed tomography

Chest CT measures x-ray attenuation coefficients to determine lung tissue density and is routinely used in clinic to evaluate a variety of respiratory diseases, including but not limited to pneumonia, emphysema and interstitial lung disease. CT may be used to discern between pulmonary diseases and evaluate severity, as well as monitor disease progression and treatment response.

Positron emission tomography

PET is a quantitative molecular imaging modality that utilizes radiotracers to measure pulmonary ventilation, perfusion and blood flow as well as metabolic activity. The radiotracers used in PET imaging emit positrons, which will typically travel a few millimeters before colliding with an electron in the surrounding tissue, resulting in the emission of two $511\,keV$ photons in opposite directions that are then detected by PET detectors oriented around the patient. In the context of lung disease, PET imaging is mostly commonly used to detect cancerous cells. However, PET radiotracers may also be used to evaluate the deposition of inhaled substances using carbon-11 (^{11}C) and fluorine-18 (^{18}F) (62), inflammation using ^{18}F-fluorodeoxyglucose (^{18}F-FDG) (63) and ventilation using nitrogen-13 (^{13}N$_2$) (63).

Pulmonary functional MRI

Hyperpolarized noble gas, oxygen-enhanced ^{1}H, free-breathing ^{1}H and arterial spin labelling methods are employed in order to acquire pulmonary functional MR images. These MRI techniques, in combination, may be used to measure lung ventilation and/or perfusion (64, 65), and are sensitive to early disease changes not easily detected with conventional pulmonary function tests (64, 66).

Hyperpolarized noble gas MRI allows for the visualization of the gas distribution *in vivo*, and is sensitive to the transitional and respiratory zones of the lung (67). Hyperpolarized ^{3}He or ^{129}Xe act as inhaled contrast agents, providing three-dimensional images with increased spatial and temporal resolution as compared to other functional imaging methods such as PET (68, 69), ^{133}Xe scintigraphy (70), Xe-enhanced CT (71) and single photon emission computed tomography (SPECT) (68, 69). Ventilation abnormalities depicted on hyperpolarized MR images are associated with pulmonary structural abnormalities including luminal plugging, air trapping, airway inflammation and emphysema (68, 72–76). Hyperpolarized gas MRI has mainly been used as a research tool to investigate respiratory diseases including COPD (77), asthma (67), cystic fibrosis (78) and idiopathic pulmonary fibrosis (79). Recently, hyperpolarized ^{129}Xe was approved by the US FDA for clinical use (80).

Oxygen-enhanced ^{1}H MRI is a research technique that relies on the weak paramagnetic properties in oxygen, which decreases the longitudinal relaxation time (T1) of tissues, thus increasing the signal (81). By using O_2 as a contrast agent, the presence and/or absence of ventilation may be evaluated and specific ventilation may be quantified (81). Oxygen-enhanced MRI has previously been used to investigate patients with COPD (82), interstitial lung disease (83) and asthma (84).

Free-breathing ^{1}H MRI is also mainly a research tool which can be used to evaluate pulmonary ventilation and perfusion without the need for an endogenous contrast agent. This technique has been applied to patients with various cardiopulmonary diseases including COPD, CTEPH and asthma (85–88). These images are post-processed using non-rigid registration which compensates for respiratory motion, eliminating the need for patient compliance during a breathhold. Commonly used image processing techniques for free-breathing MRI include phase-resolved functional lung (PREFUL) (88) and matrix pencil decomposition (89).

Arterial spin labelling MRI can be used to evaluate pulmonary blood flow and its heterogeneity. This technique involves the acquisition of two images, which are subsequently subtracted to remove the blood signal from the tissue (90). This results in a perfusion-weighted image of the blood delivered within one cardiac cycle.

Results

Participants and report characteristics

A flowchart is provided in Figure 5; Supplementary Figure S2 which show that the PubMed search conducted on June 132,023 yielded 445 manuscripts, one of which was identified as a duplicate and was excluded. The abstracts of the remaining 444 manuscripts were screened and an additional 235 manuscripts were excluded. Following a thorough full text evaluation, 45 review articles, 27

FIGURE 5
Flow chart of study search and selection process.
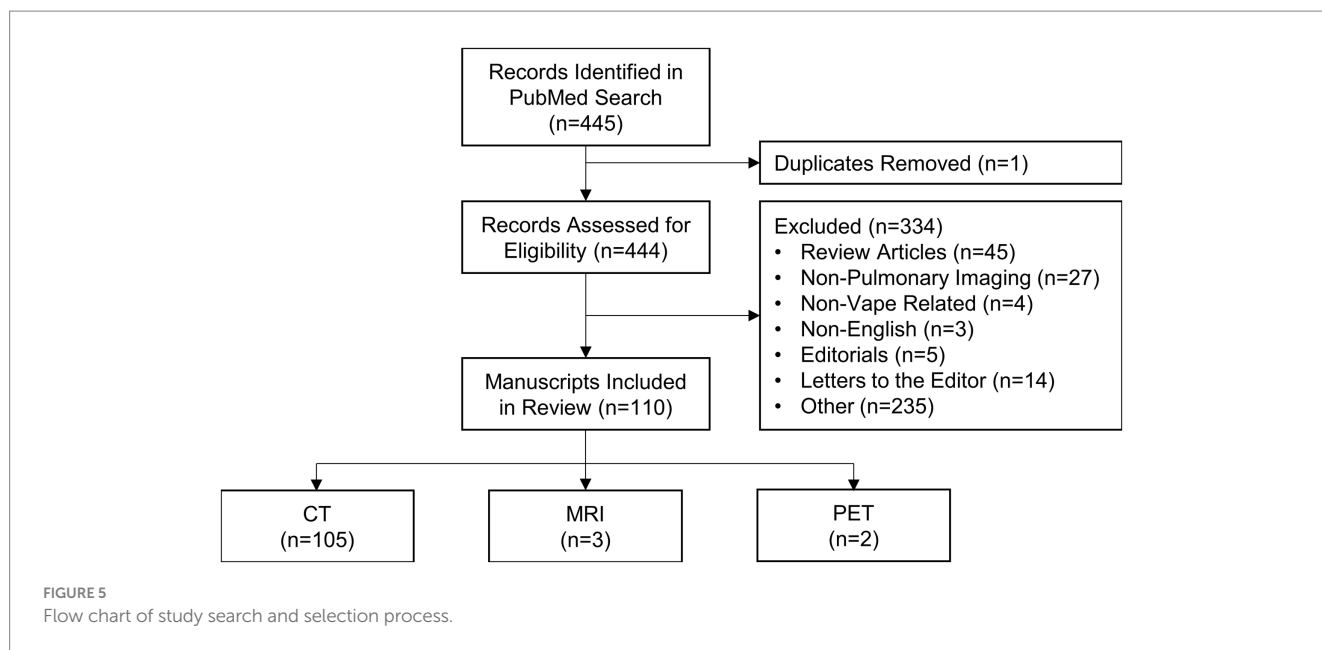
TABLE 3 Active ingredients in e-liquid reported in included studies.

Vape component n (%)	Manuscripts n = 110	Participants n = 668[a]
Nicotine	35 (32)	101 (15)
THC	46 (42)	247 (37)
Dual		
Nicotine + THC	39 (35)	230 (34)
Nicotine + other	3 (3)	3 (0.4)
Triple		
Nicotine + THC + other	3 (3)	10 (1)
Unspecified	29 (26)	71 (11)
Other	4 (4)	6 (1)

THC, Δ⁹-tetrahydrocannabinol. [a]A single participant was described in two separate case studies.

non-pulmonary articles, 15 letters to the editor, five editorials, four non-vape related and three non-English articles were excluded. Ultimately, 110 manuscripts met the inclusion criteria and were included in this review.

The characteristics of the included studies are listed in Supplementary Tables S1, S2. The vast majority of studies were conducted in North America ($n = 90$) with the remaining studies having been conducted in Europe ($n = 15$), Asia ($n = 4$) and South America ($n = 1$). A total of 668 e-cigarette users or patients that were exposed to vaping performed MRI, CT and/or PET scans; the primary imaging modality for 625 patients was CT, for 20 patients was PET and for 22 patients was MRI; a single participant was presented in two separate manuscripts, where one report described the clinical case and CT findings (91) and the other described research MRI and quantitative CT findings (61). In addition, each PET scan was accompanied by an anatomical CT. The minimum number of patients included in each study was 1 and the maximum was 160. Nearly all

($n = 109$) studies evaluated current or former e-cigarette users; one study (92) prospectively evaluated the acute effects of nicotine-based vaping in 15 healthy adults. The reported active ingredients in the e-liquid used by patients is summarized in Table 3. A total of 354 participants exclusively used nicotine ($n = 101$), THC ($n = 247$) or other substances ($n = 6$) in vape devices, 233 participants used both nicotine and either THC ($n = 230$) or containing other substances ($n = 3$) and 10 participants used a combination of three substances, while the contents of the vape devices of the remaining 71 participants was unspecified or unknown.

CT findings

Of the included manuscripts, CT was the most common imaging modality used to evaluate pulmonary abnormalities in vaping-related lung injury (105/110 manuscripts; 95%). Table 4; Supplementary Table S1 provide summaries of the findings from the included CT manuscripts. The majority of manuscripts reported on patients that were adolescents or young adults (61/105; 58%), with patient age ranging from 13 to 68 years old. Of the CT manuscripts included, 92 (88%) were case studies and 13 (12%) were retrospective studies. Acute lung injury was reported in 97 (92%) studies and 600 (98%) patients, chronic lung injury was reported in 7 (7%) studies and 13 (2%) patients, and both were reported in 2 (2%) studies. Nineteen (18%) studies involved participants who had previous respiratory diseases including: asthma ($n = 15$), tuberculosis ($n = 1$), allergic rhinitis ($n = 1$) and COPD ($n = 1$).

Ground-glass opacities, defined as a hazy increase in lung density observed on chest CT (93), were the most common CT finding in the included manuscripts, with 84 (80%) studies and 347 (57%) participants reporting evidence of this pulmonary abnormality. Consolidation opacities (37 studies/176 participants), subpleural sparing (32 studies/230 participants), septal thickening (24 studies/145 participants) and patchy opacities (22 studies/44 participants) were also common.

TABLE 4 CT findings.

CT study type & findings	Number of studies N = 105	Number of participants N = 612
Study types		
Cases	92 (88)	201 (33)
Retrospective	13 (12)	411 (67)
Lung injury type		
Acute lung injury	97 (92)	600 (98)
Chronic lung injury	7 (7)	12 (2)
Acute and chronic lung injury	1 (1)	24
Major CT findings		
Ground-glass opacities	84 (80)	347 (57)
Consolidation opacities	37 (35)	176 (29)
Subpleural sparing	32 (30)	230 (38)
Septal thickening	24 (23)	145 (24)
Patchy opacities	22 (21)	44 (8)
Pleural effusion	22 (21)	88 (14)
Pneumomediastinum	14 (13)	28 (5)
Lymphadenopathy	12 (11)	135 (22)
Bronchial wall thickening	10 (10)	45 (7)
Pneumothorax	10 (10)	12 (2)
Centrilobular nodules	9 (9)	72 (12)
Organizing pneumonia	9 (9)	42 (7)
Diffuse pulmonary nodules	7 (7)	8 (1)
Crazy paving	7 (7)	43 (7)
Pulmonary infiltrates	7 (7)	10 (2)
Tree in bud	5 (5)	7 (1)
Reverse halo sign	5 (5)	16 (3)
Bronchiolitis	5 (5)	9 (2)
Unspecified opacities	4 (4)	94 (15)
Emphysema	4 (4)	18 (3)
Parenchymal sparing	3 (3)	4 (1)
Peripheral sparing	3 (3)	3 (0.5)
Bullae	3 (3)	3 (0.5)
Pericardial effusions	3 (3)	5 (1)
Pneumonitis	3 (3)	7 (1)
Peribronchovascular sparing	3 (3)	3 (0.5)
Interlobular thickening	3 (3)	7 (1)
Uncommon CT findings		
Parenchymal opacities	2 (2)	25 (4)
Interstitial opacities	2 (2)	7 (1)
Honey combing	2 (2)	3 (0.5)
Prominent mosaicism	2 (2)	5 (1)
Bronchiectasis	2 (2)	4 (1)
Solid nodules	2 (2)	3 (0.5)
Bilateral patchy Infiltrates	2 (2)	92 (15)
Subpleural cysts	2 (2)	2 (0.3)
Interstitial thickening	2 (2)	2 (0.3)
Mosaic attenuations	2 (2)	5 (1)
Bronchocentric opacities	1 (1)	1 (0.2)
Bronchiolar dilation	1 (1)	9 (2)
Fissural displacement	1 (1)	6 (1)
Pulmonary embolism	1 (1)	1 (0.2)
Alveolar infiltrates	1 (1)	1 (0.2)
Necrotizing pneumonia	1 (1)	1 (0.2)
Tracheomalacia	1 (1)	1 (0.2)
Centrolobular thickening	1 (1)	1 (0.2)
Peripheral opacities	1 (1)	5 (1)
Peribronchovascular opacities	1 (1)	6 (1)
Miliary pattern	1 (1)	1 (0.2)
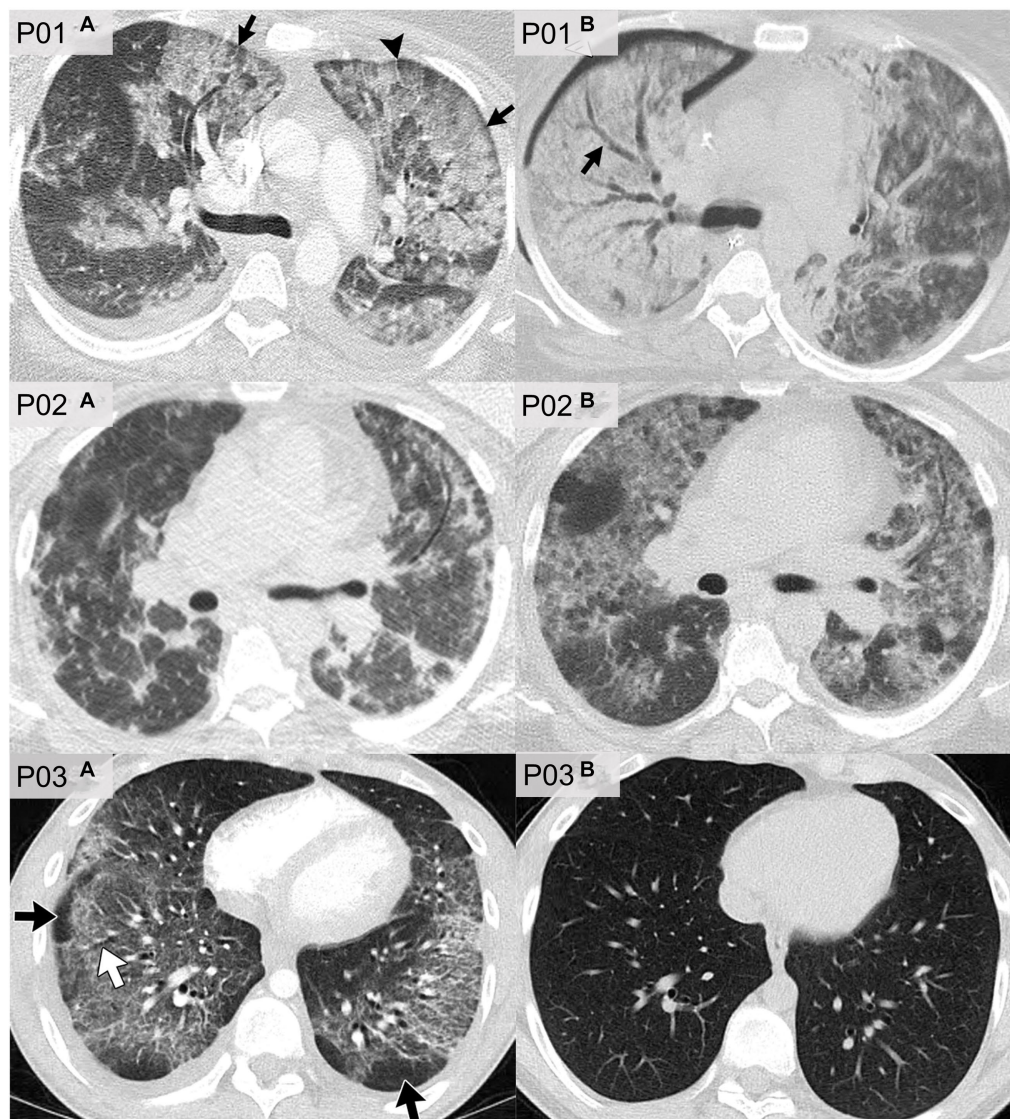
FIGURE 6
CT imaging of e-cigarette-related or vaping-associated lung injury (EVALI) with short-term follow-up. P01: axial CT of a 35 years-old female with diffuse alveolar damage pattern who vaped THC. **(A)** CT imaging showed ground-glass opacities with areas of consolidation, subpleural and perilobular sparing (arrows) and septal thickening (arrowhead). **(B)** CT 2 weeks later showed extensive right lung consolidation with areas of bronchial dilation (arrow) and internal development of right pneumothorax. Patient died 5 days later. P02: axial CT of a 51 years-old female showing multiple episodes of EVALI following repeated vaping of nicotine with mint flavoring. **(A)** CT imaging showed scattered areas of ground-glass opacities with subpleural sparing. **(B)** Two months later, the patient returned to emergency department with dyspnea and fever. CT findings included more extensive ground-glass opacities with areas of lobular and subpleural sparing. Septal thickening is present creating "crazy paving" pattern. Patient's condition deteriorated, and was complicated by aspiration pneumonia and bilateral lower-lobe collapse. P03: axial CT of organizing pneumonia pattern in a 20 years-old male who vapes nicotine and THC products daily. **(A)** CT imaging showed peribronchiolar ground-glass opacities with subpleural sparing (black arrows). Areas of bronchial dilation are seen in areas of ground-glass opacities (white arrow). **(B)** Four weeks following steroid therapy, the patient's CT scan was normal. Images reproduced with permission from Kligerman et al. (40).

Figure 6 shows axial CT slices of three patients, each at two separate time-points, previously published and described by Kligerman et al. (94). P01 was a 35 years-old female who reported the use of THC-based e-cigarettes and who presented with CT evidence of ground-glass opacities and consolidations as well as subpleural and perilobular sparing (P01a). CT images that were acquired 2 weeks later (P01b) showed extensive consolidation along with areas of bronchial dilation and the development of a right pneumothorax; this patient died 5 days later. P02 was a 51 years-old female who reported the use of nicotine-based e-cigarettes and who presented with CT evidence of

ground-glass opacities and subpleural sparing (P02a). Two months later (P02b), CT images revealed more extensive ground-glass opacities and septal thickening which presented as "crazy paving" pattern. P03 was a 20 years-old male who reported the use of both THC- and nicotine-based e-cigarettes. CT imaging (P03a) in this patient revealed organizing pneumonia, peribronchiolar ground-glass opacities and subpleural sparing. CT imaging 4 weeks later (P03b) normalized, after the patient was treated with steroid therapy.

Table 5 summarizes CT findings from the 13 included retrospective studies. All patients described in these studies presented

TABLE 5 Retrospective CT manuscripts summary.

Author	Objective	Location	Summary
Aberegg et al. (95)	Describe findings and outcomes of EVALI	Salt Lake City, United States	• Patients commonly presented with organizing pneumonia pattern on CT • Radiographic opacities resolved within 30 days
Artunduaga et al. (96)	Evaluate chest radiographic and CT findings of EVALI	Dallas, United States	• EVALI characterized by bilateral ground-glass opacities, consolidation on CT
Carroll et al. (97)	Evaluate short-term outcomes in EVALI	Milwaukee, United States	• Ground-glass and patchy opacities common
Chidambaram et al. (98)	Present clinical and imaging findings in adolescents with respiratory symptoms	Philadelphia, United States	• Imaging findings include ground-glass opacities, subpleural sparing and basilar opacities
Harry-Hernandez et al. (99)	Define pathologic findings in patients with EVALI	Multicenter—United States	• Pathologic BAL and biopsy findings correlated with CT findings
Helfgott et al. (100)	Report on radiological findings in adolescents with EVALI and COVID-19 symptoms	New Brunswick, United States	• CT findings of EVALI similar to COVID-19
Kalininskiy et al. (101)	Summarize clinical presentation of patients with probable or confirmed EVALI	Rochester, United States	• CT findings were resolved at follow-up
Kligerman et al. (94)	Describe frequency of imaging findings in EVALI	Multicenter—United States	• Ground-glass opacities, and subpleural, lobular and peribronchovascular sparing were common • Increased vaping frequency associated with more severe injury
Layden et al. (102)	Summarize clinical characteristics of EVALI patients	Wisconsin and Illinois, United States	• Bilateral infiltrates on imaging in all patients • Ground-glass opacities and subpleural sparing were common
Pajak et al. (103)	Describe imaging findings in EVALI	Delaware, United States	• Subpleural sparing, opacities and consolidation were common
Panse et al. (104)	Describe CT patterns in EVALI	Arizona, United States	• EVALI acutely presents as ground-glass opacity and consolidation • Longitudinal CT pattern resembles subacute hypersensitivity pneumonitis
Rao et al. (105)	Describe diagnosis, evaluation and management of EVALI in adolescents	Dallas, United States	• All patients demonstrated bilateral ground-glass opacities on CT
Wang et al. (106)	Describe most common CT findings in EVALI	Houston, United States	• Opacities with subpleural and peribronchovascular sparing commonly observed

EVALI, e-cigarette-related or vaping-associated lung injury; CT, computed tomography; BAL, bronchoalveolar lavage.

with probable or confirmed EVALI. Six of these studies reported on pediatric patients and four reported findings in adult patients, while the remaining three described findings in both pediatric and adult populations.

Nearly all (10/13) of the retrospective studies reported that ground-glass opacities, subpleural sparing and/or consolidation were among the most common CT findings in patients with EVALI (94–98, 102–106).

Three investigations included short-term (within 30 days of initial presentation) follow-up imaging (95, 101, 104). Kalininskiy et al. (101) reported that CT abnormalities in patients treated with antibiotics were largely resolved within 17 days of their initial emergency department visit; Aberegg et al. (95) also reported resolution of CT opacities after treatment with antibiotics and/or corticosteroids. CT abnormalities cleared in some patients described by Panse et al. (104) despite extensive injury at initial

TABLE 6 Manuscripts with PET and MRI endpoints.

Author	Objective	Location	Summary
Eddy et al. (61)	Evaluate teenage male post ECMO for EVALI	London, Canada	• Persistent, chronic, irreversible airflow limitations and gas trapping requiring mechanical ventilation and ECMO • Continuous abnormal ventilation pattern dissimilar to ventilation heterogeneity observed in asthma or COPD
Kizhakke Puliyakote et al. (107)	Assess ventilation-perfusion mismatch in asymptomatic e-cigarette users using MRI to determine a correlation	San Diego, United States	• Impaired ventilation-perfusion matching in the lung through an alteration of both ventilation and perfusion • Degree of disruption after a session matches that of patients with COPD
Nyilas et al. (108)	To examine the immediate effect of ENDs exposure and tobacco smoke on lung ventilation and perfusion by functional MRI and lung function tests	Bern, Switzerland	• Local perfusion increased in participants who used ENDS after exposure • No change in perfusion was detected in the group of participants who used nicotine-free e-liquids • No change in lung function compared to baseline was observed (nicotine and non-nicotine) • Ventilation perfusion mismatch in ENDS users
Wall et al. (92)	Investigate the distribution and deposition of inhaled [^{11}C] nicotine using freebase nicotine and lactate salt nicotine	Uppsala, Sweden	• Freebase nicotine exhibited higher uptake and deposition in the upper respiratory pathways • Lactate nicotine showed lower tracer uptake and accumulation in the upper respiratory pathways and an earlier peak and a steeper decline in the lung
Wetherill et al. (109)	Used PET to quantify inducible nitric oxide synthase expression to characterize oxidative stress and inflammation in the lungs *in vivo*	Pennsylvania, United States	• ENDS users showed greater ^{18}F-NOS non-displaceable binding potential (BP_{ND}) than cigarette smokers ($p = 0.03$) and controls ($p = 0.01$) • ^{18}F-NOS lung tissue delivery and inducible nitric oxide synthase distribution volume did not significantly differ among groups

ECMO, extracorporeal membrane oxygenation; EVALI, e-cigarette-related or vaping-associated lung injury; COPD, chronic obstructive pulmonary disease; MRI, magnetic resonance imaging; ENDS, electronic nicotine delivery systems; PET, positron emission tomography.

presentation, although some patients described showed evidence of residual imaging findings. Prospective longitudinal imaging studies will help elucidate whether such residual abnormalities are the result of slow healing areas of injury or point to the development of permanent scarring.

PET findings

Table 6 summarizes the included PET studies ($n = 2$). In total, 20 patients were either current or former e-cigarette users, or were acutely exposed to e-cigarettes.

Wetherill et al. (109) used the radiotracer ^{18}F-nitric oxide synthase (NOS) to characterize oxidative stress and inflammation in the lungs of e-cigarette users ($n = 5$) and compare with combustible cigarette smokers ($n = 5$) and healthy controls ($n = 5$). They revealed that e-cigarette users showed greater non-displaceable binding potential (Figure 7A) than combustible cigarette smokers and healthy controls, which is associated with the pro-inflammatory cytokine TNF-alpha that is involved in the inflammatory cascade of acute lung injury. They

also observed that pulmonary inflammation was greater in e-cigarette users than combustible cigarette smokers and healthy controls.

In contrast, Wall et al. (92) investigated the acute deposition of inhaled ^{11}C-nicotine in 15 healthy adults using two nicotine formulations—freebase and lactate salt. In this study, the authors showed that freebase nicotine exhibited high uptake and deposition in the upper respiratory pathways while lactate nicotine was deposited throughout the entire lung and bronchial tree (Figure 7B). Lactate nicotine was also distributed more rapidly than freebase nicotine.

MRI findings

Table 6 also summarizes the included MRI reports ($n = 3$). A total of 23 patients who were current or former e-cigarette users were reported.

Figure 8 shows hyperpolarized ^{129}Xe MRI ventilation (cyan) co-registered with anatomical ^{1}H MRI (greyscale) for one participant with EVALI, one asymptomatic e-cigarette user and one healthy
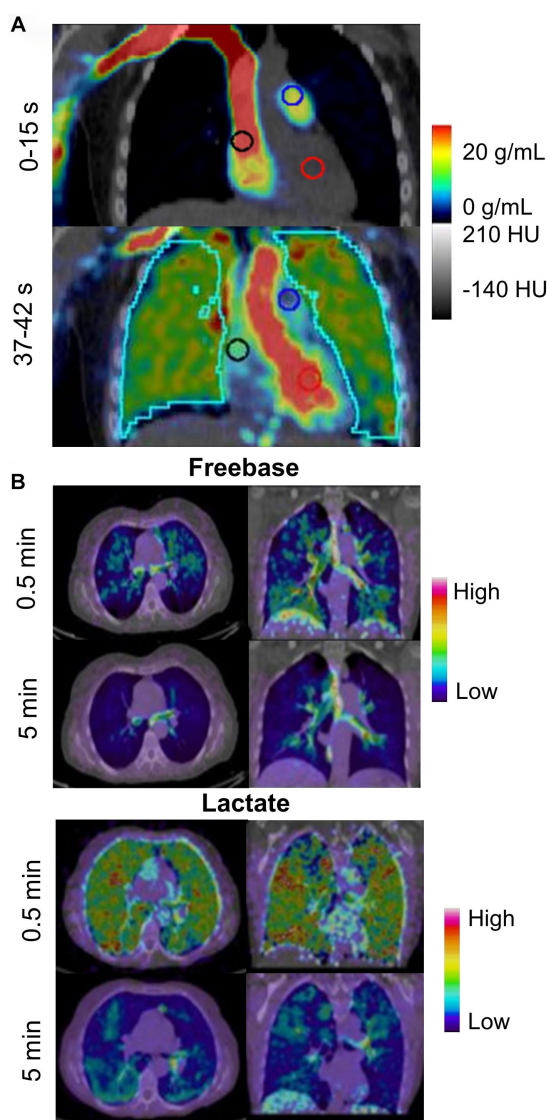
FIGURE 7
PET imaging tracking metabolic activity in e-cigarette users and smokers. **(A)** Oronal PET/CT ^{18}F-NOS in an e-cigarette user (unspecified active ingredient). Uptake is shown 0−15 s and 37−42 s after injection of ^{18}F-NOS to quantify oxidative stress and inflammation in the lungs. Adapted from Wetherill et al. (109) under Creative Commons License. **(B)** Representative distribution of ^{11}C-nicotine freebase and lactate in lungs at 0.5 and 5 min after inhalation of the tagged nicotine salts through an e-cigarette in a healthy adult smoker. Adapted from Wall et al. (92) under Creative Commons License.

participant; all images were acquired by the authors' group. The participant with EVALI demonstrates a heterogenous ventilation pattern with multiple ventilation defects, represented as black regions, throughout the entire lung. In contrast, the asymptomatic e-cigarette user demonstrates a relatively homogenous ventilation pattern with evidence of small, peripheral ventilation defects in the lung apices. Finally, the healthy participant has a homogenous ventilation distribution with no ventilation defects.

Two studies compared the acute effects of e-cigarette use on ventilation. In the first, alveolar ventilation and ventilation heterogeneity in asymptomatic e-cigarette users were similar to

healthy controls prior to exposure, however both measurements worsened following exposure to vaping in e-cigarette users (107). In contrast, while the second study observed a nominal increase in ventilation impairment post-exposure, this change was not statistically significant (108). In a similar manner, the [129]Xe MRI ventilation pattern of a patient recovering from EVALI was highly abnormal and dissimilar to the patterns previously observed in either asthma or COPD (61). Furthermore, these ventilation abnormalities persisted for at least 8 months (61).

MRI perfusion maps are shown for four e-cigarette users in Figure 9, demonstrating changes in regional perfusion following exposure to e-cigarettes. Two studies have evaluated pulmonary perfusion in e-cigarette users prior to and following acute exposure to e-cigarettes and both observed similar findings. Perfusion prior to exposure was similar between healthy controls and e-cigarette users, and following exposure, perfusion was increased (107) while perfusion heterogeneity was decreased (108). Furthermore, e-cigarette users who were exposed to nicotine-based e-liquids demonstrated significant increases in perfusion while those who were exposed to nicotine-free e-liquids were not. Finally, ventilation/perfusion heterogeneity was increased throughout the lungs of e-cigarette users as compared to healthy controls, both prior to and following acute exposure to e-cigarettes (107); the extent of this disruption in ventilation/perfusion matching was similar to what has previously been observed in COPD.

Discussion

Despite nearly 3,000 EVALI cases having been reported in the US in less than 12 months in 2019, the impact of e-cigarette use or vaping on respiratory health is still largely unknown. This is, in part, due to a lack of prospective studies designed to sensitively evaluate the pathological mechanisms responsible for the development of this disease. In this systematic review we summarize in e-cigarette users and patients with EVALI the following points: (1) CT sensitively revealed pulmonary structural abnormalities, (2) PET measured pulmonary inflammation and the deposition of e-cigarette aerosolized particles, and (3) ventilation and perfusion measured via MRI was abnormal. These findings demonstrated that pulmonary functional and structural abnormalities were common in both patients with vaping-induced lung injury and in e-cigarette users with or without respiratory symptoms, and that these abnormalities may be sensitively measured using CT, PET and pulmonary functional MRI. Furthermore, these findings support the use of imaging modalities in prospective studies to help uncover the pathological drivers and mechanisms underlying respiratory symptoms and the development of EVALI in previously healthy adolescents and young adults.

Our systematic review included any report which used CT, PET or MRI to evaluate patients with EVALI or e-cigarette users. CT was the most commonly reported imaging modality, with 105 of the 110 included studies primarily reporting CT findings. Ground-glass opacities, consolidation opacities, subpleural sparing, septal thickening and patchy opacities were among the most common CT findings, in patients with EVALI and e-cigarette users. These findings demonstrated that CT is highly useful in the diagnosis and monitoring of patients with EVALI. Moreover, the worsening or resolution of CT
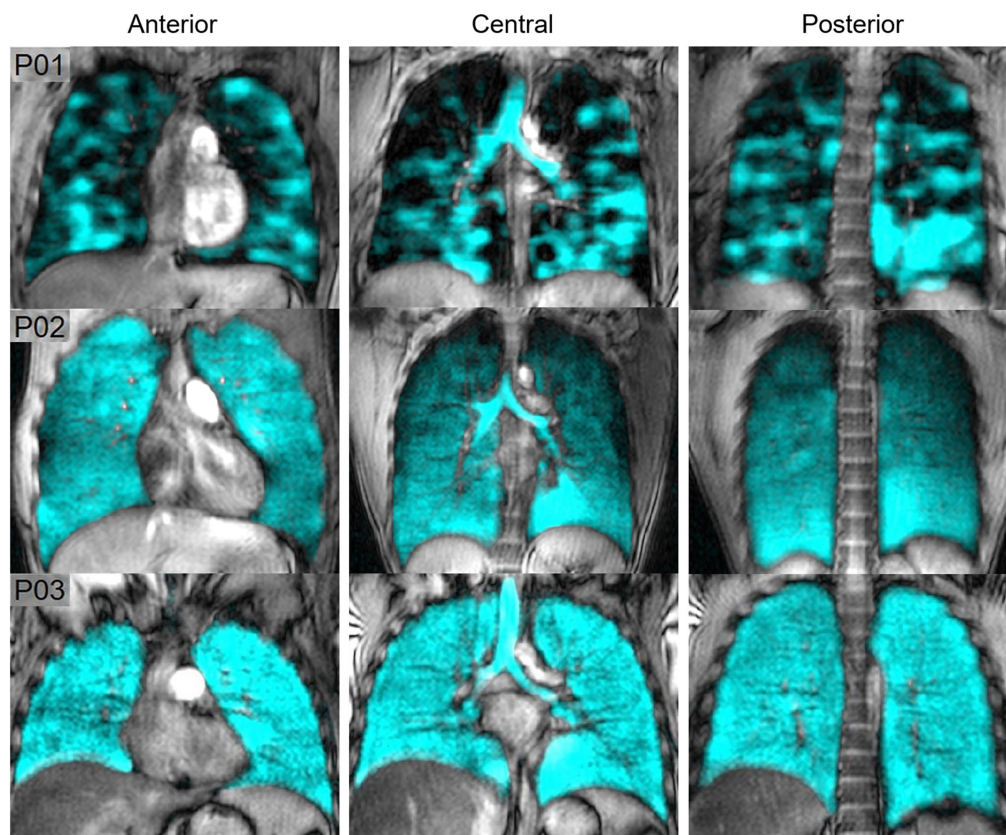
FIGURE 8
Hyperpolarized ^{129}Xe MR ventilation imaging in chronic vapers and similar-aged healthy volunteer. Anterior, central, and posterior coronal slices of ^{129}Xe ventilation (cyan) co-registered with anatomical ^{1}H (greyscale) MRI. All images were acquired by the authors' group. P01 is an 18 years-old male with severe bronchiolitis and respiratory failure caused by e-cigarette use, 1 month post discharge after a 6 months history of vaping (VDP = 21%). P02 is a 29 years-old male with a 2.5 years history of vaping nicotine (3/4 pods daily) (VDP = 4.7%). P03 is a 22 years-old male with no history of chronic respiratory abnormalities, vaping or combustible cigarette use (VDP = 1.5%).

abnormalities, with or without therapeutic intervention, may help inform on the pathological drivers of EVALI, point towards potential treatment options or help identify characteristics of patients at risk of developing serious and/or permanent lung injury as a result of e-cigarette use.

PET showed that there were differences in the deposition distribution of freebase or lactate salt nicotine aerosolized particles (92). This is important because early e-liquids primarily used freebase nicotine however, since 2017, lactate salt nicotine is more commonly used (110). The health risks of lactate salt and freebase nicotine have not yet been studied. PET also revealed a unique inflammatory response to acute e-cigarette exposure, in comparison with acute combustible cigarette exposure or in healthy controls. These findings point towards possible mechanisms of acute lung injury following e-cigarette use. Together, these preliminary studies may drive further hypothesis-driven research into the pulmonary health effects associated with e-cigarette use.

Unlike CT and PET, pulmonary functional MRI does not require ionizing radiation and multiple acquisitions can be completed over either short or long periods of time. In the context of vaping, pulmonary functional MRI sensitively revealed ventilation and perfusion abnormalities in e-cigarette users following acute exposure to e-cigarettes (107, 108), and in a patient recovering from EVALI

(61). Ventilation abnormalities in e-cigarette users may be driven by airflow obstruction via inflammation, impaired mucus clearance, constriction or collapse (107). In addition, perfusion abnormalities were worse in patients using nicotine-based e-liquids as compared to those using nicotine-free e-liquids, which points towards known impacts of nicotine on cardiovascular hemodynamics, increases in heart rate, systolic blood pressure and cardiac output (107). Furthermore, findings of increased perfusion and decreased perfusion heterogeneity in e-cigarette users, regardless of e-liquid content, may suggest hypoxic vasoconstriction as a compensatory mechanism in the presence of abnormal ventilation (107). This idea is further supported by the observation of increased ventilation/perfusion heterogeneity. Together, these findings show that MRI may be used to sensitively measure both the acute and chronic effects of e-cigarette use on both pulmonary ventilation and perfusion.

Limitations

The reviewed manuscripts have a number of important limitations. First, our search criteria were quite specific and thus we may have excluded relevant studies which did not use our precise terminology. Second, technical CT parameters were not commonly
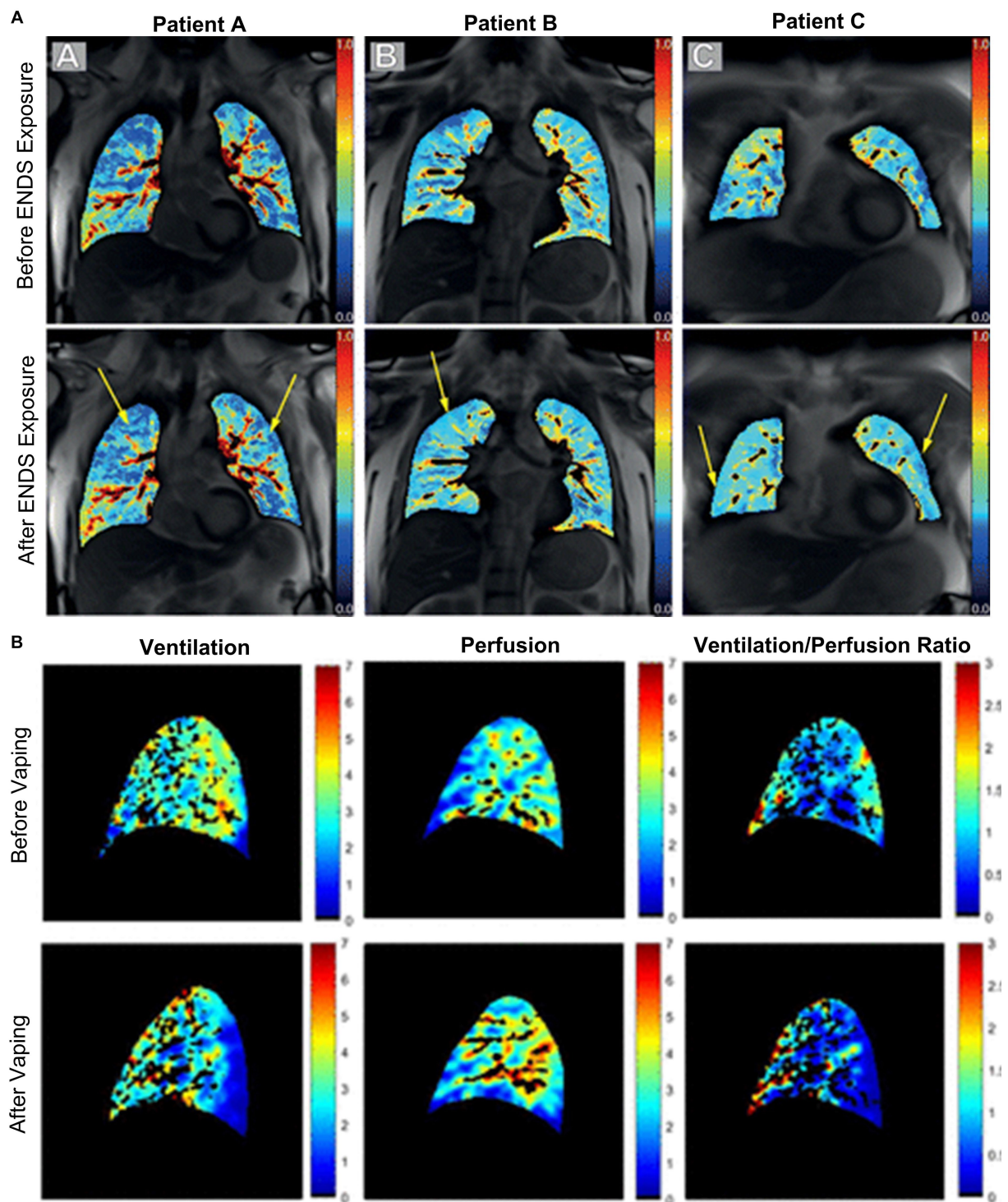
FIGURE 9
Free-breathing ^{1}H MRI of acute effects of vaping. **(A)** Pulmonary perfusion images obtained by using non-contrast matrix pencil MRI in three electronic nicotine vape users before and after exposure. The arrows indicate lung regions with increased regional perfusion post-exposure. Red corresponds to greater perfusion amplitude and blue corresponds to lower values (108). **(B)** Spatial maps of ventilation, perfusion, and ventilation-perfusion ratios, before and after vaping, in a single sagittal slice of one representative subject. Black regions within the lung field represent regions excluded from analysis, including conducting airways and vessels, and regions with poor signal to noise (107).

described. The variation of CT parameters between studies and patients may have impacted the radiologist's findings and thus, it may be possible that some CT abnormalities were undetected. Third, the

commercialization and distribution of e-cigarette devices is recent and thus, the majority of the included studies reported on data acquired within the past 4 years. It is not possible at this time to comment on

the long-term health effects associated with either continuous or sporadic e-cigarette use as this has not yet been investigated and/or reported. Fourth, a large proportion (26% of studies and 11% of patients) of the included studies did not report or were not able to ascertain the specific ingredients present in the e-liquids used by the studied patients. With the wide variety of e-liquids available, as well as the potential combination of multiple e-liquids by users, the effect of individual or combinations of e-liquids and their ingredients on respiratory health is largely unknown. Finally, e-cigarette devices have evolved and many components of these may be customized, including the voltage of the battery, the e-liquid temperature, nicotine yield or puff volume (111–113), introducing a wide number of variables that may contribute to vaping-induced lung injury.

Conclusion

With the growing popularity of e-cigarettes for recreational use in people who have not previously smoked combustible cigarettes, it is important to understand the short- and long-term effects of vaping on respiratory health. This systematic review, while revealing only a handful of functional imaging studies, showed that pulmonary imaging and in particular functional imaging is helpful in generating a better understanding of the acute and chronic effects of vaping on pulmonary structure and function. The time is right to explore larger scale, multi-centre studies using pulmonary functional imaging methods to uncover the pathological mechanisms driving vaping-induced lung injury and respiratory symptoms, and to evaluate interventions.

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

JH: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. VP: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. JZ: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. MS: Supervision, Visualization, Writing – review &

editing. GP: Conceptualization, Investigation, Methodology, Supervision, Writing – review & editing. MM: Supervision, Visualization, Writing – review & editing.

Conflict of interest

GP discloses investigator initiated study funding from AstraZeneca, Novartis and GSK, study funding from the Ministry of Health and Long-term Care Ontario, and honoraria from AstraZeneca for speaking engagements, outside the submitted 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.

The reviewer LS declared a past co-authorship with the authors MS and GP to the handling editor.

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1285361/full#supplementary-material

References

1. Siegel DA, Jatlaoui TC, Koumans EH, Kiernan EA, Layer M, Cates JE, et al. Update: interim guidance for health care providers evaluating and caring for patients with suspected e-cigarette, or vaping, product use associated lung injury—United States, October 2019. *MMWR Morb Mortal Wkly Rep*. (2019) 68:919–27. doi: 10.15585/mmwr.mm6841e3

2. Layden JE, Ghinai I, Pray I, Kimball A, Layer M, Tenforde MW, et al. Pulmonary illness related to e-cigarette use in Illinois and Wisconsin—final report. *N Engl J Med*. (2020) 382:903–16. doi: 10.1056/NEJMoa1911614

3. Centers for Disease Control and Prevention. Outbreak of lung injury associated with the use of e-cigarette, or vaping, products. (2021). Available at: https://www.cdc.gov/tobacco/basic_information/e-cigarettes/severe-lung-disease.html

4. Schaffer S, Strang A, Saul D, Krishnan V, Chidekel A. Adolescent e-cigarette or vaping use-associated lung injury in the Delaware Valley: a review of hospital-based

presentation, management, and outcomes. *Cureus*. (2022) 14:e21988. doi: 10.7759/cureus.21988

5. Hall W, Gartner C, Bonevski B. Lessons from the public health responses to the US outbreak of vaping-related lung injury. *Addiction*. (2021) 116:985–93. doi: 10.1111/add.15108

6. Zulfiqar H, Sankari A, Rahman O. Vaping-associated pulmonary injury In: *StatPearls*. Treasure Island, FL: StatPearls Publishing LLC (2023)

7. National Center for Chronic Disease Prevention and Health Promotion. *e-cigarette use among youth and young adults: a report of the surgeon general*. Rockville, MD: National Center for Chronic Disease Prevention and Health Promotion (U.S.) (2016).

8. National Academies of Sciences, Engineering, and Medicine. *Public health consequences of e-cigarettes*. Washington DC: The National Academies Press (2018).

9. Wang JB, Olgin JE, Nah G, Vittinghoff E, Cataldo JK, Pletcher MJ, et al. Cigarette and e-cigarette dual use and risk of cardiopulmonary symptoms in the health eHeart study. *PLoS One*. (2018) 13:e0198681. doi: 10.1371/journal.pone.0198681

10. Moheimani RS, Bhetraratana M, Yin F, Peters KM, Gornbein J, Araujo JA, et al. Increased cardiac sympathetic activity and oxidative stress in habitual electronic cigarette users: implications for cardiovascular risk. *JAMA Cardiol*. (2017) 2:278–84. doi: 10.1001/jamacardio.2016.5303

11. Wills TA, Soneji SS, Choi K, Jaspers I, Tam EK. e-cigarette use and respiratory disorders: an integrative review of converging evidence from epidemiological and laboratory studies. *Eur Respir J*. (2021) 57:1901815. doi: 10.1183/13993003.01815-2019

12. Marques P, Piqueras L, Sanz MJ. An updated overview of e-cigarette impact on human health. *Respir Res*. (2021) 22:151. doi: 10.1186/s12931-021-01737-5

13. Banks E, Yazidjoglou A, Brown S, Nguyen M, Martin M, Beckwith K, et al. Electronic cigarettes and health outcomes: umbrella and systematic review of the global evidence. *Med J Aust*. (2023) 218:267–75. doi: 10.5694/mja2.51890

14. Canada Go. Canadian tobacco and nicotine survey (CTNS): summary of results for 2021 (2021) Available at: https://www.canada.ca/en/health-canada/services/canadian-tobacco-nicotine-survey/2021-summary.html

15. White AV, Wambui DW, Pokhrel LR. Risk assessment of inhaled diacetyl from electronic cigarette use among teens and adults. *Sci Total Environ*. (2021) 772:145486. doi: 10.1016/j.scitotenv.2021.145486

16. Cao DJ, Aldy K, Hsu S, McGetrick M, Verbeck G, De Silva I, et al. Review of health consequences of electronic cigarettes and the outbreak of electronic cigarette, or vaping, product use-associated lung injury. *J Med Toxicol*. (2020) 16:295–310. doi: 10.1007/s13181-020-00772-w

17. Gotts JE, Jordt SE, McConnell R, Tarran R. What are the respiratory effects of e-cigarettes? *BMJ*. (2019) 366:l5275. doi: 10.1136/bmj.l5275

18. Jonas A. Impact of vaping on respiratory health. *BMJ*. (2022) 378:e065997. doi: 10.1136/bmj-2021-065997

19. Blount BC, Karwowski MP, Shields PG, Morel-Espinosa M, Valentin-Blasini L, Gardner M, et al. Vitamin E acetate in bronchoalveolar-lavage fluid associated with EVALI. *N Engl J Med*. (2020) 382:697–705. doi: 10.1056/NEJMoa1916433

20. Clapp PW, Pawlak EA, Lackey JT, Keating JE, Reeber SL, Glish GL, et al. Flavored e-cigarette liquids and cinnamaldehyde impair respiratory innate immune cell function. *Am J Physiol Lung Cell Mol Physiol*. (2017) 313:L278–92. doi: 10.1152/ajplung.00452.2016

21. Stefaniak AB, LeBouf RF, Ranpara AC, Leonard SS. Toxicology of flavoring- and cannabis-containing e-liquids used in electronic delivery systems. *Pharmacol Ther*. (2021) 224:107838. doi: 10.1016/j.pharmthera.2021.107838

22. Smith MR, Jarrell ZR, Orr M, Liu KH, Go Y-M, Jones DP. Metabolome-wide association study of flavorant vanillin exposure in bronchial epithelial cells reveals disease-related perturbations in metabolism. *Environ Int*. (2021) 147:106323. doi: 10.1016/j.envint.2020.106323

23. Hickman E, Herrera CA, Jaspers I. Common e-cigarette flavoring chemicals impair neutrophil phagocytosis and oxidative burst. *Chem Res Toxicol*. (2019) 32:982–5. doi: 10.1021/acs.chemrestox.9b00171

24. Durrani K, Din S-MAE, Sun Y, Rule AM, Bressler J. Ethyl maltol enhances copper mediated cytotoxicity in lung epithelial cells. *Toxicol Appl Pharmacol*. (2021) 410:115354. doi: 10.1016/j.taap.2020.115354

25. Nair V, Tran M, Behar RZ, Zhai S, Cui X, Phandthong R, et al. Menthol in electronic cigarettes: a contributor to respiratory disease? *Toxicol Appl Pharmacol*. (2020) 407:115238. doi: 10.1016/j.taap.2020.115238

26. Cherinka B, Andrews BH, Sanchez-Gallego J, Brownstein J, Argudo-Fernandez M, Blanton M, et al. Marvin: a tool kit for streamlined access and visualization of the SDSS-IV MaNGA data set. *Astron J*. (2019) 158:74. doi: 10.3847/1538-3881/ab2634

27. Traboulsi H, Cherian M, Rjeili MA, Preteroti M, Bourbeau J, Smith BM, et al. Inhalation toxicology of vaping products and implications for pulmonary health. *Int J Mol Sci*. (2020) 21:3495. doi: 10.3390/ijms21103495

28. Federal Food, Drug, and Cosmetic Act. *Federal Food, Drug, and Cosmetic Act. Section 201*. Washington DC: US Food and Drug Administration. (1938). 409 p.

29. Cherian SV, Kumar A, Estrada YMRM. e-cigarette or vaping product-associated lung injury: a review. *Am J Med*. (2020) 133:657–63. doi: 10.1016/j.amjmed.2020.02.004

30. Kreiss K, Gomaa A, Kullman G, Fedan K, Simoes EJ, Enright PL. Clinical bronchiolitis obliterans in workers at a microwave-popcorn plant. *N Engl J Med*. (2002) 347:330–8. doi: 10.1056/NEJMoa020300

31. Bracken-Clarke D, Kapoor D, Baird AM, Buchanan PJ, Gately K, Cuffe S, et al. Vaping and lung cancer—a review of current data and recommendations. *Lung Cancer*. (2021) 153:11–20. doi: 10.1016/j.lungcan.2020.12.030

32. Jacob SE, Scheman A, McGowan MA. Propylene glycol. *Dermatitis*. (2018) 29:3–5. doi: 10.1097/DER.0000000000000315

33. Komura M, Sato T, Yoshikawa H, Nitta NA, Suzuki Y, Koike K, et al. Propylene glycol, a component of electronic cigarette liquid, damages epithelial cells in human small airways. *Respir Res*. (2022) 23:216. doi: 10.1186/s12931-022-02142-2

34. Kim MD, Chung S, Dennis JS, Yoshida M, Aguiar C, Aller SP, et al. Vegetable glycerin e-cigarette aerosols cause airway inflammation and ion channel dysfunction. *Front Pharmacol*. (2022) 13:1012723. doi: 10.3389/fphar.2022.1012723

35. Massey JB, She HS, Pownall HJ. Interaction of vitamin E with saturated phospholipid bilayers. *Biochem Biophys Res Commun*. (1982) 106:842–7. doi: 10.1016/0006-291X(82)91787-9

36. Kamal MA, Raghunathan VA. Modulated phases of phospholipid bilayers induced by tocopherols. *Biochim Biophys Acta*. (2012) 1818:2486–93. doi: 10.1016/j.bbamem.2012.06.016

37. Casals C, Canadas O. Role of lipid ordered/disordered phase coexistence in pulmonary surfactant function. *Biochim Biophys Acta*. (2012) 1818:2550–62. doi: 10.1016/j.bbamem.2012.05.024

38. Van Bavel N, Lai P, Amrein M, Prenner EJ. Pulmonary surfactant function and molecular architecture is disrupted in the presence of vaping additives. *Colloids Surf B*. (2023) 222:113132. doi: 10.1016/j.colsurfb.2023.113132

39. Winnicka L, Shenoy MA. EVALI and the pulmonary toxicity of electronic cigarettes: a review. *J Gen Intern Med*. (2020) 35:2130–5. doi: 10.1007/s11606-020-05813-2

40. Kligerman S, Raptis C, Larsen B, Henry TS, Caporale A, Tazelaar H, et al. Radiologic, pathologic, clinical, and physiologic findings of electronic cigarette or vaping product use-associated lung injury (EVALI): evolving knowledge and remaining questions. *Radiology*. (2020) 294:491–505. doi: 10.1148/radiol.2020192585

41. Lee HW, Park SH, Weng MW, Wang HT, Huang WC, Lepor H, et al. e-cigarette smoke damages DNA and reduces repair activity in mouse lung, heart, and bladder as well as in human lung and bladder cells. *Proc Natl Acad Sci U S A*. (2018) 115:E1560–9. doi: 10.1073/pnas.1718135115

42. Ghosh A, Coakley RD, Ghio AJ, Muhlebach MS, Esther CR Jr, Alexis NE, et al. Chronic e-cigarette use increases neutrophil elastase and matrix metalloprotease levels in the lung. *Am J Respir Crit Care Med*. (2019) 200:1392–401. doi: 10.1164/rccm.201903-0615OC

43. Garcia-Arcos I, Geraghty P, Baumlin N, Campos M, Dabo AJ, Jundi B, et al. Chronic electronic cigarette exposure in mice induces features of COPD in a nicotine-dependent manner. *Thorax*. (2016) 71:1119–29. doi: 10.1136/thoraxjnl-2015-208039

44. Herriges MJ, Pinkhasov R, Shapiro O, Jacob JM, Basnet A, Bratslavsky G, et al. e-cigarette use and the risk of bladder and lung cancer. *J Clin Oncol*. (2022) 40:443. doi: 10.1200/JCO.2022.40.6_suppl.443

45. Saji S, Patil SS, Alleyn M, Lockey R, Kolliputi N. Nicotine in e-cigarette smoke: cancer culprit? *J Cell Commun Signal*. (2020) 14:127–8. doi: 10.1007/s12079-019-00519-5

46. Schaal CM, Bora-Singhal N, Kumar DM, Chellappan SP. Regulation of Sox2 and stemness by nicotine and electronic-cigarettes in non-small cell lung cancer. *Mol Cancer*. (2018) 17:149. doi: 10.1186/s12943-018-0901-2

47. Salzman GA, Alqawasma M, Asad H. Vaping associated lung injury (EVALI): an explosive United States epidemic. *Mo Med*. (2019) 116:492–6.

48. Muthumalage T, Lamb T, Friedman MR, Rahman I. e-cigarette flavored pods induce inflammation, epithelial barrier dysfunction, and DNA damage in lung epithelial cells and monocytes. *Sci Rep*. (2019) 9:19035. doi: 10.1038/s41598-019-51643-6

49. Bitzer ZT, Goel R, Reilly SM, Elias RJ, Silakov A, Foulds J, et al. Effect of flavoring chemicals on free radical formation in electronic cigarette aerosols. *Free Radic Biol Med*. (2018) 120:72–9. doi: 10.1016/j.freeradbiomed.2018.03.020

50. Tsai M, Byun MK, Shin J, Crotty Alexander LE. Effects of e-cigarettes and vaping devices on cardiac and pulmonary physiology. *J Physiol*. (2020) 598:5039–62. doi: 10.1113/JP279754

51. Bodas M, Van Westphal C, Carpenter-Thompson R, Arnold EC 4th, Johnson SC, Harb A. Nicotine exposure induces bronchial epithelial cell apoptosis and senescence via ROS mediated autophagy-impairment. *Free Radic Biol Med*. (2016) 97:441–53. doi: 10.1016/j.freeradbiomed.2016.06.017

52. Krishna R, Anjum F, Oliver TI. Bronchiolitis obliterans In: *StatPearls*. Treasure Island, FL: StatPearls (2023)

53. Yang X, Kong J, Yang MX, Xie WP, Wang H, Rong R, et al. Analysis of clinical characteristics of 21 cases of acute fibrinous and organizing pneumonia. *Zhonghua Jie He He Hu Xi Za Zhi*. (2020) 43:670–6. doi: 10.3760/cma.j.cn112147-20191014-00685

54. Jeon J, Zhang Q, Chepaitis PS, Greenwald R, Black M, Wright C. Toxicological assessment of particulate and metal hazards associated with vaping frequency and device age. *Toxics*. (2023) 11:155. doi: 10.3390/toxics11020155

55. Taucher E, Mykoliuk I, Lindenmann J, Smolle-Juettner FM. Implications of the immune landscape in COPD and lung cancer: smoking versus other causes. *Front Immunol*. (2022) 13:846605. doi: 10.3389/fimmu.2022.846605

56. Fei QY, Weng XQ, Liu K, Liu S, Chen JM, Guo XR, et al. The relationship between metal exposure and chronic obstructive pulmonary disease in the general US population: NHANES 2015–2016. *Int J Environ Res Pub Health*. (2022) 19:2085. doi: 10.3390/ijerph19042085

57. McRae N, Gennings C, Rivera NR, Tamayo-Ortiz M, Pantic I, Amarasiriwardena C, et al. Association between prenatal metal exposure and adverse respiratory symptoms in childhood. *Environ Res*. (2022) 205:112448. doi: 10.1016/j.envres.2021.112448

58. Ko TJ, Kim SA. Effect of heating on physicochemical property of aerosols during vaping. *Int J Environ Res Public Health*. (2022) 19:1892. doi: 10.3390/ijerph19031892

59. Guerrini V, Panettieri RA Jr, Gennaro ML. Lipid-laden macrophages as biomarkers of vaping-associated lung injury. *Lancet Respir Med*. (2020) 8:e6. doi: 10.1016/S2213-2600(19)30476-X

60. Ghosh A, Ahmad S, Coakley RD, Sassano MF, Alexis NE, Tarran R. Lipid-laden macrophages are not unique to patients with e-cigarette or vaping product use-associated lung injury. *Am J Respir Crit Care Med.* (2021) 203:1030–3. doi: 10.1164/rccm.202009-3507LE

61. Eddy RL, Serajeddini H, Knipping D, Landman ST, Bosma KJ, Mackenzie CA, et al. Pulmonary functional MRI and CT in a survivor of bronchiolitis and respiratory failure caused by e-cigarette use. *Chest.* (2020) 158:e147–51. doi: 10.1016/j.chest.2020.06.005

62. Conway J. Lung imaging - two dimensional gamma scintigraphy, SPECT, CT and PET. *Adv Drug Deliv Rev.* (2012) 64:357–68. doi: 10.1016/j.addr.2012.01.013

63. Iking J, Staniszewska M, Kessler L, Klose JM, Luckerath K, Fendler WP, et al. Imaging inflammation with positron emission tomography. *Biomedicine.* (2021) 9:212. doi: 10.3390/biomedicines9020212

64. Nyilas S, Bauman G, Pusterla O, Ramsey K, Singer F, Stranzinger E, et al. Ventilation and perfusion assessed by functional MRI in children with CF: reproducibility in comparison to lung function. *J Cyst Fibros.* (2019) 18:543–50. doi: 10.1016/j.jcf.2018.10.003

65. Nyilas S, Bauman G, Pusterla O, Sommer G, Singer F, Stranzinger E, et al. Structural and functional lung impairment in primary ciliary dyskinesia. Assessment with magnetic resonance imaging and multiple breath washout in comparison to spirometry. *Ann Am Thorac Soc.* (2018) 15:1434–42. doi: 10.1513/AnnalsATS.201712-967OC

66. Ebner L, Kammerman J, Driehuys B, Schiebler ML, Cadman RV, Fain SB. The role of hyperpolarized ^{129}xenon in MR imaging of pulmonary function. *Eur J Radiol.* (2017) 86:343–52. doi: 10.1016/j.ejrad.2016.09.015

67. Kooner HK, McIntosh MJ, Desaigoudar V, Rayment JH, Eddy RL, Driehuys B, et al. Pulmonary functional MRI: detecting the structure-function pathologies that drive asthma symptoms and quality of life. *Respirology.* (2022) 27:114–33. doi: 10.1111/resp.14197

68. Fain S, Schiebler ML, McCormack DG, Parraga G. Imaging of lung function using hyperpolarized helium-3 magnetic resonance imaging: review of current and emerging translational methods and applications. *J Magn Reson Imaging.* (2010) 32:1398–408. doi: 10.1002/jmri.22375

69. Vidal Melo MF, Layfield D, Harris RS, O'Neill K, Musch G, Richter T, et al. Quantification of regional ventilation-perfusion ratios with PET. *J Nucl Med.* (2003) 44:1982–91.

70. Altes TA, Rehm PK, Harrell F, Salerno M, Daniel TM, De Lange EE. Ventilation imaging of the lung: comparison of hyperpolarized helium-3 MR imaging with Xe-133 scintigraphy. *Acad Radiol.* (2004) 11:729–34. doi: 10.1016/j.acra.2004.04.001

71. Santyr GE, Couch MJ, Lam WW, Ouriadov A, Drangova M, McCormack DG, et al. Comparison of hyperpolarized ^{3}He MRI with Xe-enhanced computed tomography imaging for ventilation mapping of rat lung. *NMR Biomed.* (2011) 24:1073–80. doi: 10.1002/nbm.1659

72. Mummy DG, Dunican EM, Carey KJ, Evans MD, Elicker BM, Newell JD Jr, et al. Mucus plugs in asthma at CT associated with regional ventilation defects at helium 3 MRI. *Radiology.* (2021) 303:184–90. doi: 10.1148/radiol.2021204616

73. Svenningsen S, Eddy RL, Lim HF, Cox PG, Nair P, Parraga G. Sputum eosinophilia and magnetic resonance imaging ventilation heterogeneity in severe asthma. *Am J Respir Crit Care Med.* (2018) 197:876–84. doi: 10.1164/rccm.201709-1948OC

74. Svenningsen S, Haider E, Boylan C, Mukherjee M, Eddy RL, Capaldi DPI, et al. CT and functional MRI to evaluate airway mucus in severe asthma. *Chest.* (2019) 155:1178–89. doi: 10.1016/j.chest.2019.02.403

75. Svenningsen S, Kirby M, Starr D, Coxson HO, Paterson NA, McCormack DG, et al. What are ventilation defects in asthma? *Thorax.* (2014) 69:63–71. doi: 10.1136/thoraxjnl-2013-203711

76. Kirby M, Svenningsen S, Kanhere N, Owrangi A, Wheatley A, Coxson HO, et al. Pulmonary ventilation visualized using hyperpolarized helium-3 and xenon-129 magnetic resonance imaging: differences in COPD and relationship to emphysema. *J Appl Physiol.* (2013) 114:707–15. doi: 10.1152/japplphysiol.01206.2012

77. Sheikh K, Coxson HO, Parraga G. This is what COPD looks like. *Respirology.* (2016) 21:224–36. doi: 10.1111/resp.12611

78. Santyr G, Kanhere N, Morgado F, Rayment JH, Ratjen F, Couch MJ. Hyperpolarized gas magnetic resonance imaging of pediatric cystic fibrosis lung disease. *Acad Radiol.* (2019) 26:344–54. doi: 10.1016/j.acra.2018.04.024

79. Mammarappallil JG, Rankine L, Wild JM, Driehuys B. New developments in imaging idiopathic pulmonary fibrosis with hyperpolarized xenon magnetic resonance imaging. *J Thorac Imaging.* (2019) 34:136–50. doi: 10.1097/RTI.0000000000000392

80. *FDA approves Polarean's XENOVIEW^{TM} (xenon Xe 129 hyperpolarized) for use with MRI for the evaluation of lung ventilation.* North Carolina USA: Polarean Imaging PLC. (Press release). (2022).

81. Sa RC, Cronin MV, Henderson AC, Holverda S, Theilmann RJ, Arai TJ, et al. Vertical distribution of specific ventilation in normal supine humans measured by oxygen-enhanced proton MRI. *J Appl Physiol.* (1985) 109:1950–9. doi: 10.1152/japplphysiol.00220.2010

82. Morgan AR, Parker GJ, Roberts C, Buonaccorsi GA, Maguire NC, Hubbard Cristinacce PL, et al. Feasibility assessment of using oxygen-enhanced magnetic resonance imaging for evaluating the effect of pharmacological treatment in COPD. *Eur J Radiol.* (2014) 83:2093–101. doi: 10.1016/j.ejrad.2014.08.004

83. Ohno Y, Nishio M, Koyama H, Yoshikawa T, Matsumoto S, Seki S, et al. Oxygen-enhanced MRI for patients with connective tissue diseases: comparison with thin-section CT of capability for pulmonary functional and disease severity assessment. *Eur J Radiol.* (2014) 83:391–7. doi: 10.1016/j.ejrad.2013.11.001

84. Zhang WJ, Niven RM, Young SS, Liu YZ, Parker GJ, Naish JH. Dynamic oxygen-enhanced magnetic resonance imaging of the lung in asthma—initial experience. *Eur J Radiol.* (2015) 84:318–26. doi: 10.1016/j.ejrad.2014.10.021

85. Capaldi DP, Sheikh K, Guo F, Svenningsen S, Etemad-Rezai R, Coxson HO, et al. Free-breathing pulmonary 1H and hyperpolarized 3He MRI: comparison in COPD and bronchiectasis. *Acad Radiol.* (2015) 22:320–9. doi: 10.1016/j.acra.2014.10.003

86. Klimes F, Voskrebenzev A, Gutberlet M, Kern AL, Behrendt L, Grimm R, et al. 3D phase-resolved functional lung ventilation MR imaging in healthy volunteers and patients with chronic pulmonary disease. *Magn Reson Med.* (2021) 85:912–25. doi: 10.1002/mrm.28482

87. Matheson AM, Capaldi DP, Guo F, Eddy RL, McCormack DG, Parraga G. Fourier decomposition free-breathing 1H MRI perfusion maps in asthma Medical Imaging 2019: Image Processing; (2019)

88. Voskrebenzev A, Gutberlet M, Klimes F, Kaireit TF, Schonfeld C, Rotarmel A, et al. Feasibility of quantitative regional ventilation and perfusion mapping with phase-resolved functional lung (PREFUL) MRI in healthy volunteers and COPD, CTEPH, and CF patients. *Magn Reson Med.* (2018) 79:2306–14. doi: 10.1002/mrm.26893

89. Bauman G, Bieri O. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. *Magn Reson Med.* (2017) 77:336–42. doi: 10.1002/mrm.26096

90. Bolar DS, Levin DL, Hopkins SR, Frank LF, Liu TT, Wong EC, et al. Quantification of regional pulmonary blood flow using ASL-FAIRER. *Magn Reson Med.* (2006) 55:1308–17. doi: 10.1002/mrm.20891

91. Landman ST, Dhaliwal I, Mackenzie CA, Martinu T, Steel A, Bosma KJ. Life-threatening bronchiolitis related to electronic cigarette use in a Canadian youth. *CMAJ.* (2019) 191:E1321–31. doi: 10.1503/cmaj.191402

92. Wall A, Roslin S, Borg B, McDermott S, Walele T, Nahde T, et al. e-cigarette aerosol deposition and disposition of [^{11}C]nicotine using positron emission tomography: a comparison of nicotine uptake in lungs and brain using two different nicotine formulations. *Pharmaceuticals.* (2022) 15:367. doi: 10.3390/ph15030367

93. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J. Fleischner society: glossary of terms for thoracic imaging. *Radiology.* (2008) 246:697–722. doi: 10.1148/radiol.2462070712

94. Kligerman SJ, Kay FU, Raptis CA, Henry TS, Sechrist JW, Walker CM, et al. CT findings and patterns of e-cigarette or vaping product use-associated lung injury: a multicenter cohort of 160 cases. *Chest.* (2021) 160:1492–511. doi: 10.1016/j.chest.2021.04.054

95. Aberegg SK, Cirulis MM, Maddock SD, Freeman A, Keenan LM, Pirozzi CS, et al. Clinical, bronchoscopic, and imaging findings of e-cigarette, or vaping, product use-associated lung injury among patients treated at an Academic Medical Center. *JAMA Netw Open.* (2020) 3:e2019176. doi: 10.1001/jamanetworkopen.2020.19176

96. Artunduaga M, Rao D, Friedman J, Kwon JK, Pfeifer CM, Dettori A, et al. Pediatric chest radiographic and CT findings of electronic cigarette or vaping product use-associated lung injury (EVALI). *Radiology.* (2020) 295:430–8. doi: 10.1148/radiol.2020192778

97. Carroll BJ, Kim M, Hemyari A, Thakrar P, Kump TE, Wade T, et al. Impaired lung function following e-cigarette or vaping product use associated lung injury in the first cohort of hospitalized adolescents. *Pediatr Pulmonol.* (2020) 55:1712–8. doi: 10.1002/ppul.24787

98. Chidambaram AG, Dennis RA, Biko DM, Hook M, Allen J, Rapp JB. Clinical and radiological characteristics of e-cigarette or vaping product use associated lung injury. *Emerg Radiol.* (2020) 27:495–501. doi: 10.1007/s10140-020-01796-z

99. Harry-Hernandez S, Thiboutot J, Wahidi MM, Giovacchini CX, De Cardenas J, Meldrum C, et al. Bronchoalveolar lavage (BAL) and pathologic assessment of electronic cigarette or vaping product use-associated lung injury (EVALI): the EVALI-BAL study, a multicenter cohort. *J Bronchology Interv Pulmonol.* (2023) 30:144–54. doi: 10.1097/LBR.0000000000000890

100. Helfgott D, Capozzoli G, Madray J, Baig A, Uppaluri L, Gaur S, et al. e-cigarette or vaping product use associated lung injury (EVALI) in the time of COVID-19: a clinical dilemma. *Pediatr Pulmonol.* (2022) 57:623–30. doi: 10.1002/ppul.25804

101. Kalininskiy A, Bach CT, Nacca NE, Ginsberg G, Marraffa J, Navarette KA, et al. e-cigarette, or vaping, product use associated lung injury (EVALI): case series and diagnostic approach. *Lancet Respir Med.* (2019) 7:1017–26. doi: 10.1016/S2213-2600(19)30415-1

102. Layden JE, Meiman JG. Pulmonary disease related to e-cigarette use. Reply. *N Engl J Med.* (2020) 383:793. doi: 10.1056/NEJMc2009932

103. Pajak A, Bascoy S, Li JC, Benninghoff M, Deitchman A. e-cigarette or vaping product use associated lung injury among three young adults: a retrospective case series from Delaware. *Cureus.* (2020) 12:e11031. doi: 10.7759/cureus.11031

104. Panse PM, Feller FF, Butt YM, Smith ML, Larsen BT, Tazelaar HD, et al. Pulmonary injury resulting from vaping or e-cigarette use: imaging appearances at presentation and follow-up. *Radiol Cardiothorac Imaging.* (2020) 2:e200081. doi: 10.1148/ryct.2020200081

105. Rao DR, Maple KL, Dettori A, Afolabi F, Francis JKR, Artunduaga M, et al. Clinical features of e-cigarette, or vaping, product use-associated lung injury in teenagers. *Pediatrics*. (2020) 146:e20194104. doi: 10.1542/peds.2019-4104

106. Wang KY, Jadhav SP, Yenduri NJS, Lee SA, Farber HJ, Guillerman RP. e-cigarette or vaping product use-associated lung injury in the pediatric population: imaging features at presentation and short-term follow-up. *Pediatr Radiol*. (2020) 50:1231–9. doi: 10.1007/s00247-020-04698-x

107. Kizhakke Puliyakote AS, Elliott AR, Sá RC, Anderson KM, Crotty Alexander LE, Hopkins SR. Vaping disrupts ventilation-perfusion matching in asymptomatic users. *J Appl Physiol*. (1985) 130:308–17. doi: 10.1152/japplphysiol.00709.2020

108. Nyilas S, Bauman G, Korten I, Pusterla O, Singer F, Ith M, et al. MRI shows lung perfusion changes after vaping and smoking. *Radiology*. (2022) 304:195–204. doi: 10.1148/radiol.211327

109. Wetherill RR, Doot RK, Young AJ, Lee H, Schubert EK, Wiers CE, et al. Molecular imaging of pulmonary inflammation in users of electronic and combustible cigarettes: a pilot study. *J Nucl Med*. (2023) 64:797–802. doi: 10.2967/jnumed.122.264529

110. Han S, Liu C, Chen H, Fu Y, Zhang Y, Miao R, et al. Pharmacokinetics of freebase nicotine and nicotine salts following subcutaneous administration in male rats. *Drug Test Anal*. (2022) 15:1099–106. doi: 10.1002/dta.3363

111. Etter JF. Explaining the effects of electronic cigarettes on craving for tobacco in recent quitters. *Drug Alcohol Depend*. (2015) 148:102–8. doi: 10.1016/j.drugalcdep.2014.12.030

112. Etter JF. Characteristics of users and usage of different types of electronic cigarettes: findings from an online survey. *Addiction*. (2016) 111:724–33. doi: 10.1111/add.13240

113. Talih S, Balhas Z, Eissenberg T, Salman R, Karaoghlanian N, El Hellani A, et al. Effects of user puff topography, device voltage, and liquid nicotine concentration on electronic cigarette nicotine yield: measurements and model predictions. *Nicotine Tob Res*. (2015) 17:150–7. doi: 10.1093/ntr/ntu174

In vivo low-dose phase-contrast CT for quantification of functional and anatomical alterations in lungs of an experimental allergic airway disease mouse model

Christian Dullin[1,2,3]*, Jonas Albers[1,4], Aishwarya Tagat[5], Andrea Lorenzon[6], Lorenzo D'Amico[7,8], Sabina Chiriotti[9], Nicola Sodini[7], Diego Dreossi[7], Frauke Alves[1,2,10], Anna Bergamaschi[9] and Giuliana Tromba[7]

[1]Institute for Diagnostic and Interventional Radiology, University Medical Center Göttingen, Göttingen, Germany, [2]Translational Molecular Imaging, Max-Plank-Institute for Multidisciplinary Sciences, Göttingen, Germany, [3]Diagnostic and Interventional Radiology, University Hospital Heidelberg, Heidelberg, Germany, [4]European Molecular Biology Laboratory, Hamburg Unit c/o Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany, [5]Department of Urology, University Hospital of Saarland, Homburg, Germany, [6]Innova S.p.A., Trieste, Italy, [7]Elettra-Sincrotrone Trieste S.C.p.A., Trieste, Italy, [8]Department of Physics, University of Trieste, Trieste, Italy, [9]PSD Detector Science and Characterization Group, Paul Scherrer Institute, Villingen, Switzerland, [10]Department of Haematology and Medical Oncology, University Medical Center Göttingen, Göttingen, Germany

Introduction: Synchrotron-based propagation-based imaging (PBI) is ideally suited for lung imaging and has successfully been applied in a variety of *in vivo* small animal studies. Virtually all these experiments were tailored to achieve extremely high spatial resolution close to the alveolar level while delivering high x-ray doses that would not permit longitudinal studies. However, the main rationale for performing lung imaging studies *in vivo* in small animal models is the ability to follow disease progression or monitor treatment response in the same animal over time. Thus, an *in vivo* imaging strategy should ideally allow performing longitudinal studies.

Methods: Here, we demonstrate our findings of using PBI-based planar and CT imaging with two different detectors—MÖNCH 0.3 direct conversion detector and a complementary metal-oxide-semiconductor (CMOS) detector (Photonics Science)—in an Ovalbumin induced experimental allergic airway disease mouse model in comparison with healthy controls. The mice were imaged free breathing under isoflurane anesthesia.

Results: At x-ray dose levels below those once used by commercial small animal CT devices at similar spatial resolutions, we were able to resolve structural changes at a pixel size down to 25 μm and demonstrate the reduction in elastic recoil in the asthmatic mice in cinematic planar x-ray imaging with a frame rate of up to 100 fps.

Discussion: Thus, we believe that our approach will permit longitudinal small animal lung disease studies, closely following the mice over longer time spans.

KEYWORDS

phase contrast, allergic airway disease models, lung function, longitudinal experiments, x-ray dose

1 Introduction

Apart from classical radiography and computed tomography in which the image content depends on tissue-specific differences in the attenuation of the incident x-ray beam, other techniques such as dark-field and phase-contrast imaging have evolved, which exploit the wave nature of the x-rays. Lung imaging is ideally suited for these new methods as the strong gradients in the refractive index between lung tissue and aerated lung areas causes strong scattering and larger phase shifts. Since the phase shift of waves does not change their intensities, different strategies to record this information have been implemented such as grating-based phase-contrast, edge-illumination, and propagation-based imaging (PBI) to only name the most used ones (1). Among them, PBI can be considered to be the simplest approach as no additional optical elements are required, which in addition renders PBI very dose-efficient. Synchrotron propagation-based phase-contrast CT imaging has proven to provide strongly elevated soft tissue contrast especially in lung imaging (2). Thus, PBI was successfully employed in a large variety of such studies (2–4). However, studies utilizing small animal lung disease models in mice or rats typically aimed to resolve the structural alterations in the lung at a scale of few micrometers (3, 5, 6). Even though PBI is known to enable a significant reduction in the required x-ray dose (2), the x-ray dose levels needed in such experiments would still prevent using these techniques for longitudinal studies—scanning the animals multiple times over a longer period. In addition, to deal with the breathing motion during data acquisition, the laboratory animals are often ventilated using intubation or even tracheotomy. Not only can forced ventilation shadow important disease-related features of the altered breathing patterns, but it also prevents or complicates a longitudinal scanning approach.

Although limitations and differences exists in human asthma, experimental allergic airway disease (EAAD) mouse models remain widely used. Among them, the ovalbumin (OVA)-induced mouse model is often applied. Based on the schedule and the amount of OVA used, an acute phase can be generated characterized by allergy-mediated responses such as varying degrees of inflammation, mucus production, accumulation of eosinophiles, and impairment of lung function (7). Here, we use the OVA-induced mouse model first introduced by (8), which results in severe asthma-related symptoms starting at 24 hours after the last intranasal OVA challenge. Especially in EAAD, in which symptoms quickly change over time—like in human asthma—longitudinal non-invasive imaging is of great importance to be able to monitor disease progression and treatment response while keeping the numbers of required laboratory mice to a minimum. In such mice with EAAD, we already successfully applied in vivo x-ray-based imaging to address changes in lung function over time and in response to therapy (9). However, the level of structural changes that can be assessed by classical in vivo CT imaging remains limited due to x-ray dose constrains.

While lung imaging is typically a domain for x-ray-based methods in which the air acts as a natural negative contrast agent, magnetic resonance imaging (MRI) can also be applied. Since MRI does not require the use of ionizing radiation, it would be ideally suited for longitudinal studies. However, as the lung posses a strong challenge for imaging in MRI, the achieved spatial resolution is typically lower than that used in CT, especially in small animal models. Bianchi et al. (10), for instance, achieved longitudinal imaging in an OVA mouse model over 75 days utilizing three MRI imaging sessions. However, only a 2D axial protocol was used, which might pose a limitation in cases of heterogeneous distribution of lung inflammation.

Here, we realized low-dose PBI planar and CT imaging in free-breathing anesthetized mice to monitor both functional and structural alterations in EAAD mice mimicking human asthma in comparison with healthy controls. Moreover, by validating our findings with classical in vivo microCT and subsequent histological analysis, we confirmed the correlation of the obtained functional parameter with the severity of the EAAD model.

2 Methods

2.1 Ethics

All animal in vivo procedures were performed at the animal facility of the University of Trieste, Italy, as per the guidelines of the European Directive and Italian ethical laws (n°600/2018-PR, Direzione Generale della Sanita Animale e dei Farmaci Veterinari). The study is reported according to the ARRIVE guidelines.

2.2 Experimental allergic airways disease mouse model

Thirteen female BALB/c mice (6 weeks of age) were purchased from Charles River (Italy). All animals were housed in a controlled environment with a regular 12-hour dark/light cycle at 22°C and maintained with ad libitum food and water. An experimental allergic airway disease (EAAD) mouse model was generated to mimic severe allergic asthma (SAA, N = 6) (8). The mice were sensitized on days 0 and 14 intraperitoneally (i.p.) with a mixture of 50 μg ovalbumin (OVA) and 0.5% of aluminum hydroxide adjuvant (Alum) in a volume of 200 μl phosphate-buffered saline (PBS), as well as intranasally (i.n.) with 50 μg OVA in 25 μl PBS. The mice were further treated i.n. with a solution of 250 μg OVA/50 μl PBS/mouse at days 28, 29, 30, and 33 in order to trigger a severe allergic reaction. Healthy age and gender matched mice served as controls (CN, N = 7). The workflow including the setup of EAAD as well as the schedule of the performed imaging and lung extraction is summarized in Figure 1.

2.3 Experimental setup—Synchrotron

The experiments were carried out at the SYRMEP (Synchrotron Radiation for Medical Physics) beamline of the ELETTRA synchrotron radiation facility (Trieste, Italy). Two types of acquisitions were performed: (a) 2D cinematic x-ray imaging of the lung region (without rotation) and (b) CT scans. For the latter, 540° scans with a rotation speed of 18° s^{-1} were performed with a total scanning time of 30 s, using monochromatic x-rays with an energy of 22 keV in combination with a sample-to-detector distance of 1.5 m. We compared the performance of two different detectors: a

FIGURE 1
Animal model: experimental allergic airway disease mouse model was induced by first sensitization of the immune system to Ovalbumin (OVA) at day 0 and day 14 with an i.p. injection of 50 μg OVA+0.5% Alum in combination with an intranasal application of 50 μg OVA. At days 28, 29, 30, and 33, 250 μg OVA was applied i.n. to provoke an allergic reaction, which based on prior experiments reaches the peak of inflammation about 48 hrs later. Thus, both planar PBI and PBI CT imaging were performed at days 35 and 36. Directly after the non-invasive imaging, the mice were euthanized, and an *ex vivo* broncho alveolar lavage was performed, followed by extraction of the lung, filling it with 1-ml formalin via a syringe and processing it for subsequent specimen CT and histology. In six mice, EAAD was induced. Seven aged and gender-matched healthy mice served as controls.

MÖNCH 0.3 (Figure 2A, D1) direct conversion detector (11) and a complementary metal-oxide-semiconductor (CMOS) detector (Photonics Science, Figure 2A, D2) combined with a 13 μm thick Gadox:Tb scintillator (12). Table 1 summarizes the main characteristics of the two detectors. The two detectors are similar concerning the quantum efficiency and the pixel size. However, the larger field of view (FOV) of the CMOS camera facilitates the alignment of the sample for the tomographic measurements, while the FOV of MÖNCH is smaller than the diameter of the mice. Thus, it was not possible to target the same area within the lung in all mice. The fast frame rate of MÖNCH allows following the breathing dynamics of the mouse. In this experiment, the images taken with MÖNCH were acquired at 1,000 fps and binned to 100 fps by summing up 10 consecutive images to increase the photon statistic. The mice were mounted in an upright position on a custom-made mouse holder (Figure 2A, M) at the center of the hexapod (Figure 2A, H) and anesthetized by inhalation via a mask with 2% isoflurane, 1 L/h oxygen for the duration of the scan. An isoflurane anesthesia device (Figure 2B, A) was used, comprised of a standard isoflurane vaporizer, an oxygen flask, and tubing to either connect to a box to initialize the anesthesia or to the mask of the mouse to maintain the anesthesia. The exhausts of both mask and box were connected to charcoal filters. A flow pump was used to remove the exhausted anesthesia gas. A calibrated ionization chamber (Figure 2B, I) was employed to track the entrance dose of each acquisition. The setup was completed with a lamp and a camera to monitor the anesthesia (Figures 2A, C). A Styrofoam cube colored partially in black was attached to the abdomen of the mouse with double-sided tape. Motion tracking of this cube was used to adjust the breathing frequency to about 0.7 Hz by slightly adjusting the isoflurane concentration in the vaporizer.

TABLE 1 Comparison of the main characteristics for the two detectors used in this study.

	Photonic science CMOS	MÖNCH 0.3
QE 22 keV	23%	21%
FOV (mm^2)	63.9 x 63.9	10 x 10
Pixel size (μm)	31.2	25
Frame rate (fps)	40	1,300

2.4 Sample preparation, broncho-alveolar lavage, and histological analysis

Following x-ray-based imaging, the mice were euthanized on days 35 and 36 by an overdose of isoflurane, followed by cervical dislocation. Bronchoalveolar lavage (BAL) was performed by gently washing the airways thrice with 250 μl of 2% fetal calf serum (FCS)/PBS after revealing and cannulating the trachea. The volumes obtained in the three washing steps were pooled and rewashed with the same buffer, and the cells post-recovery were counted in a hemocytometer. In total, 3 x 10^4 cells were used for cytospins, followed by Giemsa staining for differential cell counting using an Axioskop 2 microscope as previously described (13).

The lungs were excised post-BAL filling them with 1-ml formalin and processed according to the standard procedure for formalin-fixed and paraffin-embedded (FFPE) tissue specimens, which were cut into 2-μm thick slices. The tissue sections were
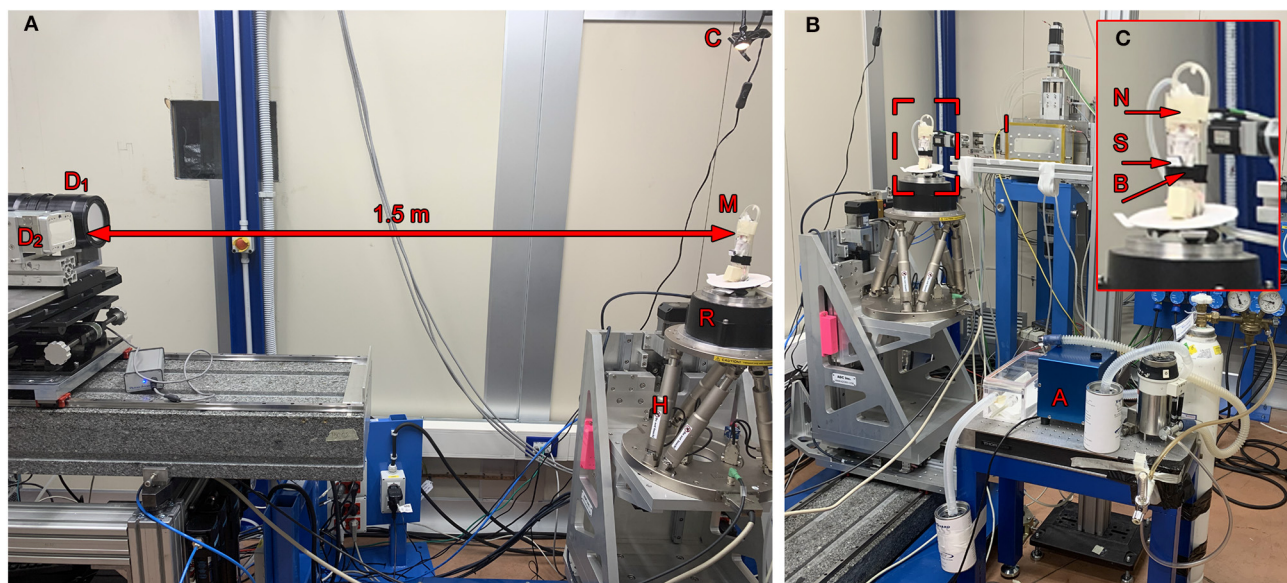
FIGURE 2
Setup: **(A)** The anesthetized mouse was mounted upright in a custom-made mouse holder (M) on top of the rotator (R), which itself could be adjusted using the hexapod (H). In a sample-to-detector distance of 1.5 m, the Photonics detector (D₁) and the MÖNCH detector (D₂) were mounted on a stage to place either one or the other in the light path. A camera and a lamp were mounted on top of the mouse holder to provide a live feedback of the breathing frequency (C). **(B)** The mouse was kept under anesthesia but free-breathing using a mask (N) taped to the top of the holder. An isoflurane anesthesia system (A) comprised of a vaporizer, an oxygen flask, an exhaust pump, an induction box, and two charcoal filters was used to keep the mouse under anesthesia during the acquisition and to adjust the breathing frequency. A calibrated ionization chamber (I) was utilized to monitor the entrance dose. **(C)** Shows a close-up of the mouse holder with a mounted mouse, the anesthesia mask (N) including tubes for gas influx and exhaust, a flexible belt (B) to restrain the mouse in upright position as well as a Styrofoam block (S) used for the optical breathing frequency measurement. The disk on the basement of the holder was used to provide a control coil up and down of the pipes.

deparaffinized, and hematoxylin-eosin (H&E) or Alcian blue—periodic acid schiff (PAS) staining was performed as previously described (7). An Axioskop 2 (Carl Zeiss Microscopy GmbH) microscope in combination with a Leica DC 100 camera was used for visualization of the stained sections. Sections were scored in a blinded fashion from 0 to 4 to address the amount of infiltrating immune cells seen in H&E (designated HE-score) and the amount of mucus-producing goblet cells seen in the PAS staining (designated PAS-score), with 0 = no infiltrating cells and no mucus, to 4 = high infiltrating cell count and strong mucus production. The scoring was done by five independent readers, two with more than seven years of experience in the analysis of experimental allergic airway disease (EAAD) mouse models. The presented values represent the median of the five scores for each individual specimen.

2.5 Image quality parameter

To assess image quality in the reconstructed CT data sets, contrast-to-noise ratio (CNR), coefficient of variation in soft tissue (COV), and edge sharpness by the means of the full width half maximum of a differentiated edge profile (FWHM) have been calculated. For the calculation of CNR, we used the following equation: $CNR = |g_1 - g_2| / \sqrt{0.5\left(\sigma_1^2 + \sigma_2^2\right)}$, with g_i the average gray value in a supposedly homogeneous region and σ_i the standard deviation of the gray values in the same region. COV was measured by $COV = g/\sigma$, with g the average gray value and σ the variance in a soft tissue region, respectively.

2.6 Software

Phase retrieval and reconstruction of the synchrotron data were performed with SYRMEP Tomo Project (STP) (14) using the TIE-HOM algorithm (15) with a delta-to-beta ratio of 1,000 prior reconstruction using filtered back projection. To suppress artifacts from the truncated projection data (local area tomography), the projection data were expanded to 150% of its width by zero padding. The resulting data were reconstructed with a pixel size of 31.2 μm on a 1,417 x 1,417 matrix and 25 μm on a 400 x 400 matrix for the Photonics Science CMOS and MÖNCH detector, respectively. In the case of the Photonics Science CMOS detector, the projection data were cropped to roughly the size of the x-ray beam and are therefore smaller than the entire matrix of 2,048 x 2,048 px.

To analyze the data, custom-made scripts in Python 3 were employed including Matplotlib 3.7.1 https://matplotlib.org/, Scikit-Image, seaborn 0.12.2 https://seaborn.pydata.org/, Numpy 1.23.5 https://numpy.org/, and statannot 0.2.3 https://pypi.org/project/statannot/. The generation of representative images was done using Fiji https://fiji.sc/. 3D rendering was

conducted using VGStudio 2.2 https://www.volumegraphics.com/en/products/vgstudio.html. Statistical analysis was performed with Statannot https://pypi.org/project/statannot/ and Prism https://www.graphpad.com using a single-sided Welch's t-test with unequal variations. P-values are denoted on the figures as (ns = not significant p > 0.05, *p < 0.05, **p < 0.01, ***p < 0.001, ***p < 0.0001).

3 Results

3.1 Retrieving functional information from planar cinematic radiography—Photonics detector

Relevant temporal parameters of the breathing patterns can be extracted from simple planar cinematic radiography movies. In order to not override those features by using forced ventilation, we performed the experiments with free-breathing anesthetized mice. We gently strapped the mice to a custom made holder in upright position. After anesthesia was induced in an induction chamber (Figure 2B, A) using 5% isoflurane, the mouse was moved to the holder, and anesthesia was maintained at approximately 2% isoflurane in 1 L/h oxygen. Due to the live feedback from the optical camera in combination with the Styrofoam cube attached to the abdomen of the mouse, the breathing rate was adjusted to about 0.7 Hz. Movies at 40 fps were recorded for 30 s. Figure 3A shows one frame of such a movie. A region of interest was placed over the lung (red), and a reference region (blue) was placed next to the mouse holder only containing air. The average x-ray attenuation in the measurement region subtracted by the average x-ray attenuation in the reference region is shown in Figure 3B (blue) for a healthy control mouse. Clearly, breathing events can be seen and isolated. The modulation of the function at the baseline is attributed to fluctuations in the x-ray beam, noise, and the motion of the beating heart. A rolling-average filtered version of this curve [Figure 3B (red)] was used to identify end inspiration peaks (red asterisk) and begin-of-inspiration points (black dots). For comparison, Figure 3D shows the same data for a SAA mouse. The data of all detected fully recorded breathing events were overlaid at the peak position to fit two functions of the type $f(t) = c_0 * 1/(c_1 * t) + c_2$ to the inspiration (black curve) and expiration phase (red curve), for the CN and SAA mouse respectively. Figures 3C, E shows the obtained results. The breathing intervals, which we aimed to keep constant at 1,400 ms, demonstrated a large variation. This was attributed to the fact that changes in the isoflurane concentration modulate the breathing frequency with a large delay. Moreover, the isoflurane concentration was adjusted manually before preparing for the acquisition, which started ∼5 mins later. During that time, the concentration could not be modified. However, the measured breathing frequency did not showed significant differences between the asthmatic (SAA, $N = 3$) and the control (CN, $N = 4$) group. The area under the breathing curve (AUC) was significantly enlarged, and the fall constant was significantly smaller in SAA (Figure 3F), while the raise constant showed the tendency to be smaller but without being significantly different.

3.2 Retrieving functional information from planar cinematic radiography—MÖNCH detector

Another set of eight mice (CN = 4, SAA = 4) was imaged with the MÖNCH detector to assess lung function, binning the data to a frame rate of 100 fps. Setting up the mice for the experiment and adjustment of the anesthesia was done in the same way as described above. Since the MÖNCH detector has an active area of 1 x 1 cm^2, the FOV was smaller than the diameter of the mouse (Figure 4A). This condition, plus the limited motion options of our mouse holder and the fact that conversion of the data was not instant, rendered positioning of the mouse (respectively, the lung region) challenging. Thus, different parts of the lung were imaged for each mouse, and a statistical analysis was impossible. However, due to the much higher frame rate of the MÖNCH detector in comparison with the Photonics detector, the breathing curves are better sampled (Figure 4B). Since no background region only consisting of air could be used, a baseline correction based on rolling-average filtering was applied. Single breathing events were isolated as described above (Section 3.1), and the data were overlaid to fit a function in the expiratory phase. Figures 4C, D demonstrates exemplary data from one control (Figure 4C) and one asthmatic mouse (Figure 4D). As shown in Figure 4C, the expiration phase shows a change in curvature at around 10 ms. Thus, we fitted a sum of two Gaussians to the curve $f(t) = c_2 * exp(-c_1 * t^2) + c_4 * exp(-c_3 * t^2) + c_0$. The here-observed change in the curvature of the x-ray attenuation in the expiration phase is most likely attributed to the fact that the expiration, which is a consequence of the elastic recoil, gets dampened at a certain time by the diaphragm. Here, we define this point as a minima in the 3rd derivative of the fitted function (Figures 4C, D right panel). Since asthma is known for a loss of elastic recoil of the lung, we expected a larger AUC and a later onset of the dampening effect, which can be seen in Figure 4D. For the visualization of this important physiological parameter in mice, the high frame rate of 100 fps of the MÖNCH detector was essential.

3.3 Visualization of structural changes within the lung by CT—Photonics detector

Subsequent to the planar acquisitions, CT scans were acquired. Three consecutive scans over 180° were performed to account for breathing motion in the data. However, the mouse also showed some bulk movement due to imperfections in the mouse holder. Therefore, performing retrospective gating was not possible. Thus, only the data covering the first 180° were used. These data spanning 10 s (Figure 5) contained in average 15 breathing events. Partially due to the isoflurane anesthesia, such events are characterized by a sharp inspiration peak and a longer resting phase. However, we found that this motion did not resulted in major artifacts using classical filtered back projection for reconstruction. Already at the pixel size of 31.2 μm, the motion of the lung tissue caused by the heart beat was apparent. Thus, finally, every second projection was used that resulted in reconstructed data with less apparent motion artifacts. Thus, the entire acquisition took 30 s with an approximate dose of 170 mGy, while the reconstructed data were achieved
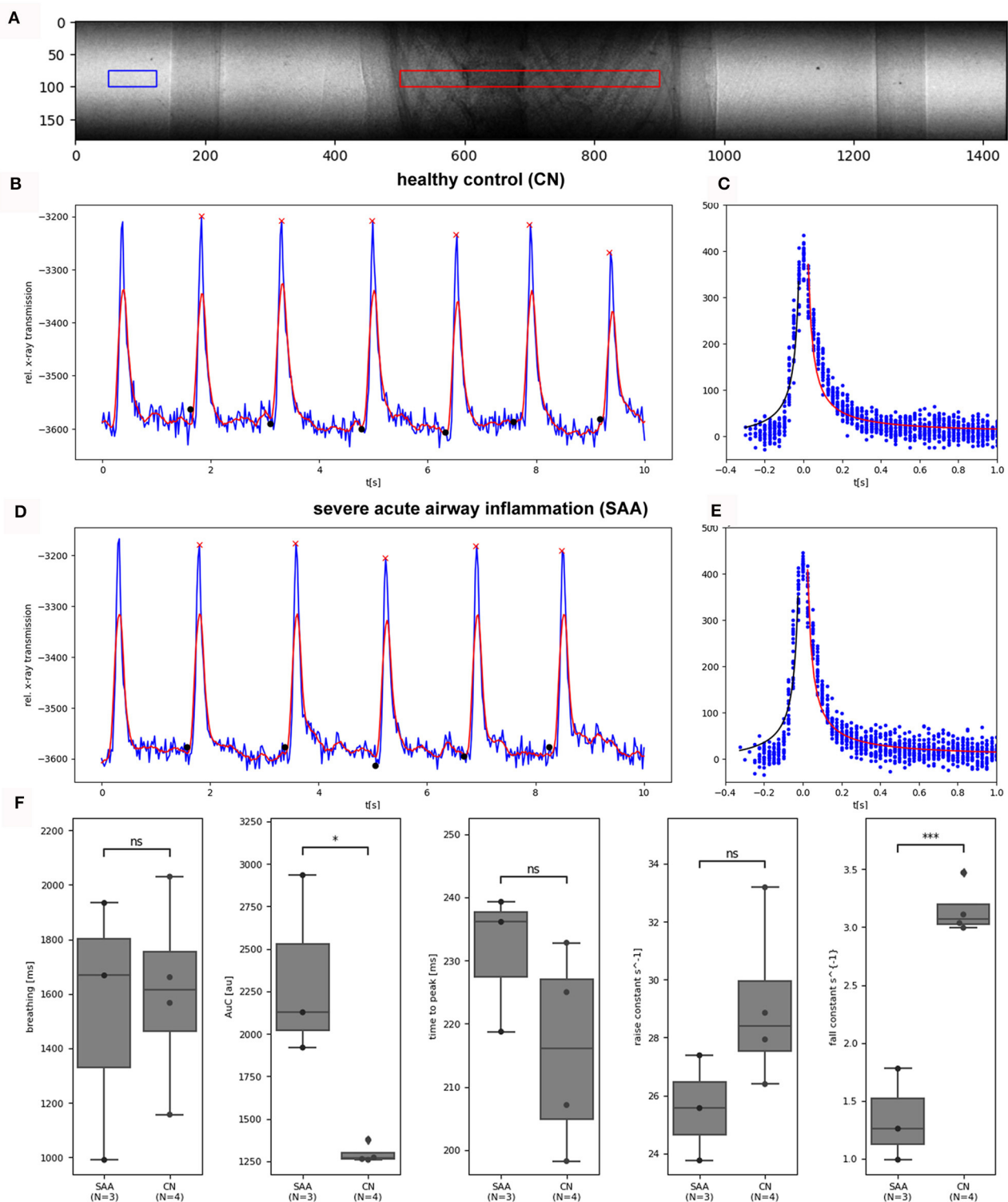
FIGURE 3

Analysis photonics detector planar: **(A)** shows a frame of an acquired planar image sequence (31.2 μm pixel size, 40 fps, 1500 frames, dose rate 2.8 mGy/s). The average x-ray transmission was measured in a ROI (red) over the lung. To normalize for fluctuations in the incident x-ray beam, the average of a reference region (blue) was subtracted from the data. **(B)** Shows the resulting curve for a healthy control mouse (blue) for 10 s, demonstrating 10 breathing events. After the application of a rolling average filter (red curve), end-inspiration peaks (red asterisks) and breathing intervals (black dots) were identified. **(C)** The overlaid breathing (setting the individual peak position to 0) was used to fit one 1/x functions in the inspiration phase (black) and in the expiration phase (red). **(D, E)** Show the same data for an exemplary SAA-mouse. **(F)** Shows the comparison of the retrieved parameter between the two groups (SAA and CN). The breathing rates showed large variations but no significant differences between the groups as intended. The area under the curve (AUC) of the breathing peaks was significantly larger and the fall constant of the expiration phase was significantly smaller in SAA, which is in agreement with previous studies and published data (9). The "raise" time constant is reduced in SAA however not significantly. N represents the number of mice analyzed per group. *$p < 0.05$, ***$p < 0.001$.
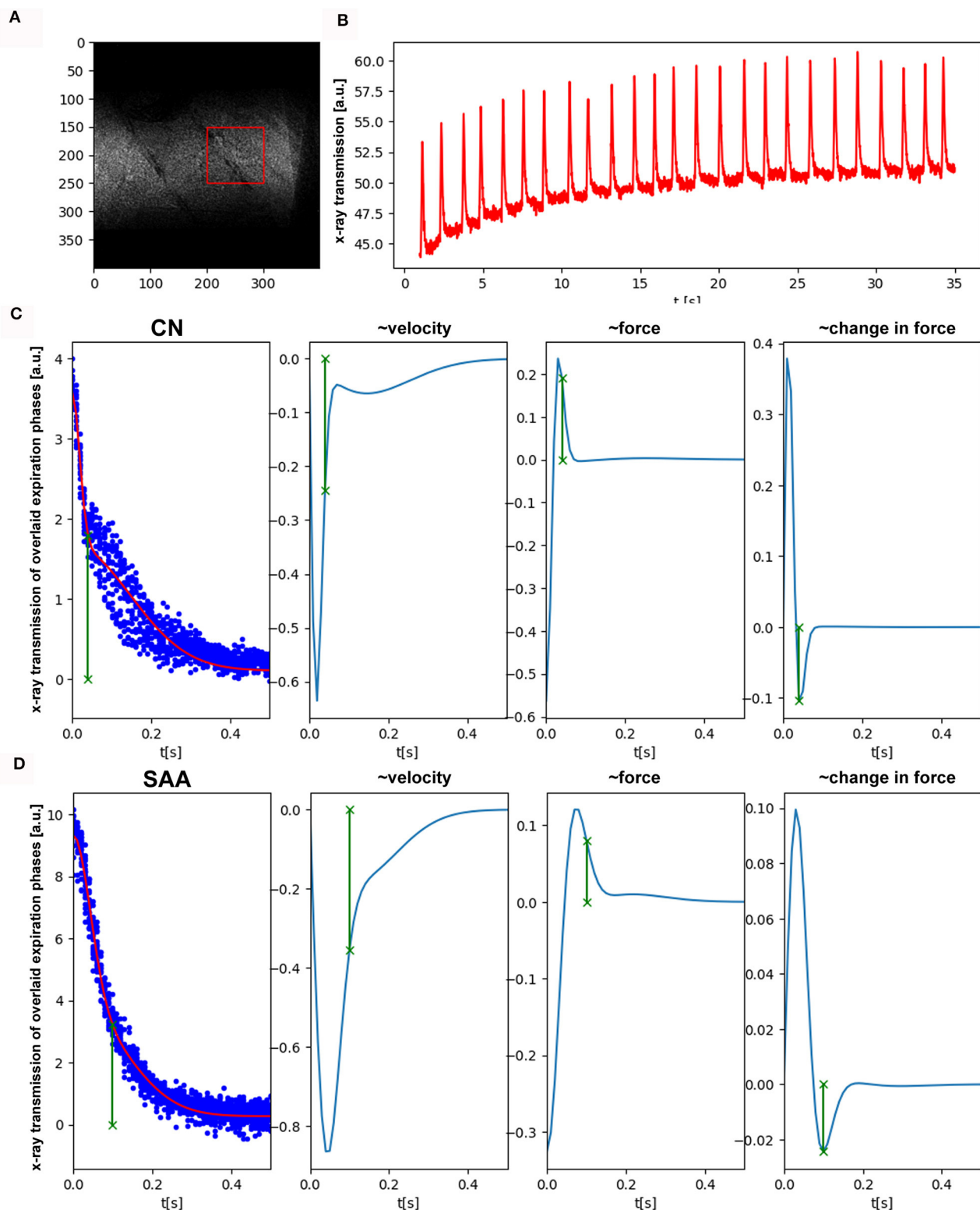
FIGURE 4
Functional analysis by planar imaging with the MÖNCH detector: **(A)** One frame of the acquired image sequence (25 μm pixel size, 100 fps, 3,900 frames, dose rate 2.8 mGy/s). Due to the smaller FOV of 1x1 cm^2, only a region over the lung (red) but no reference region was selected. **(B)** Shows the average x-ray transmission of this region over time. Clearly, breathing events are visible as well as a build-up of the baseline. **(C, D)** Depict individual expiration phases overlaid, after baseline correction and peak detection (blue dots). We fitted a sum of two Gaussian to the data to account for the different phases in expiration. The other columns showing the 1st, 2nd, and 3rd derivation of this fitted function. In comparison with the control **(C)** in SAA **(D)**, the time at which the elastic recoil of the lung gets dampened by the diaphragm (minima in the 3rd derivative, green) is much larger. This explains the increase in the AUC seen with the Photonics detector and is in agreement with the known loss of elastic recoil in asthma.

FIGURE 5
Anatomical analysis using the Photonics detector: The CT's of the lungs were acquired over 180° with a total acquisition time of 20 s resulting in an entrance dose of roughly 60 mGy. **(A)** Shows a cross-section through a scan of an asthmatic and **(B)** of a healthy mouse. Clearly strong signs of inflammation (consolidated areas and mucus plugging of the airways) can be seen in **(A)** (asterisk). **(C, D)** Show volume renderings of the same mice. Bone is rendered in gray and the aerated parts of the lung in pink/black. In the asthmatic example **(C)** a largely non-aerated part close to the main bronchi (asterisk) can be seen corresponding to the regions indicated in **(A)**. The analysis of the aerated lung fraction in the lungs **(E)** shows a significant decrease in SAA in comparison with the healthy controls. *$p < 0.05$.

with roughly 60 mGy (30 mGy if accounted for the use of every 2nd projection). Figures 5A, B shows two example reconstructed slices of (Figure 5A) an asthmatic mouse and (Figure 5B) a healthy control. Clearly, strong signs of inflammation causing mucus plugging can be seen in Figure 5A (asterisk). Rendering the aerated lung of the same two examples in 3D (Figures 5C, D, pink) also revealed non aerated areas in the asthmatic mouse (Figure 5C, asterisk). In order to quantify these alterations in the lung structure, six non-overlapping ROI's were placed within each lung. This ensured the reproducibility of the measurement in different samples using ImageJ. The mean value between air and soft tissue was chosen as a threshold to discriminate between them, and the threshold was kept constant for all ROI's and samples. The values of the six ROI's per sample were averaged, and the results are represented in Figure 5E, demonstrating a significant reduction in the air volume fraction in the asthmatic mice.

The reconstructed CT data sets showed a CNR of 15.0 ± 1.9, a COV of approximately 7E-4 and a edge sharpness of FWHM 156 ± 13 μm measured between soft tissue and air within a larger bronchi.

3.4 Visualization of structural changes within the lung by CT—MÖNCH detector

The same acquisition protocol described in 3.3 was used to acquire lung CTs in combination with the MÖNCH detector. The acquired RAW data were binned to 100 fps. Also, in this case,

only the first 180° rotation out of the three acquired was used, resulting in the same dose of 170 mGy for the entire acquisition and 60 mGy for the used data. Figures 6A, B shows that due to the smaller pixel size of 25 μm, the MÖNCH detector revealed that, also in the asthma case (Figure 6A), the observed mucus plugging did not completely block the airways, as seen in the data acquired with the Photonics detector. In analogy to the planar acquisitions, the problems with positioning the scan in combination with the smaller FOV resulted in vastly different lung regions being depicted and did not permit a statistical analysis. However, the 3D renderings of the aerated lung regions (Figure 6C asthmatic and Figure 6D healthy) demonstrate that alterations in the lung structure can be depicted *in vivo* at this comparable low x-ray dose rate.

The reconstructed CT data sets showed a CNR of 10.9 ± 8.3, a COV of approximately 1E-3 and a edge sharpness of FWHM 72 ± 6 μm measured between soft tissue and air within a larger bronchi.

3.5 Dose measurement

In addition to the calibrated ionization chamber (Figure 2B, I), thermoluminescence dosimeter crystals (TLD) were utilized. Four TLD's were wrapped in parafilm to shield them from moisture and were enclosed within the chest of a dead mouse. This mouse corpse was exposed to the same imaging protocols, and an average dose of 84 mGy for planar imaging and a dose of 168 mGy for the CT acquisitions was measured with the dose rate being 2.8 mGy/s.

3.6 Validation of the pathological features of the EAAD mouse model by broncho alveolar lavage and histology

Following the imaging, the mice were euthanized on days 35 or 36 by an overdose of isoflurane, followed by cervical dislocation, and a BAL was performed. Figure 7A shows the subset of cells that were counted in (SAA) and in the controls (CN). In the control, most of the cells were identified as macrophages (Figure 7B), while in SAA predominately eosinophiles were found (Figure 7C).

Figures 7D, E shows examples for HE-stained sections of (Figure 7D) a control and (Figure 7E) an asthmatic mouse. Airways are indicated as (aw). Several regions of densely packed immune cells (eosinophiles) can be seen in HE-stained sections of SAA mice (Figure 7E, asterisk). The PAS staining revealed elevated mucus production in lungs of the SAA mice (Figure 7H, arrow head) compared with the healthy control, where almost no mucus production was observed (Figure 7G). Both findings point to a strong acute airway inflammation. The severity of both effects was scored between $0 =$ no accumulation of immune cells/no mucus production up and $4 =$ large amount of immune cells/strong mucus production by five independent readers in a blind manner. The average of these scoring results is presented in Figures 7F, I, together with the eosinophile counts, which confirmed the successful induction of the EAAD model in all mice used in this experiment.

4 Discussion

We presented our results of low-dose *in vivo* imaging of spontaneously breathing mice under isoflurane anesthesia using both a Photonics detector with a 31.2 μm and the MÖNCH detector with 25 μm pixel size in mice from an EAAD model and healthy controls. We demonstrated that the breathing motion can be analyzed by planar imaging, revealing significant differences between asthmatic and healthy mice. Moreover, we validated that, despite the breathing motion, CT reconstructions can be performed showing disease-specific structural features such as consolidations and mucus plugging. This was all done at a low-dose rate of 2.8 mGy/s using a monochromatic x-ray beam of 22 keV and propagation-based imaging with a sample-to-detector distance of 1.5 m. The used high frame rate of 100 fps and the comparable small pixel size of 25 μm of the MÖNCH detector allowed to measure the later onset of the dampening effect by the diaphragm in asthmatic mice and revealed patterns of mucus plugging and consolidation only partially blocking smaller airways in contrast to results obtained with the Photonics detector.

The effects of x-ray dose exposure in preclinical *in vivo* CT imaging are critically discussed (16). However, the topic seems of rather low interest, as in many animal studies, the x-ray dose is not even reported. It is certainly true that in mice the same dose constraints like in clinical CT imaging do not apply, which mainly take the induction of an elevated cancer risk by the CT examination over the entire lifespan of the patient into account. Moreover, it has been demonstrated that mice can recovery from sublethal dose levels (17). In mice, a LD50 to lung damage (lethal dose for 50% of a mouse population) between 9 and 12 Gy was reported by (18). In comparison, the LD50 for whole-body exposure of

humans is in the range of 4.5 Gy (19). This further suggests that mice can withstand slightly higher dose levels than humans. Nevertheless, Nowosielska et al. (20) showed anti-neoplastic and immunostimulatory effects of low-dose whole-body irradiation of 10 sessions between 0.01 and 0.1 Gy each. Thus, in preclinical studies, the applied x-ray dose should neither interfere with the well-being of the animal nor with experimental conditions to justify *in vivo* imaging. The ability to perform longitudinal *in vivo* imaging is especially suited to study progression and monitor therapy response while at same time limit the animal numbers. In this perspective, Vande Velde et al. (16) reported no signs of radiotoxicity for biweekly scans with a dose of 1.64 Gy over 5 weeks as assessed by comparison of structural biomarkers and collagen deposition. While this is certainly encouraging for longitudinal studies, no information regarding long-term effects, changes in blood parameters, inflammatory responses, and induction of fibrosis or cancer are presented. In addition, the time point of analysis of collagen deposition directly after the last imaging session seems to be too early to exclude long-term radiation-induced pulmonary fibrosis. Foster and Ford showed no signs of radiotoxicity in tumor-bearing mice for CT scans up to 0.5 Gy (21), and for repeated mouse lung CT scans of 0.27 Gy, they found no indication of adverse effects (22). Current preclinical lung imaging studies by Zaw Thin et al. (23) and Pennati et al. (24) reported a x-ray dose for retrospective gated lung CT of 927 mGy and 4 mins total scanning time resulting in 3D data sets with 50 μm resolution. While in the latter publication, although the x-ray dose was not measured, but as they used the same device with the settings, we speculate that the x-ray dose was similar. For comparison, Yarnold (25) used a x-ray dose of 2 Gy per day for fractionated clinically applied x-ray radiation therapy for patients with breast cancer. Thus, ~1 Gy per imaging session seems rather high or at least not compatible with longer longitudinal experiments. Of course, these are protocols optimized to calculate more than one phase of the breathing cycle in contrast to only the expiration phase as we did in this study. Moreover, the applied x-ray dose strongly differs between different CT devices based on the used combination of x-ray tube and detector and the targeted spatial resolution. However, we reached 31.2 or 25 μm voxel size with only a x-ray dose of 60 mGy, so less than 1/15 of the dose delivered in the previously reported *in vivo* CT scans using commercial small animal CT systems. The large variety of x-ray dose levels in the presented studies, in combination with the reported immunostimulatory effects at very low doses, suggest that mouse lung CT imaging can be safely performed a higher dose levels, but further evidence is needed to estimate the impact of longitudinal CT imaging on the experimental conditions. In this perspective, the here-presented PBI-based approach could be a valuable alternative to x-ray attenuation-based CT techniques.

PBI has been shown to provide excellent soft tissue contrast at low-dose levels especially in lung imaging (2). Therefore, PBI has already been used for *in vivo* imaging at a large variety of setups, resolutions, and dose rates (3–5) to name a few. To our knowledge, our experiment does not use the lowest spatial resolution of *in vivo* small animal CT studies, but the lowest dose rate and a comparable fast acquisition. Ford et al. (26) describe the fundamental dose limits in classical attenuation-based microCT systems. In comparison with the results achieved here is challenging as the image content in our case is related to the phase shift of the
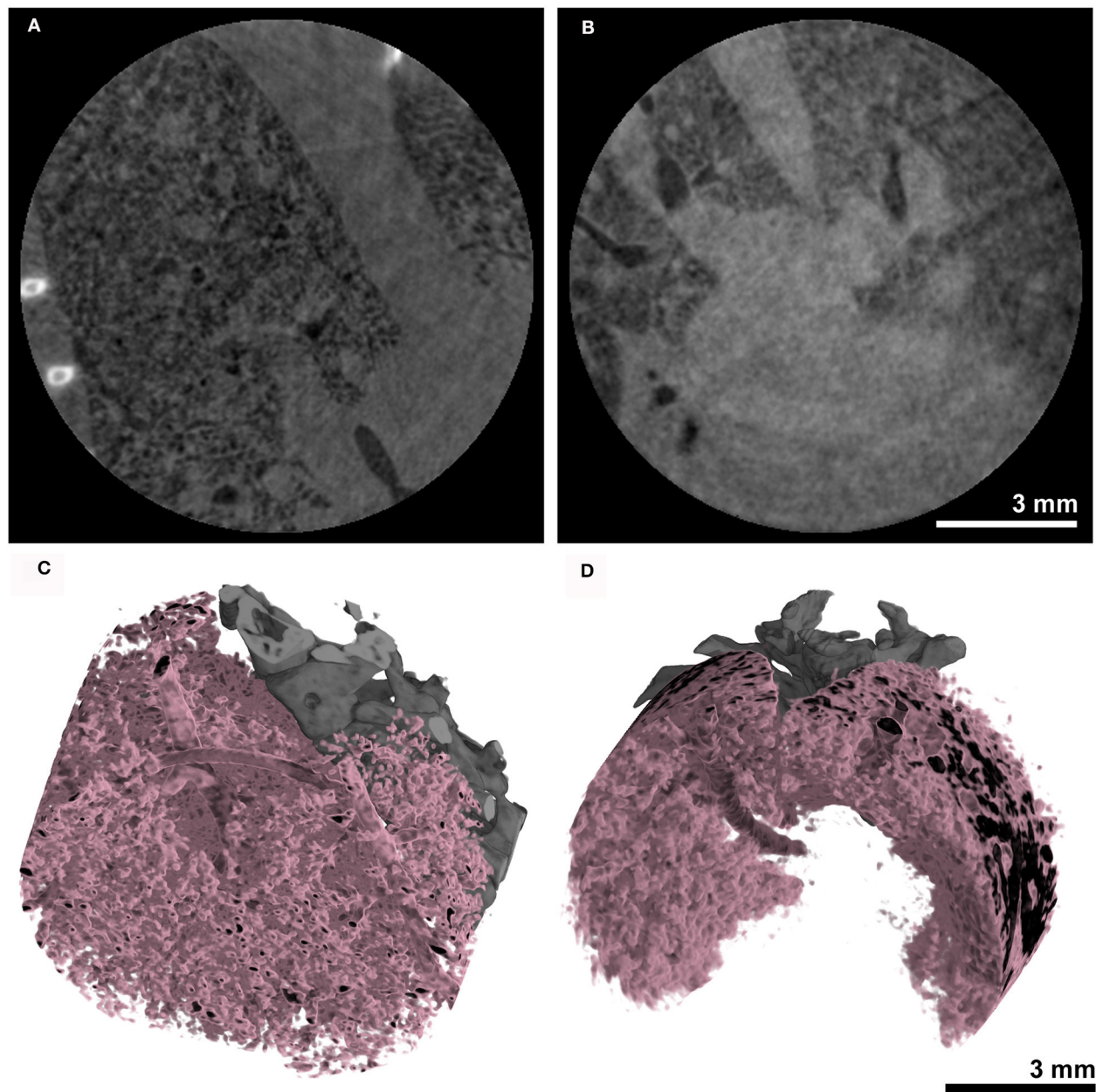
FIGURE 6
Anatomical analysis using the MÖNCH detector: The CT's of the lungs were acquired over 180° with a total acquisition time of 20 s resulting in an entrance dose of ∼60 mGy. The cross-sections of **(A)** an asthmatic and **(B)** a healthy subject show signs of inflammation only in the asthmatic mouse. The 3D renderings in **(C, D)** demonstrate that the aerated lung can be resolved *in vivo* at a pixel size of 25 μm at this low x-ray dose.

x-ray beam rather than its attenuation. Moreover, the applied single distance TIE-HOM algorithm (15) is known to introduce a low-pass filtering effect. Therefore, we can consider the measured edge sharpness of approximately 160 and 70 μm of the Photonics and MÖNCH detector respectively as effective resolution to account for this effect. The comparison is further complicated by the differences in dose measurements. While Ford et al. (26) used the entrance dose, we measured the dose within the mouse. Putting all these limitations aside, we surpass the theoretical limit of a classical microCT. This finding is in agreement with the theoretical

consideration of noise in propagation-based imaging by Gureyev et al. (27) and is further supported by the experimental data provided by Kitchen et al. (2) in performing ex vivo phase-contrast lung CT imaging in newborn rabbits.

Functional readouts in animal lung disease models are mostly done with plethysmography. However, plethysmography is either invasive or, in terms of non-invasive whole-body plethysmography depends on strong and complex assumptions about flow rate, humidity, and temperature in the chamber. It has been shown by for instance Hülsmann et al. (28) that the precision of the
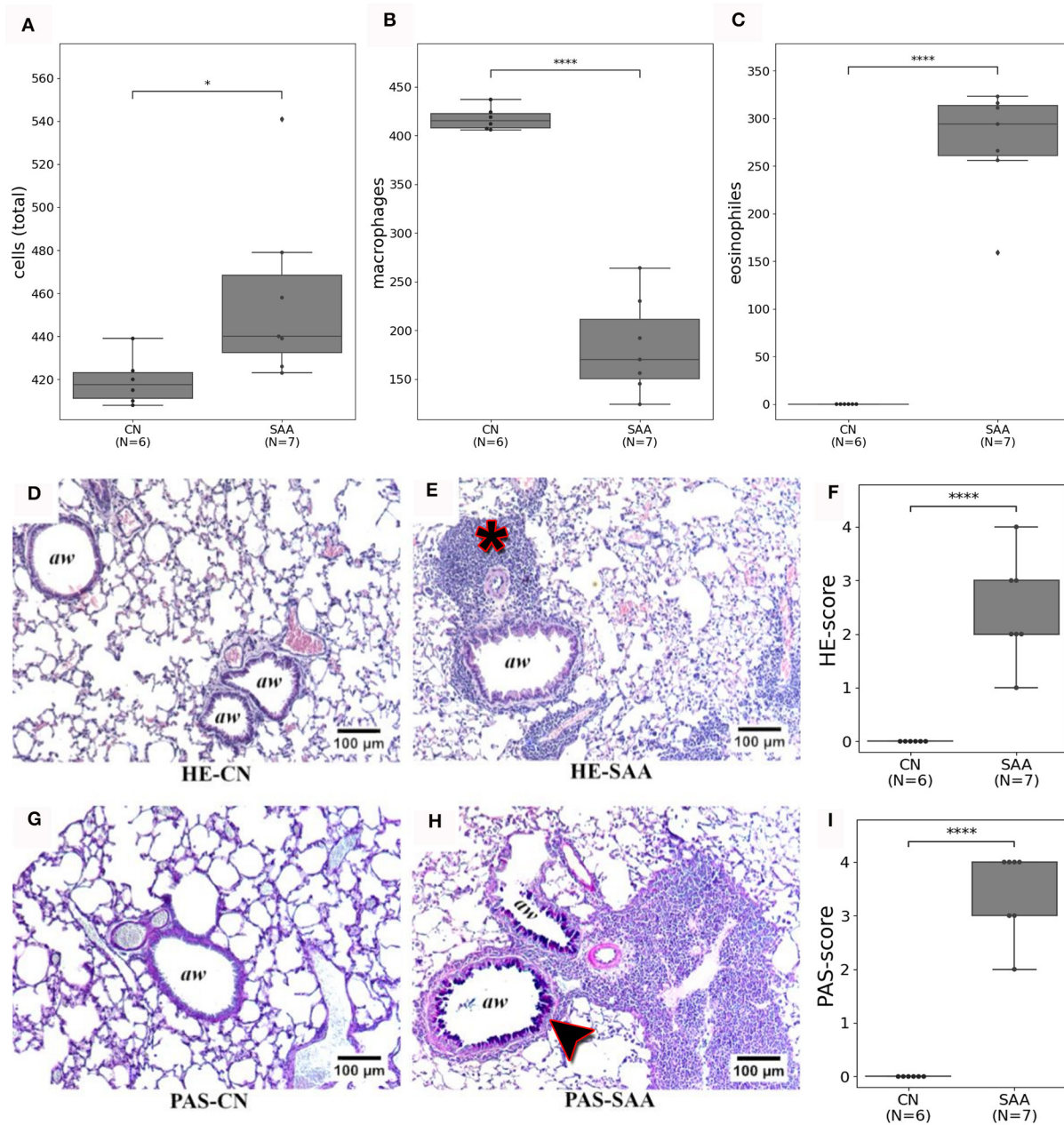
FIGURE 7
Validation of the *in vivo* imaging findings by BAL and histology: **(A)** the number of cells from the specific BAL's that were counted. **(B)** Shows that these cells are predominately macrophages in CN. While **(C)** shows that in SAA most of the cells were eosinophiles. **(D, E)** Depict HE stained sections of a healthy lung (HE-CN) and the lung of an asthmatic mouse (HE-SAA) respectively. The HE-SAA specimen shows a strong infiltration of immune cells close to the airways (aw, asterisk). **(F)** The HE score revealed a significant stronger immune reaction in lungs of SAA compared to CN. **(G, H)** Depicts Periodic Acid Schiff (PAS) stained from a healthy lung (PAS-CN) and an asthmatic lung (PAS-SAA) respectively. Only in PAS-SAA strong mucus production was found within the airways (purple staining, arrow head). **(I)** PAS score shows a significant increase in mucus production for SAA compared to CN (aw denotes airways).

derived parameters is at least questionable. In comparison, we demonstrated that with a rather simple approach of cinematic planar phase contrast imaging, we derived significant differences in lung function parameters that allowed discriminating between asthmatic and healthy mice. Even more importantly, we were able to demonstrate that with a temporal resolution of 100 fps, the phase of elastic recoil of the lung and the dampening effect of the

diaphragm could reliably be detected, a characteristic feature of asthmatic mice (29), that so far we were only able to detect with optical cameras at the same high frame rate of 100 Hz (30). The concept of extraction breathing motion by analyzing the average x-ray transmission in a region of interest typically on top of the lung-diaphragm interface is a well-known approach introduced for instance by Ford and Bartling (31, 32). We were building on

this approach, which is typically used to suppress motion artifacts in lung CT by performing retrospective gating but can also be employed to retrieve parameter characterizing lung function (9).

It has to be mentioned that the here-presented pilot study has a couple of shortcomings and therefore a potential for further improvement. We are convinced that the analysis of mouse lung disease models shows that free-breathing under anesthesia is preferable to intubation and ventilation for multiple reasons, such as avoiding damage to the trachea during intubation that would prevent a longitudinal application of the approach and "overriding" the pathological breathing pattern with a forced ventilation pattern. Nevertheless, in the free-breathing approach, maintaining a constant and defined breathing frequency in mice is challenging and would require the development of an anesthesia system with an active feedback cycle. In this study, we were not able to keep the breathing constant for all animals. A future improved setup will therefore contain an optical system to target the lung region and a mouse holder that can be centered on top of the rotator.

To fully exploit our approach of low-dose phase-contrast lung imaging, more synchrotron sites need to be equipped with animal facilities and need to consider that ability to perform fractionated/longitudinal experiments in her way of approving experiments.

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 animal study was approved by Ministero della Salute, Direzione Generale della Sanita Animale e dei Farmaci Veterinari N° 600/2018-PR. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

CD: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. JA: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. AT: Investigation, Writing – original draft, Conceptualization. AL: Investigation, Writing – review & editing. LD'A: Data curation, Validation, Writing – review & editing. SC: Investigation, Methodology, Writing – review & editing. NS: Conceptualization, Methodology, Writing – review & editing. DD: Conceptualization, Methodology, Writing – review & editing. FA: Conceptualization, Writing – original draft, Writing – review & editing. AB: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing. GT: Investigation, Methodology, Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

SC and AL are now working for companies but worked for public research organizations during duration of the experiment. AL was employed by Innova S.p.A.

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.

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.

References

1. Momose A. X-ray phase imaging reaching clinical uses. *Physica Medica*. (2020) 79:93–102. doi: 10.1016/j.ejmp.2020.11.003

2. Kitchen MJ, Buckley GA, Gureyev TE, Wallace MJ, Andres-Thio N, Uesugi K, et al. CT dose reduction factors in the thousands using X-ray phase contrast. *Sci Rep*. (2017) 7:15953. doi: 10.1038/s41598-017-16264-x

3. Bayat S, Fardin L, Cercos-Pita JL, Perchiazzi G, Bravin A. Imaging regional lung structure and function in small animals using synchrotron radiation phase-contrast and K-edge subtraction computed tomography. *Front Physiol*. (2022) 13:825433. doi: 10.3389/fphys.2022.825433

4. Lewis RA, Yagi N, Kitchen MJ, Morgan MJ, Paganin D, Siu KKW, et al. Dynamic imaging of the lungs using x-ray phase contrast. *Phys Med Biol*. (2005) 50:5031. doi: 10.1088/0031-9155/50/21/006

5. Fardin L, Broche L, Lovric G, Mittone A, Larsson A, Bravin A, et al. Within-breath dynamics of atelectrauma during mechanical ventilation studied by *in vivo* 4D microscopy in injured rabbit lung. *Eur Resp Soc*. (2020) 56:1933. doi: 10.1183/13993003.congress-2020.1933

6. Lovric G, Mokso R, Arcadu F, Vogiatzis Oikonomidis I, Schittny JC, Roth-Kleiner M, et al. Tomographic *in vivo* microscopy for the study of lung physiology at the alveolar level. *Sci Rep*. (2017) 7:12545. doi: 10.1038/s41598-017-12886-3

7. Dullin C, Larsson E, Tromba G, Markus AM, Alves F. Phase-contrast computed tomography for quantification of structural changes in lungs of asthma mouse models of different severity. *J Synchrotron Radiat*. (2015) 22:1106–11. doi: 10.1107/S1600577515006177

8. Nabe T, Zindl CL, Jung YW, Stephens R, Sakamoto A, Kohno S, et al. Induction of a late asthmatic response associated with airway inflammation in mice. *Eur J Pharmacol*. (2005) 521:144–55. doi: 10.1016/j.ejphar.2005.08.015

9. Dullin C, Markus MA, Larsson E, Tromba G, Hülsmann S, Alves F. X-Ray based Lung Function measurement-a sensitive technique to quantify lung function in allergic airway inflammation mouse models. *Sci Rep*. (2016) 6:36297. doi: 10.1038/srep36297

10. Bianchi A, Ozier A, Ousova O, Raffard G, Crémillieux Y. Ultrashort-TE MRI longitudinal study and characterization of a chronic model of asthma in mice: inflammation and bronchial remodeling assessment. *NMR Biomed*. (2013) 26:1451–9. doi: 10.1002/nbm.2975

11. Ramilli M, Bergamaschi A, Andrae M. Brückner M, Cartier S, Dinapoli R, et al. Measurements with MÖNCH, a 25 μm pixel pitch hybrid pixel detector. *J Instrumentat*. (2017) 12:C01071. doi: 10.1088/1748-0221/12/01/C01071

12. Van Eijk CW. Inorganic scintillators in medical imaging. *Phys Med Biol*. (2002) 47:R85. doi: 10.1088/0031-9155/47/8/201

13. Biffi S, Dal Monego S, Dullin C, Garrovo C, Bosnjak B, Licha K, et al. Dendritic polyglycerolsulfate near infrared fluorescent (NIRF) dye conjugate for non-invasively monitoring of inflammation in an allergic asthma mouse model. *PLoS ONE*. (2013) 8:e57150. doi: 10.1371/journal.pone.0057150

14. Brun F, Pacilè S, Accardo A, Kourousias G, Dreossi D, Mancini L, et al. Enhanced and flexible software tools for X-ray computed tomography at the Italian synchrotron radiation facility. *Elettra Fundamenta Informaticae*. (2015) 141:233–43. doi: 10.3233/FI-2015-1273

15. Paganin D, Mayo SC, Gureyev TE, Miller PR, Wilkins SW. Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *J Microsc*. (2002) 206:33–40. doi: 10.1046/j.1365-2818.2002.01010.x

16. Vande Velde G, De Langhe E, Poelmans J, Bruyndonckx P, d'Agostino E, Verbeken E, et al. Longitudinal *in vivo* microcomputed tomography of mouse lungs: No evidence for radiotoxicity. *Am J Physiol-Lung Cell Mol Physiol*. (2015) 309:L271–9. doi: 10.1152/ajplung.00098.2015

17. Mole R. Quantitative observations on recovery from whole body irradiation in mice part II. *Br J Radiol*. (1957) 30:40–6. doi: 10.1259/0007-1285-30-349-40

18. Safwat A, Nielsen OS, El-Badawy S, Overgaard J. Effect of radiation dose rate and cyclophosphamide on pulmonary toxicity after total body irradiation in a mouse model. *Int J Radiat Oncol Biol Phys*. (1996) 34:85–91. doi: 10.1016/0360-3016(95)02078-0

19. López M, Martín M. Medical management of the acute radiation syndrome. *Rep Pract Oncol Radiother*. (2011) 16:138–46. doi: 10.1016/j.rpor.2011.05.001

20. Nowosielska EM, Cheda A, Wrembel-Wargocka J, Janiak MK. Anti-neoplastic and immunostimulatory effects of low-dose X-ray fractions in mice. *Int J Radiat Biol*. (2011) 87:202–12. doi: 10.3109/09553002.2010.519422

21. Foster WK, Ford NL. Investigating the effect of longitudinal micro-CT imaging on tumour growth in mice. *Phys Med Biol*. (2010) 56:315. doi: 10.1088/0031-9155/56/2/002

22. Ford NL, Martin EL, Lewis JF, Veldhuizen RA, Drangova M, Holdsworth DW. *In vivo* characterization of lung morphology and function in anesthetized free-breathing mice using micro-computed tomography. *J Appl Physiol*. (2007) 102:2046–55. doi: 10.1152/japplphysiol.00629.2006

23. Zaw Thin M, Moore C, Snoeks T, Kalber T, Downward J, Behrens A. Micro-CT acquisition and image processing to track and characterize pulmonary nodules in mice. *Nat Protoc*. (2023) 18:990–1015. doi: 10.1038/s41596-022-00769-5

24. Pennati F, Leo L, Ferrini E, Sverzellati N, Bernardi D, Stellari FF, et al. Micro-CT-derived ventilation biomarkers for the longitudinal assessment of pathology and response to therapy in a mouse model of lung fibrosis. *Sci Rep*. (2023) 13:4462. doi: 10.1038/s41598-023-30402-8

25. Yarnold J. Changes in radiotherapy fractionation–breast cancer. *Br J Radiol*. (2018) 92:20170849. doi: 10.1259/bjr.20170849

26. Ford N, Thornton M, Holdsworth D. Fundamental image quality limits for microcomputed tomography in small animals. *Med Phys*. (2003) 30:2869–77. doi: 10.1118/1.1617353

27. Gureyev TE, Nesterets YI, Kozlov A, Paganin DM, Quiney HM. On the "unreasonable" effectiveness of transport of intensity imaging and optical deconvolution. *JOSA A*. (2017) 34:2251–60. doi: 10.1364/JOSAA.34.002251

28. Hülsmann S, Khan A, Hagos L, Hindermann M, Nägel T, Dullin C. Evaluation of a mechanical lung model to test small animal whole body plethysmography. *Sci Rep*. (2021) 11:17099. doi: 10.1038/s41598-021-96355-y

29. Markus MA, Borowik S, Reichardt M, Tromba G, Alves F, Dullin C. X-ray-based lung function measurement reveals persistent loss of lung tissue elasticity in mice recovered from allergic airway inflammation. *Am J Physiol Lung Cell Mol Physiol*. (2017) 313:L763–71. doi: 10.1152/ajplung.00136.2017

30. Svetlove A, Albers J, Hülsmann S, Markus MA, Zschüntzsch J, Alves F, et al. Non-invasive optical motion tracking allows monitoring of respiratory dynamics in dystrophin-deficient mice. *Cells*. (2022) 11:918. doi: 10.3390/cells11050918

31. Ford NL, Detombe S, Wheatley A, Holdsworth DW, Drangova M. Optimization of retrospective respiratory-gated high speed micro-CT for free-breathing mice. *Med Imag*. (2006) 6143:24–31. doi: 10.1117/12.653993

32. Bartling SH, Stiller W, Grasruck M, Schmidt B, Peschke P, Semmler W, et al. Retrospective motion gating in small animal CT of mice and rats. *Invest Radiol*. (2007) 42:704–14. doi: 10.1097/RLI.0b013e318070dcad

Capturing COPD heterogeneity: anomaly detection and parametric response mapping comparison for phenotyping on chest computed tomography

Silvia D. Almeida[1,2,3,4]*, Tobias Norajitra[1,2], Carsten T. Lüth[5,6], Tassilo Wald[1,6], Vivienn Weru[7], Marco Nolden[1,8], Paul F. Jäger[5,6], Oyunbileg von Stackelberg[2,9], Claus Peter Heußel[2,9,10], Oliver Weinheimer[2,9], Jürgen Biederer[2,9,11,12], Hans-Ulrich Kauczor[2,9] and Klaus Maier-Hein[1,2,4,6,8]*

[1]Division of Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany, [2]Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), Heidelberg, Germany, [3]Medical Faculty, Heidelberg University, Heidelberg, Germany, [4]National Center for Tumor Diseases (NCT), NCT Heidelberg, A Partnership Between DKFZ and Heidelberg University Medical Center, Heidelberg, Germany, [5]Interactive Machine Learning Group (IML), German Cancer Research Center (DKFZ), Heidelberg, Germany, [6]Helmholtz Imaging, German Cancer Research Center (DKFZ), Heidelberg, Germany, [7]Division of Biostatistics, German Cancer Research Center (DKFZ), Heidelberg, Germany, [8]Pattern Analysis and Learning Group, Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany, [9]Diagnostic and Interventional Radiology, Heidelberg University Hospital, Heidelberg, Germany, [10]Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik at University Hospital, Heidelberg, Germany, [11]Faculty of Medicine, University of Latvia, Riga, Latvia, [12]Faculty of Medicine, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

Background: Chronic obstructive pulmonary disease (COPD) poses a substantial global health burden, demanding advanced diagnostic tools for early detection and accurate phenotyping. In this line, this study seeks to enhance COPD characterization on chest computed tomography (CT) by comparing the spatial and quantitative relationships between traditional parametric response mapping (PRM) and a novel self-supervised anomaly detection approach, and to unveil potential additional insights into the dynamic transitional stages of COPD.

Methods: Non-contrast inspiratory and expiratory CT of 1,310 never-smoker and GOLD 0 individuals and COPD patients (GOLD 1–4) from the COPDGene dataset were retrospectively evaluated. A novel self-supervised anomaly detection approach was applied to quantify lung abnormalities associated with COPD, as regional deviations. These regional anomaly scores were qualitatively and quantitatively compared, per GOLD class, to PRM volumes (emphysema: PRM^{Emph}, functional small-airway disease: PRM^{fSAD}) and to a Principal Component Analysis (PCA) and Clustering, applied on the self-supervised latent space. Its relationships to pulmonary function tests (PFTs) were also evaluated.

Results: Initial t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization of the self-supervised latent space highlighted distinct spatial patterns, revealing clear separations between regions with and without emphysema and air trapping. Four stable clusters were identified among this latent space by the PCA and Cluster Analysis. As the GOLD stage increased, PRM^{Emph}, PRM^{fSAD}, anomaly score, and Cluster 3 volumes exhibited escalating trends, contrasting with a decline in Cluster 2. The patient-wise anomaly scores significantly differed across GOLD

stages ($p < 0.01$), except for never-smokers and GOLD 0 patients. In contrast, PRM^{Emph}, PRM^{fSAD}, and cluster classes showed fewer significant differences. Pearson correlation coefficients revealed moderate anomaly score correlations to PFTs (0.41−0.68), except for the functional residual capacity and smoking duration. The anomaly score was correlated with PRM^{Emph} ($r = 0.66$, $p < 0.01$) and PRM^{fSAD} ($r = 0.61$, $p < 0.01$). Anomaly scores significantly improved fitting of PRM-adjusted multivariate models for predicting clinical parameters ($p < 0.001$). Bland−Altman plots revealed that volume agreement between PRM-derived volumes and clusters was not constant across the range of measurements.

Conclusion: Our study highlights the synergistic utility of the anomaly detection approach and traditional PRM in capturing the nuanced heterogeneity of COPD. The observed disparities in spatial patterns, cluster dynamics, and correlations with PFTs underscore the distinct − yet complementary − strengths of these methods. Integrating anomaly detection and PRM offers a promising avenue for understanding of COPD pathophysiology, potentially informing more tailored diagnostic and intervention approaches to improve patient outcomes.

1 Introduction

Chronic obstructive pulmonary disease (COPD) remains a major global health burden, ranking as the third leading cause of mortality worldwide (1). Characterized by progressive airflow limitation, COPD predominantly arises from prolonged exposure to harmful airborne particles, particularly in individuals with a history of smoking (2, 3). Early detection and accurate phenotyping of COPD are paramount, as timely intervention, including smoking cessation and appropriate treatments, may slow disease progression and improve patient outcomes. Although pulmonary function testing (PFT) and, in particular, gold standard spirometry, play a central role in COPD diagnosis, its ability to detect early-stage disease and reliably characterize its heterogeneity remains limited (4, 5).

Recent advances in imaging technology, particularly computed tomography (CT), have been instrumental to gain insights into COPD pathophysiology, by quantifying emphysema and small-airway disease. Several emphysema quantification methods are based on measuring the relative area of the lungs below a specific Hounsfield unit (HU) threshold, i.e., low attenuation areas (LAA), on inspiratory CT scans, showing significant correlations with pulmonary function test (PFT) parameters (6−8).

While large airways down to the first sub-segmental generations are clearly visible on CT, smaller, more distal subsegmental airways and respiratory bronchioles cannot be detected unless becoming more conspicuous due to mucus retention or peribronchial inflammation. However, small airways disease can be indirectly assessed and quantified with CT scans in expiration, when small airway obstruction results in air trapping (8). Several methods, including measuring the LAA below −856 HU (LAA-856) on expiratory scans, have been proposed for evaluating air trapping (8). However, the optimal CT-based method to assess small-airway disease remains a subject of ongoing debate. In recent years, parametric response mapping (PRM) has emerged as a novel approach to phenotyping COPD by utilizing both inspiration and expiration CT scans (9). PRM allows differentiation between emphysematous and non-emphysematous air trapping regions. For small-airway disease, PRM identifies lung areas with densities greater than or equal to −950 HU on inspiration CT and less than −856 HU on expiration CT. However, PRM's current method for small-airway disease assessment focuses on slight dynamic density changes within each voxel and may not fully consider emphysema's potential contribution to air trapping assessment. As it is a mutually exclusive voxel-wise method, each voxel is assigned exclusively to either the emphysematous or non-emphysematous category, potentially oversimplifying the intricate interplay between emphysema and small-airway disease. Additionally, its method's dependence on fixed thresholds introduces limitations in capturing variations across diverse patient populations, while the reliance on registration may be susceptible to specific methods and anatomical variations. Moreover, the sensitivity to CT-protocol variations poses a challenge in ensuring consistent and reproducible results across different imaging settings.

With advances in deep learning (DL) based artificial intelligence (AI), several algorithms have been developed to target the direct interpretation of CT scans, without the need to pre-define COPD features of interest. While the majority relies on supervised learning (10−14), self-supervised (15) and unsupervised methods (16) have been gaining a lot of attention. Their ability to capture complex disease heterogeneity without the need for explicit annotation, makes them particularly advantageous in scenarios where manual labeling may be challenging or subjective. Moreover, these methods have the potential to discover novel disease subtypes or manifestations that may not be predefined in the training dataset, enhancing their adaptability to the diverse and evolving nature of COPD.

In particular, Almeida et al. (15) introduced a self-supervised DL anomaly-detection approach that identifies COPD lung regions as anomalies, reflecting the varied manifestations of COPD across different phenotypes, including large and small airway disease,

parenchymal scars, and emphysema. The method harnesses informative latent representations and a generative model to identify deviations from the distribution of normal-appearing lung regions of individuals without airflow obstruction. These deviations are then presented as a lung anomaly map with anomaly scores. Importantly, it has demonstrated its value in distinguishing normal individuals from those with COPD and predicting lung function decline (17).

While being a promising tool for the assessment of the severity of COPD, the regional anomaly scores derived from this method have not been further explored. Hence, our study aims to compare the quantitative and spatial relationships between PRM functional small airway disease and emphysema measurements to regional anomaly scores. This can further provide critical insights into COPD heterogeneity and implications for personalized patient care.

2 Materials and methods

The objective of our study, depicted in Figure 1, was to perform a quantitative and qualitative comparison of spatial relationships between traditional (1) Parametric response mapping (PRM) volumes and the (2) Self-Supervised Anomaly Detection method (15, 17). Our hypothesis posits that the anomaly scores derived from self-supervised learning align with PRM volumes, and their associations with clinical metrics are comparable. To enhance our understanding of what the self-supervised method learns, we also employ (3) Principal Component Analysis (PCA) & Clustering on the self-supervised representation level. This approach identifies stable clusters, allowing us to compare them with the aforementioned two methods.

2.1 Study cohort

In the present study, we utilized the Genetic Epidemiology of COPD (COPDGene) study. The COPDGene study (ClinicalTrials.gov Identifier NCT00608764) recruited never-smoker controls and current and former smokers, who had a smoking history of ≥10 pack-years. The enrollment period was conducted between 2008 and 2011, targeting individuals aged 45–80 years.

Comprehensive assessments were performed, including paired chest CT scans during inspiration (Insp) and expiration (Exp), pulmonary function tests (PFT), and questionnaire evaluations. Specific details regarding the CT acquisition protocol can be found in Almeida et al. (17).

Ethical approval was obtained, and written consent was acquired from all participants after the study protocol received approval from the respective clinical center's review board. The inclusion and exclusion criteria were previously described in Almeida et al. (15, 17).

In accordance with the established protocol defined in Almeida et al. (15, 17), the dataset was partitioned into distinct sets for training, evaluation and testing.
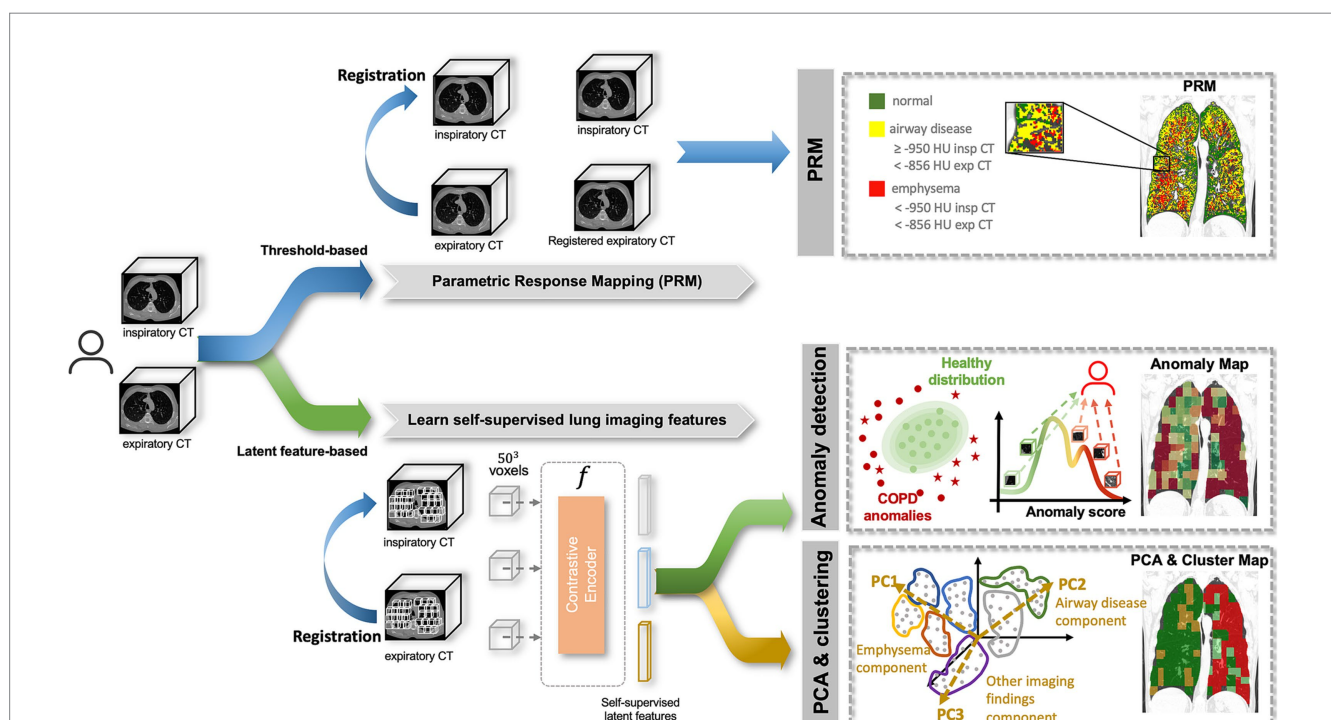
FIGURE 1
Methodology pipeline. A subject is composed of paired inspiratory and expiratory CT scans which are analyzed by 3 methods: (1) Parametric response mapping (PRM), which relies on the spatial alignment of the two scans and defines emphysema and airway disease areas based on strict thresholds; (2) Anomaly detection, which attributes region-wise anomaly scores based on the distribution of normal lung features, defined by a self-supervised contrastive method; (3) Principal Component Analysis (PCA) & Clustering, which applies dimensionality reduction (PCA) to the latent features from the self-supervised contrastive method to find stable clusters. All three methods produce lung maps, which were then compared qualitatively and quantitively.

2.2 Parametric response mapping

Parametric response mapping (PRM) was applied to the paired Insp and Exp CT scans as described in the original work (9) of the test set. This process categorized the lung parenchyma into functional small-airway disease (PRM^{fSAD}), emphysema (PRM^{Emph}), and normal lung (PRM^{Normal}). To minimize the contribution of airways and vessels, specific minimum and maximum attenuation values were defined for both scans. PRM^{Emph} was defined by voxels between $-1{,}000$ HU and -950 HU in the inspiratory CT and between $-1{,}000$ HU and -856 HU in the expiratory CT scan. PRM^{fSAD} was defined by voxels between -950 HU and -810 HU in the inspiratory CT and $-1{,}000$ HU and -856 HU in the expiratory CT. Lastly, PRM^{Normal} was defined by voxels between -950 HU and -810 HU in the inspiratory CT and -856 and -500 HU in the expiratory CT scan. The top section of Figure 1 illustrates this step.

Following the PRM analysis, the subsequent sections detail the application of (2) Anomaly Detection and (3) PCA & Clustering, both involving two common steps: quantitative pre-processing and extraction of self-supervised latent features.

2.3 Pre-processing

2.3.1 Quantitative pre-processing

After segmentation of the lung parenchyma, 3D Regions of Interest (ROIs) of size $50 \times 50 \times 50$ voxels (covering >70% of the lung parenchyma) were extracted from paired Insp and registered Exp (ExpR) CT scans, with 20% patch-overlapping, as previously described in Almeida et al. (15, 17).

2.3.2 Extraction of self-supervised latent features

Following the extraction of 3D patches, the objective was to create a representation per patch that encapsulated relevant information related to its imaging features. This representation vector, of size 1×512, was later employed for both (1) Anomaly Detection and (2) PCA & Clustering. The self-supervised learning strategy described in Almeida et al. (15, 17) was employed, where a subset of the COPDGene subjects was used to train a self-supervised contrastive model, without using any labels. This model is based on the idea of learning representations that maximize the agreement between differently augmented views of the same region via a contrastive loss, i.e., attracting regions that look similar and repel the ones that do not. Based on this, informative latent representations were generated per patch, as illustrated in Figure 1.

2.4 Visualization of the latent space: t-distributed stochastic neighbor embedding

For understanding the information conveyed in the self-supervised latent features of each region and for visualization purposes, t-Distributed Stochastic Neighbor Embedding (t-SNE) (18) was applied as a non-linear dimensionality reduction technique. t-SNE is a common dimensionality reduction technique that maps high-dimensional data into a lower-dimensional space while preserving the pairwise similarities between data points. In order to

preserve better global structure, non-standard affinity methods were initialized with the evaluation set, and later transformed to the test set. This was implemented using the openTSNE package (19).

2.5 Anomaly detection

Subsequently, having the patch-level latent representations, the anomaly detection model was applied, as illustrated in Figure 1.

The model aimed to quantify the degree to which a specific region or patient deviates from the pre-defined "normality," as defined in Almeida et al. (17). This normative baseline was defined by lung regions with less than 1% emphysema from individuals without airflow obstruction (never-smoker controls and GOLD0). The distribution of these "normal/healthy" latent features, as derived from the self-supervised contrastive approach, serves as the reference for the model. Region, i.e., patch-wise, anomaly scores are computed using the negative log-likelihood. Patient-level anomaly scores are subsequently obtained by aggregating scores from all regions. Further details can be found in Almeida et al. (15, 17).

For visualization purposes of the anomaly lung map, min-max normalization was applied to the anomaly scores, corresponding to the 5th and 95th percentiles of the dataset.

2.6 Principal component analysis and clustering

Principal Component Analysis (PCA) & Clustering served as a direct comparative approach to the anomaly detection model, as both operate on the self-supervised latent features of the test set. These representations were obtained by a region-similarity approach, which grouped similar patterns together (whether region- or intensity-wise). The objective with the PCA & Clustering was to explore and identify clusters of regions sharing similar characteristics within the informative latent features, and directly compare it to the anomaly detection method and to the reference PRM volumes.

2.6.1 PCA on the self-supervised latent features

PCA was applied to the latent vectors of each region of the test set, in order to mitigate multicollinearity along the 512 features. The eigenvectors of the covariance matrix were analyzed. Then to reduce the dimension space, the dataset was projected onto the first few uncorrelated principal components, representing dominant eigenvectors of the covariance matrix. The optimal number of Principal Components (PCs) was decided based on Horn's parallel analysis (20), through a Scree Plot, and based on the Kaiser Criterion.

2.6.2 Clustering the PCA

Once determined the subset of PCs to retain, cluster analysis was conducted on the dataset. Various clustering methods, including K-means and Gaussian finite mixture model-based methods were compared using Silhouette, Davies-Bouldin and Calinski and Harabasz scores (21). The Silhouette score aimed for maximization, providing a measure of the separation among different clusters. Davies-Bouldin, targeted for minimization, assessed the similarity of each cluster to its next closest neighbor. Calinski and Harabasz, also

targeted for minimization, estimated the cohesion and separation of points within a cluster, based on the distance between cluster centroids.

Based on these scores, the optimal clustering method and number of clusters were determined and applied for further analysis. Further information is provided in the Supplementary materials. However, it's important to note that these clusters were primarily generated for comparison purposes. They serve as a reference point to assess the efficacy of the anomaly detection method, since both employ the same self-supervised latent features.

2.7 Statistical analysis

Comparisons among PRM volumes, Patient-wise Anomaly Scores and each class of Anomaly Region Clusters were made according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) stages using the Jonckheere-terpstra test. A post-hoc Tukey-test was then applied for multiple pairwise-comparisons between GOLD stages.

The relationships between pulmonary function tests, clinical data and each PRM class, anomaly scores and cluster groups were evaluated through Pearson's correlation coefficient. Confidence intervals (CI) were calculated using bootstrap resampling method on 10,000 samples. Pulmonary function tests and clinical data included parameters such as FEV1% predicted, FEV1/FVC, Functional Residual Capacity (FRC), Total Lung Capacity (TLC), FRC/TLC, BODE index (body mass index, air-flow obstruction, dyspnea, exercise capacity), St. George's Respiratory Questionnaire (SGRQ), the 6-min-walking-test (6MWT) and smoking duration. Correlations were interpreted as follows: 0.00–0.10 (negligible), 0.10–0.39 (weak), 0.40–0.69 (moderate), 0.70–0.89 (strong), and 0.90–1.00 (very strong) (22). Differences between correlation coefficients were assessed via the R package "cocor" (23), utilizing the Zou et al. (24) method. This involved calculating the difference between correlation coefficients for each pair of groups and determining a 95% confidence interval (CI) for that difference. If the CI included zero, the null hypothesis that the two correlations are equal was retained; if the CI did not include zero, the null hypothesis was rejected.

Linear mixed effects models (LMM) were utilized to predict clinical parameters based on the PRM volumes, with adjustments made for relevant covariates including age, gender, body mass index (BMI), smoking status (0: never-smoker control, 1: former smoker, and 2: current smoker), smoking duration, and a random term for the study site. To assess the contribution of the anomaly score beyond morphological lung changes in predicting clinical variables (FEV1%, FEV/FVC, FRC, TLC, FRC/TLC, BODE, SGRQ, and 6MWT), it was then introduced as an additional predictor in the PRM-adjusted LMM models. The overall conditional coefficient of determination (R^2), adjusted for the number of regressors, was reported. Models were compared through the likelihood ratio test of nested models.

The agreements between PRM relative volumes and volumes obtained from the Cluster groups were assessed using the Bland–Altman method (25). A p-value of <0.05 was considered as statistically significant and adjustments were made for multiple comparisons, using the Holm-Bonferroni method, when applicable.

All statistical analyses were performed using R software version 4.3.4 (R Foundation for Statistical Computing, Vienna, Austria).

3 Results

Patient characteristics are summarized in Table 1.

3.1 Visualization of the latent space: t-distributed stochastic neighbor embedding

The comprehensive representation of the lung regions in the latent space, achieved through self-supervised contrastive learning, is visually depicted using t-SNE (Figure 2). Each point represents a feature vector (1×512) per lung region (3D patch). A clear separation between regions with and without emphysema and gas trapping is consistently illustrated. Also, regions exhibiting less than 1% Emphysema and less than 1% Air Trapping are distinctly discernable from diseased regions. Notably, factors such as gender information do not contribute to the separation of these latent features (Supplementary Figure S4).

3.2 PCA and cluster analysis

PCA applied to the region-wise latent vectors determined that 85 factors should be retained, meeting the criteria of obtained eigenvalues surpassing those from random data, as per Horn's parallel analysis and the Kaiser Criterion (Supplementary Figure S1). Supplementary Figure S2 provides a comparison between the number of clusters and the method applied to the 85 retained PCs. Cluster analysis identified four stable clusters (Clusters 1, 2, 3, and 4) through K-means with a mini-batch size of 1,000. More details about the clusters are available in the Supplementary materials.

3.3 Qualitative results

Figure 3 showcases representative CT images individuals from all GOLD grades. The first and second column represent a coronal slice of the inspiratory and expiratory CT scans, while the following columns are the visualization of PRM maps, maps of the anomaly score and cluster classes, overlayed on the inspiratory CT. Figure 4 represents the spatial overlap between region-wise anomaly scores and PRM[Emph] and PRM[fSAD] for two former-smoker individuals: GOLD 0 (A) and GOLD 3 (B), with more than 30 years of smoking history.

3.4 PRM volumes, anomaly score, cluster groups by GOLD score

Figure 5 depicts a comparison between PRM volumes, anomaly score and relative size of the 4 clusters, relatively to the GOLD stage. Significant differences were found between PRM volumes, anomaly scores and all cluster classes according to the GOLD stage (Table 2), except for class 4. With an increase in the GOLD stage, PRM[Emph], PRM[fSAD], Anomaly Score and relative volume of cluster 3 increased, while cluster 2 decreased. Post-hoc Tukey analysis (Supplementary Table S1) revealed significant intergroup differences in the anomaly score between all GOLD stages ($p < 0.01$ corrected),

TABLE 1 Subject demographics (sex, age, body mass index [BMI]), functional parameters (post-bronchodilator forced expiratory volume in one second [FEV1%_pred]; FEV1/Forced Vital Capacity [FVC]; Functional Residual Capacity [FRC]; Total Lung Capacity [TLC]; FRC/TLC; Body-Mass Index, Airflow Obstruction, Dyspnea, and Exercise Capacity Index [BODE]; St. George's Respiratory Questionnaire [SGRQ]; 6-min walking test [6MWT]), smoking duration and low-attenuation (LAA) percentages [LAA-950% and LAA-856%], for the evaluation population, from the COPDGene dataset.

	Never-smoker controls (N = 29)	GOLD 0 (N = 538)	GOLD 1 (N = 128)	GOLD 2 (N = 315)	GOLD 3 (N = 195)	GOLD 4 (N = 105)	p-value
Sex [f/m]	11/18	280/258	74/54	161/154	116/79	50/55	0.011
Age (y)	62.2 ± 9.1	59.5 ± 8.3	63.2 ± 8.7*	64.8 ± 8.6	66.2 ± 7.9	65.4 ± 7.0	<0.001
BMI	27.8 ± 4.3	28.6 ± 5.6	27.1 ± 4.2	28.5 ± 6.0	28.4 ± 6.5	25.4 ± 4.8*	<0.001
FEV1%_pred	100.6 ± 13.2	96.4 ± 10.7	90.4 ± 9.0*	65.8 ± 8.8*	39.4 ± 5.5*	23.2 ± 4.2*	<0.001
FEV1/FVC	0.8 ± 0.1	0.8 ± 0.0	0.6 ± 0.0*	0.6 ± 0.1*	0.4 ± 0.1*	0.3 ± 0.1*	<0.001
FRC	2.5 ± 0.6	2.9 ± 0.7	3.4 ± 0.9*	3.5 ± 0.9	4.5 ± 1.2*	5.1 ± 1.3*	<0.001
TLC	5.2 ± 1.4	5.7 ± 1.2	6.4 ± 1.5*	5.9 ± 1.3	6.3 ± 1.6	6.6 ± 1.4	<0.001
FRC/TLC	0.5 ± 0.1	0.5 ± 0.1	0.5 ± 0.1	0.6 ± 0.1*	0.7 ± 0.1*	0.8 ± 0.1*	<0.001
BODE	0.1 ± 0.3	0.3 ± 0.7	0.5 ± 0.9	1.4 ± 1.3*	3.9 ± 1.4*	5.5 ± 1.2*	<0.001
SGRQ	4.3 ± 8.1	13.0 ± 15.6	17.5 ± 17.9	31.0 ± 20.8*	45.5 ± 18.0*	55.8 ± 16.6*	<0.001
6MWT (meters)	539.4 ± 88.7	476.3 ± 100.4	468.9 ± 105.4	415.6 ± 109.0*	344.5 ± 118.8*	267.4 ± 102.5*	<0.001
Smoking duration (y)	-	32.8 ± 10.9	38.3 ± 10.2*	40.3 ± 9.5	41.3 ± 8.7	40.3 ± 8.5	<0.001
LAA-950%	1.9 ± 2.0	2.6 ± 2.9	5.1 ± 5.1	8.3 ± 8.4	15.5 ± 12.2	27.9 ± 13.4	<0.001
LAA-856%	9.4 ± 6.8	12.0 ± 9.8	19.9 ± 11.2	28.7 ± 15.9*	47.5 ± 17.3*	63.7 ± 11.8*	<0.001

Attenuation percentages were measured by VIDA Diagnostics. All data are given as mean ± standard deviation. For some cases, data was not available: 17 for 6MWT, 17 for BODE, 76 for FRC, 3 for TLC, 61 for FRC/TLC, 3 for LAA-950%, 76 for LAA-856%. p-value was obtained by the Jonckheere-Terpstra test or a logistic regression, when applicable. * denotes $p < 0.05$ vs. previous GOLD stage for the post-hoc Tukey-test comparison.

except between controls and GOLD 0. No differences were observed between controls and GOLD 0, controls and GOLD 1, and GOLD 0 and GOLD 1 for PRM^{Emph}; and controls and GOLD 0 and GOLD 4 and GOLD 3 for PRM^{fSAD}. Cluster 1 showed significant differences between high grades COPD (GOLD 3 and 4), and Cluster 2 between GOLD 2 and 3. Notably, statistically significant differences between never-smoker controls and GOLD 0 subjects were only found by Cluster 3.

3.5 Relationships between PRM volumes, anomaly score, cluster groups

Table 3 shows the Pearson correlation coefficient and p-values corrected with Holm-Bonferroni for PRM volumes, anomaly score and cluster groups. Each PRM volume showed statistically significant correlations with PFTs and clinical data. Both PRM^{fSAD}, PRM^{Emph}, Anomaly score and Cluster 3 are negatively correlated with spirometry volumes (FEV1 and FEV1/FVC) and with the walking distance, and positively correlated with FRC, TLC, FRC/TLC, BODE and smoking duration. Pearson correlation coefficients for the anomaly score are moderate, with the exception of FRC and smoking duration, and very well comparable with the ones from PRM volumes, while less strong for the Cluster 3. The anomaly score showed significantly higher correlations than PRM^{fSAD} for SGRQ and the distance walked in the 6-min walking test; and stronger than PRM^{Emph} for FRC/TLC and smoking duration. No significant differences were found between correlations of the

anomaly score and PRM^{fSAD} for FEV1, FEV1/FVC, BODE and smoking duration; and between the anomaly score and PRM^{Emph} for FEV1, FRC, SGRQ and the distance walked in the 6-min walking test. The anomaly scores also showed significantly higher correlations than all Clusters for all clinical variables, except for TLC. Clusters 1 and 2 follow the same trends as the PRM Healthy volumes: positively correlated with FEV1, FEV1/FVC and walking distances and negatively correlated with FRC, TLC, FRC/TLC, BODE, SGRQ, and smoking duration.

Figure 6 illustrates linear relationships between patient-wise anomaly scores and PRM-derived emphysema and fSAD volumes. The anomaly score shows to be significantly correlated with PRM^{Emph} ($r = 0.66$, $p < 0.01$) and PRM^{fSAD} ($r = 0.61$ $p < 0.01$). As depicted in Table 4, adding the anomaly score to the PRM baseline models statistically improves the fit of all LMM to predict clinical measures, including FEV1, FEV/FVC, FRC, FRC/TLC, BODE, SGRQ, and 6MWT (only to PRM^{fSAD}). No differences were found for predicting the TLC and the 6MWT (for PRM^{Emph}).

Bland–Altman plots for the volume (log transformed) agreement of PRM-derived volumes and clusters which showed significant correlation to the PFTs and clinical data (Clusters 1, 2, 3) are presented in Figure 7 showing the differences (D) against averages (A). Since the Bland–Altman plots displayed non-constant bias even after log transformation of the measurements, a regression-based approach (25) was used to compute the bias and limits of agreement (LoA). The mean differences are given by

$$\hat{D}_{PRM,Cluster1} = -5.04 + 1.80\,A \quad, \quad \hat{D}_{PRM,Cluster2} = -6.84 + 1.68\,A \quad,$$
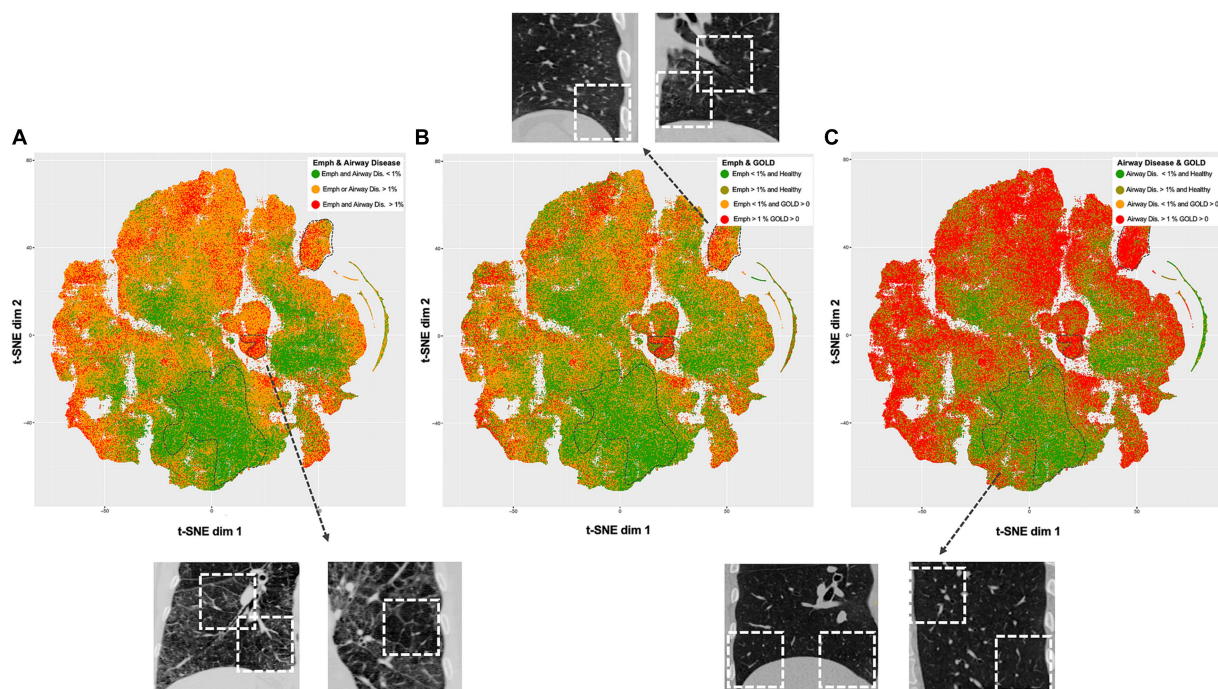
FIGURE 2
t-Distributed Stochastic Neighbor Embedding (t-SNE) visualizations of the self-supervised contrastive latent space vectors. Each dot represents a region (3D ROI) or more specifically, an embedding of its latent representation into a two-dimensional space, and its color represents a clinical or radiological characteristic (GOLD, emphysema and air trapping measures at the patch-level). The dotted regions in the t-SNE maps emphasize distinct groups (bottom – healthy, center right and upper right – diseased), serving as illustrative examples. For each dotted region, several examples are provided to illustrate representative inspiratory patches extracted from those dotted regions. Readers are guided to recognize consistent groups positioning across visualizations, enhancing overall interpretation of the t-SNE plots. **(A)** Visualization is colored by regions with and without emphysema and airway disease. **(B)** Visualization is colored by regions with less than 1% emphysema from healthy individuals (green), regions with more than 1% emphysema from healthy individuals (dark green), regions with less than 1% emphysema from COPD individuals (yellow) and regions with more than 1% emphysema from COPD individuals (red). **(C)** Same as in B but related to airway disease. Healthy individuals were defined as controls and GOLD 0, while COPD individuals as GOLD 1, 2, 3, or 4.

$\hat{D}_{PRM,Cluster\,3} = -2.17 + 0.46\,A$. The 95% LoA are then

$$\hat{D}_{PRM,Cluster\,1} = -5.04 + 1.80\,A \pm 0.80 \quad, \quad \hat{D}_{PRM,Cluster\,2} = -6.84 + 1.68\,A \pm 1.05 \quad,$$

$\hat{D}_{PRM,Cluster\,3} = -2.17 + 0.46\,A \pm 1.31$. Thus, the degree of agreement is not constant across the range of measurements.

4 Discussion

In this study, we aimed to explore the potential of a deep-learning self-supervised anomaly detection method for phenotyping Chronic Obstructive Pulmonary Disease (COPD) using computed tomography (CT) scans of 1,310 never-smoker controls and GOLD 0–4 from the COPDGene cohort. COPD remains a significant global health burden, necessitating early detection and accurate phenotyping for effective intervention. While traditional methods like parametric response mapping (PRM) have provided insights into COPD pathophysiology (27, 28), they may not fully capture the complexity and heterogeneity of the disease. This present study marks the first of its kind, comparing PRM-derived functional small-airway disease and emphysema measurements against regional anomaly scores derived from a

recently proposed self-supervised deep-learning anomaly detection approach.

Before delving into the comparison itself, it was important to understand what spatial distribution is encoded by the self-supervised latent features. As unveiled through t-distributed Stochastic Neighbor Embedding (t-SNE) analysis, it showcases a notable attraction among similar disease-related areas, concurrently distancing healthy regions from those corresponding to healthy patients. Notably, this visualization is rooted in the self-supervised contrastive learning of lung regions and, as such, does not rely on labels. It is noteworthy that this approach effectively captures features stemming from emphysema and gas trapping, while factors such as gender information do not contribute to this attraction. Furthermore, an intriguing observation surfaces in COPD patients, where even regions with low levels of emphysema exhibit spatial resemblances with areas demonstrating higher emphysema levels. This pattern potentially hints at underlying features beyond emphysema that play a role in clustering these regions together, possibly indicating the presence of early-stage COPD markers that could progress into advanced emphysematous changes.

Investigating the spatial relationships between small airway disease and emphysema manifestations captured by PRM alongside regions identified as anomalies by the self-supervised approach has revealed interesting insights. Traditionally, the progression of COPD involves the narrowing or destruction of small airways preceding the
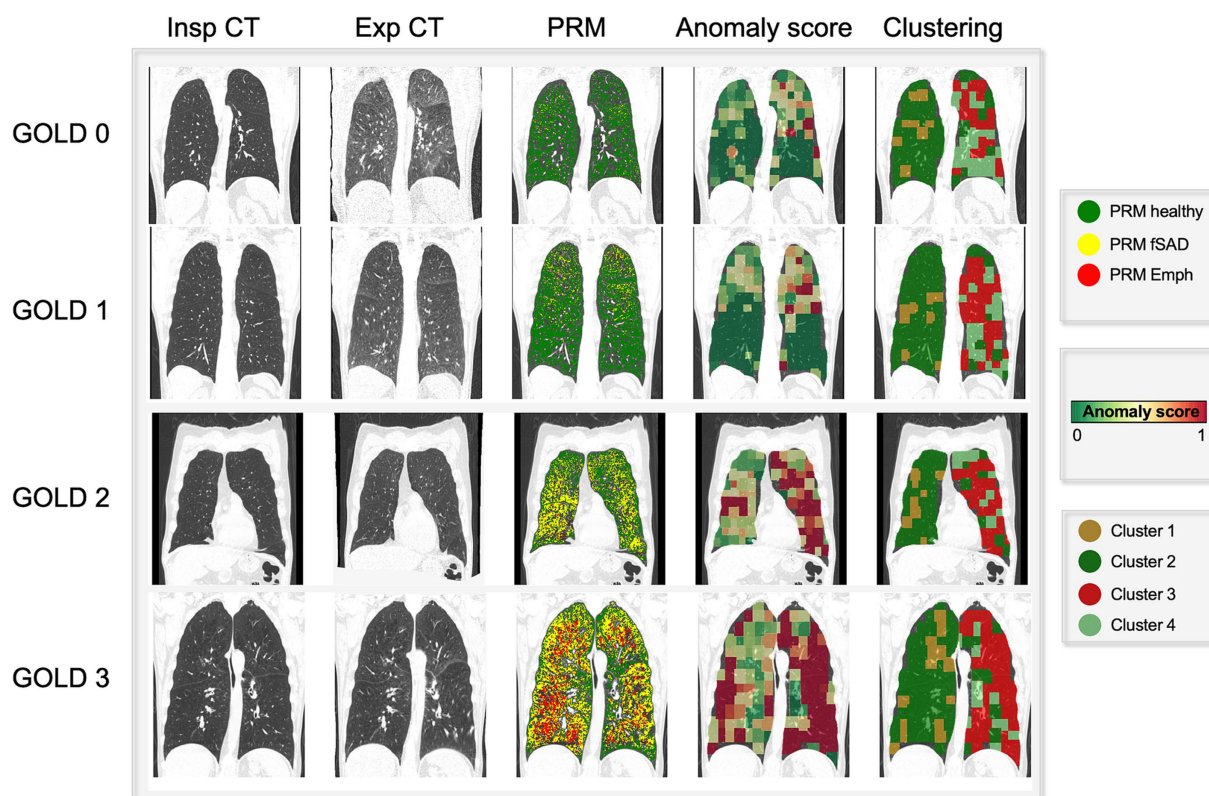
FIGURE 3
Visual comparison of PRM maps, anomaly score and cluster volumes on exemplary subjects across several GOLD stages (0−4). As the severity increases (GOLD score), so do the areas detected as emphysema and functional small airway disease by PRM. On the same fashion, the region anomaly scores also shift from green to red for more severe cases. Interestingly, these areas overlap PRM^{Emph} and PRM^{fSAD}. Particularly for low GOLD levels, it attributes higher scores to PRM^{fSAD} areas, possibly indicative of differential progressive features. Lastly, the cluster maps reveal four aggregations, indicative of different disease components.
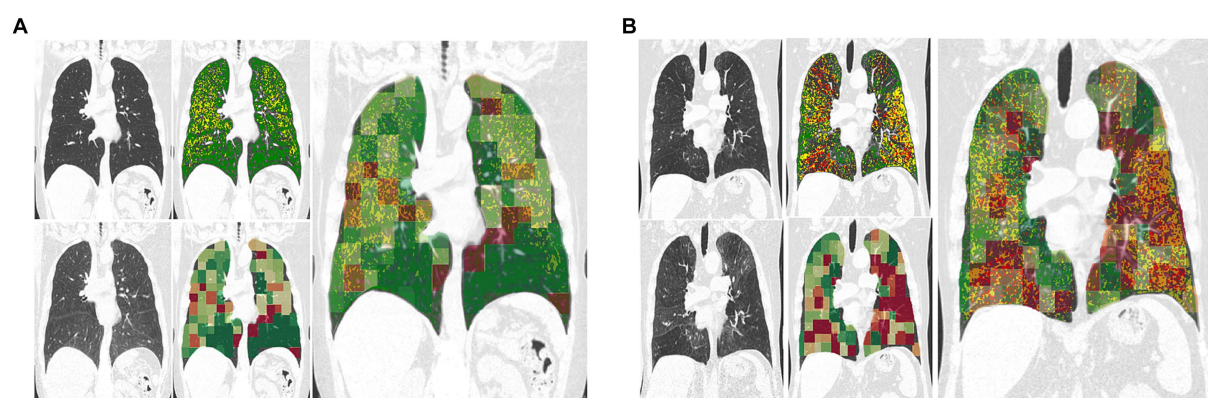
FIGURE 4
Exemplary coronal sections of overlapping of PRM and anomaly maps for two former-smoker individuals. Example **(A)** is a 64 years-old former-smoker female who does not fulfill the criteria for COPD (GOLD 0; FEV1% post-bronchodilator 126% of predicted value; FEV1/FVC 78%; smoking history of 30 years). Example **(B)** is a 72 years-old former-smoker male with severe COPD (GOLD 3; FEV1% post-bronchodilator 31% of predicted value; FEV1/FVC 34%; smoking history of 37 years). Top left corresponds to the inspiratory scans, top right to the PRM maps, bottom left to the expiratory registered scan, bottom right to the anomaly score map. The bigger image corresponds to the overlapping of the anomaly maps to the PRM maps, in which regions deviating from the norm are overlapping with PRM^{fSAD}. The closer to red, the higher the anomaly detected. PRM healthy tissue (green), PRM functional small-airway disease (yellow), and PRM emphysema (red).

development of emphysema (29), a transition not easily detectable using conventional spirometry or visible on CT scans (27, 30). Our exemplary patients' regions, colored solely by PRM^{fSAD}, often exhibit areas with higher anomaly scores in the anomaly maps, as illustrated in Figures 3, 4A. This observation raises the intriguing possibility that the anomaly detection approach serves as an additional means to
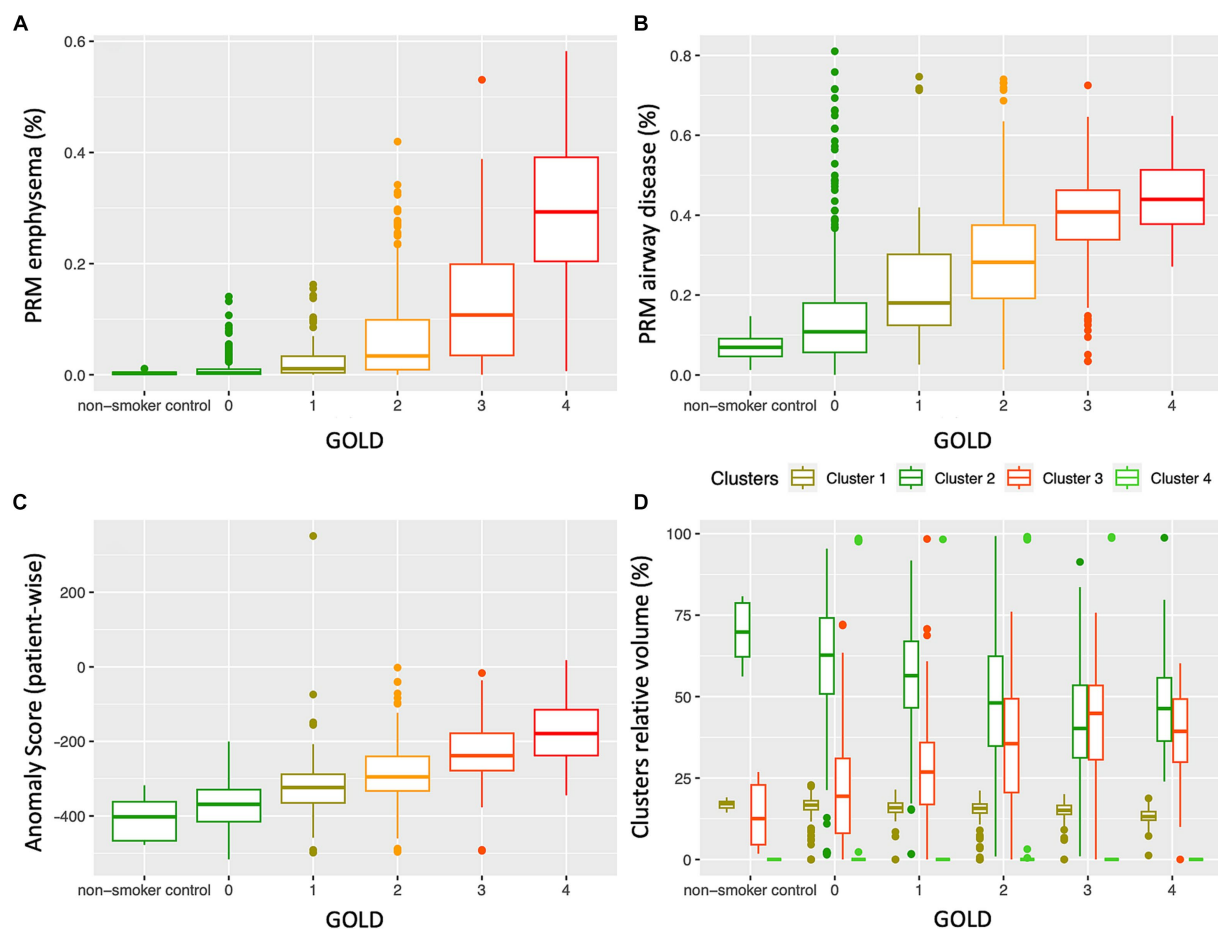
FIGURE 5

Distribution of PRM emphysema **(A)**, PRM airway disease (fSAD) **(B)**, patient-wise anomaly score **(C)** and cluster classes **(D)**, according to the GOLD stage. While PRM^Emph, PRM^fSAD and the anomaly score clearly increase with the disease severity, the relative volumes of the cluster classes showed distinctly different distributions according to GOLD severity; cluster 3 increases, whereas cluster 2 decreases with an increase in GOLD stage. GOLD, Global Initiative for Obstructive Lung Disease.

TABLE 2 Mean ± standard deviation of parametric response mapping (PRM) volumes, anomaly score and cluster groups, according to the GOLD stage.

	Controls (N = 29)	GOLD 0 (N = 538)	GOLD 1 (N = 128)	GOLD 2 (N = 315)	GOLD 3 (N = 195)	GOLD 4 (N = 105)	p-value
PRM							
Healthy (%)	92.6 ± 4.1	84.4 ± 14.8	75.4 ± 16.2*	64.2 ± 19.3*	47.3 ± 18.4*	25.4 ± 10.3*	<0.001
Functional small-airway (%)	7.1 ± 4.0	14.7 ± 13.8	22.0 ± 13.8*	28.9 ± 14.0*	39.4 ± 11.5*	45.1 ± 8.5	<0.001
Emphysema (%)	0.3 ± 0.3	0.9 ± 1.7	2.6 ± 3.6	6.8 ± 8.0*	13.3 ± 11.3*	29.5 ± 13.6*	<0.001
Anomaly detection (cOOpD)							
Anomaly score	−404.1 ± 54.3	−372.0 ± 63.8	−318.9 ± 100.8*	−287.8 ± 80.8*	−228.0 ± 77.4*	−170.9 ± 82.2*	<0.001
Clusters							
1	16.8 ± 1.4	16.2 ± 3.4	15.5 ± 3.2	15.3 ± 3.3	14.9 ± 2.7	13.2 ± 2.7*	<0.001
2	69.7 ± 9.0	60.7 ± 17.8	55.7 ± 17.7	48.9 ± 19.4	42.2 ± 16.4*	48.5 ± 15.1	<0.001
3	13.5 ± 9.3	21.1 ± 15.7*	27.7 ± 17.4	34.2 ± 18.5	41.7 ± 16.3*	38.2 ± 13.8	<0.001
4	0.0 ± 0.0	1.9 ± 13.6	1.1 ± 10.2	1.6 ± 12.5	1.2 ± 10.8	0.0 ± 0.0	n.s.

GOLD, Global Initiative for Obstructive Lung Disease. p-value was obtained by the Jonckheere-Terpstra test. * denotes $p < 0.05$ vs. previous GOLD stage for the post-hoc Tukey-test comparison.

highlight transitional stages in the disease process, paramount for understanding early pathological changes (31). It may also suggest that small airway disease undergoes dynamic changes before evolving into more advanced stages of emphysema, particularly relevant for individuals with normal lung function (GOLD 0, Figure 4A). Therefore, more importantly than the spatial similarity, the divergence

TABLE 3 Correlation of PRM volumes, anomaly score, and cluster groups with PFTs and clinical data.

Method	Pulmonary function tests and clinical data								
	FEV1	FEV1/FVC	TLC	FRC	FRC/TLC	BODE	SGRQ	6MWT	Duration smoking
PRM									
Healthy (%)	0.71**** (0.68, 0.74)	0.80**** (0.77, 0.82)	−0.30**** (−0.35, −0.24)	−0.72**** (−0.75, −0.69)	−0.73**** (−0.76, −0.70)	−0.64**** (−0.68, −0.60)	−0.46**** (−0.51, −0.41)	0.42**** (0.37, 0.47)	−0.27**** (−0.30, −0.20)
PRM^{fSAD} %	−0.61**** (−0.66, −0.57)	−0.68**** (−0.72, −0.64)	0.24**** (0.18, 0.30)	0.70**** (0.67, 0.73)	0.78**** (0.76, 0.81)	0.53**** (0.48, 0.57)	0.39**** (0.34, 0.44)	−0.34**** (−0.39, −0.30)	0.28**** (0.20, 0.31)
PRM^{Emph} %	−0.66**** (−0.69, −0.63)	−0.75**** (−0.77, −0.72)	0.30**** (0.24, 0.36)	0.56**** (0.51, 0.61)	0.44**** (0.39, 0.50)	0.64**** (0.60, 0.68)	0.43**** (0.38, 0.48)	−0.42**** (−0.47, −0.37)	8**** (0.11, 0.21)
Anomaly detection (cOOpD)									
Anomaly score	−0.65**** (−0.69, −0.60)	−0.68**** (−0.72, −0.64)	0.16**** (0.09, 0.22)	0.55**** (0.51, 0.60)	0.56**** (0.51, 0.62)	0.57**** (0.52, 0.62)	0.46**** (0.41, 0.51)	−0.41**** (−0.47, −0.35)	0.27**** (0.19, 0.30)
Clusters									
1	0.22**** (0.17, 0.29)	0.27**** (0.21, 0.34)	−0.22**** (−0.28, −0.15)	−0.34**** (−0.41, −0.27)	−0.35**** (−0.43, −0.27)	−0.23**** (−0.29, −0.16)	−0.15**** (−0.21, −0.09)	0.16**** (0.10, 0.22)	−0.10*** (−0.15, −0.03)
2	0.33**** (0.28, 0.38)	0.34**** (0.28, 0.39)	−0.06* (−0.12, −0.01)	−0.22**** (−0.28, −0.16)	−0.17**** (−0.25, −0.09)	−0.25**** (−0.30, −0.19)	−0.21**** (−0.26, −0.15)	0.20**** (0.14, 0.26)	−0.16**** (−0.19, −0.07)
3	−0.41**** (−0.46, −0.36)	−0.42**** (−0.47, −0.37)	0.11*** (0.05, 0.17)	0.32**** (0.27, 0.38)	0.29**** (0.22, 0.36)	0.32**** (0.27, 0.38)	0.27**** (0.22, 0.32)	−0.24**** (−0.30, −0.19)	0.18**** (0.10, 0.21)
4	0.04 (−0.01, 0.10)	0.04 (−0.02, 0.09)	−0.01 (−0.06, 0.04)	−0.06 (−0.10, −0.02)	−0.08* (−0.11, −0.05)	−0.04 (−0.10, 0.02)	−0.04 (−0.10, 0.01)	0.00 (−0.05, 0.06)	0.00 (−0.07, 0.04)

Confidence intervals are denoted in brackets, per each Pearson correlation coefficient. Correlations are colored according to Cohen's effect size (26) and only the significant ones are colored. None or very weak effect size is defined by $|r| < 0.3$, weak $0.3 < |r| < 0.5$, moderate $0.5 < |r| < 0.7$ and large by $|r| > 0.7$. PRM, parametric response mapping; fSAD, functional Small Airway Disease; Emph, Emphysema; FEV1, Forced Expiratory Volume in 1 s; FEV1/FVC, FEV/Forced Vital Capacity; FRC, Functional Residual Capacity; TLC, Total Lung Capacity; BODE, Body mass index, air-flow Obstruction, Dyspnea, Exercise capacity; SGRQ, St George's Respiratory Questionnaire; 6MWT, 6 min walking test. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, *$p < 0.05$.

FIGURE 6
Scatterplots depicting the relationship between the anomaly score and PRM emphysema (A) and PRM^{fSAD} (B). Both show linear regression between patient-wise anomaly score and PRM volumes in all subjects (PRM^{Emph}: $r = 0.66$, $p < 0.01$; PRM^{fSAD}: $r = 0.61$ $p < 0.01$).

TABLE 4 Linear mixed effects models were employed to predict several clinical variables including FEV1%, FEV/FVC, FRC, TLC, FRC/TLC, BODE, SGRQ, and 6MWT, with adjustments made for age, gender, BMI, smoking status, smoking duration, and study site (baseline models).

Dependent variable	Predictor	Adjusted conditional R^2	p-value adjusted
FEV1%	Age, gender, BMI, smoking status, smoking duration (center)	0.22	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.46	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.56**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.53	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	**0.58**	$p < 0.001$
FEV1/FVC	Age, gender, BMI, smoking status, smoking duration (center)	0.26	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.54	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.63**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.63	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	**0.66**	$p < 0.001$
TLC	Age, gender, BMI, smoking status, smoking duration (center)	0.50	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.54	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	0.55	n.s
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.56	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	0.56	n.s
FRC	Age, gender, BMI, smoking status, smoking duration (center)	0.32	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.67	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.69**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.54	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	**0.58**	$p < 0.001$
FRC/TLC	Age, gender, BMI, smoking status, smoking duration (center)	0.23	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.68	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.69**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.36	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	**0.41**	$p < 0.001$
BODE	Age, gender, BMI, smoking status, smoking duration (center)	0.21	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.37	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.45**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.49	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	**0.51**	$p < 0.001$
SGRQ	Age, gender, BMI, smoking status, smoking duration (center)	0.24	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.35	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.41**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.39	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	**0.42**	$p < 0.001$
6MWT	Age, gender, BMI, smoking status, smoking duration (center)	0.41	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD** (center)	0.42	
	Age, gender, BMI, smoking status, smoking duration, **PRM^fSAD, anomaly score** (center)	**0.45**	$p < 0.001$
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph** (center)	0.48	
	Age, gender, BMI, smoking status, smoking duration, **PRM^Emph, anomaly score** (center)	0.48	n.s.

PRM volumes (PRM^Emph and PRM^fSAD) were then incorporated as predictors. Subsequently, these models were compared to those further adjusted for the anomaly score. Conditional R^2 values are adjusted for the number of regressors added. Bold values indicate a higher R^2 for each dependent variable. All models were statistically significant compared to the baseline model. p-values are reported for the comparison between PRM baseline models and anomaly-adjusted PRM models, and are corrected for multiple comparisons.

FIGURE 7
Bland−Altman plots regression-based limits of agreement. Dotted line represents the mean difference between the methods **(A)** PRM and Cluster 1: $\hat{D}_{PRM,Cluster1}=-5.04 + 1.80\,A$; **(B)** PRM and Cluster 2 $\hat{D}_{PRM,Cluster2}=-6.84 + 1.68\,A$; **(C)** PRM and Cluster 3 $\hat{D}_{PRM,Cluster3}=-2.17 + 0.46\,A$. Dashed line represents the 95% Limits of Agreement: **(A)** PRM and Cluster 1: $\hat{D}_{PRM,Cluster1}=-5.04 + 1.80\,A \pm 0.80$; **(B)** PRM and Cluster 2 $\hat{D}_{PRM,Cluster2}=-6.84 + 1.68\,A \pm 1.05$; **(C)** PRM and Cluster 3 $\hat{D}_{PRM,Cluster3}=-2.17 + 0.46\,A \pm 1.31$. Red points refer to the agreement to PRM emphysema and yellow to the agreement to PRM small airway disease.

observed between anomaly maps and PRM classes indicates the potential of the anomaly detection method to detect early-stage disease manifestations not yet captured by traditional phenotyping methods, particularly in regions of individuals with a long smoking history (Figure 4). Further investigations, including the assessment of normal CT voxels alongside CT airway measurements and PRM analyses, are warranted to elucidate the underlying pathophysiological mechanisms responsible for these observations and to validate these findings in prospective longitudinal studies. Furthermore, our study revealed a significant positive correlation between the patient-wise anomaly scores and PRM-derived fSAD ($r=0.61$, $p<0.01$) and Emphysema volumes ($r=0.66$, $p<0.01$). Overall, these findings underscore the complementary nature of the two approaches and the potential benefit of integrating them to achieve a more comprehensive assessment of COPD heterogeneity.

Moreover, our study reveals a parallel finding to Hwang HJ et al. (32), where a novel emphysema air-trapping composite (EAtC) has been proposed. In particular, their functional air trapping (fAT) component captures air trapping in both emphysematous and non-emphysematous areas. Just as in their work, our investigation suggests that employing the anomaly detection comprehensive approach yields better or comparable correlations with clinical variables (SGRQ, 6MWT; =FEV1, =FEV1/FVC, =BODE and=smoking duration) compared to conventional PRM-based small-airway disease assessments, which focus solely on non-emphysematous air trapping. Similarly, the anomaly score exhibited stronger correlations than PRM^{Emph} for parameters like FRC/TLC and smoking duration, while no differences were found for FEV1, FRC, SGRQ and the distance walked in the 6-min walking test. This suggests the ability of the anomaly detection approach in

characterizing small-airway disease and emphysema more comprehensively, potentially encompassing aspects that conventional methods might overlook. Subsequent analysis confirmed this hypothesis, as adding the anomaly score to LMM adjusted for both PRM volumes, significantly improved the prediction of clinical variables (FEV1, FEV/FVC, FRC, FRC/TLC, BODE, SGRQ). This provides further evidence that the anomaly score captures nuanced features beyond the structural characteristics from the PRM analysis and underscores their complementary nature.

While the debate surrounding air trapping in emphysematous regions remains, evidence from both Hwang HJ et al. (32) and our study supports the notion that small-airway disease is not confined solely to areas with preserved alveoli but can coexist within emphysematous zones as well. This duality emphasizes the interplay between different disease manifestations and highlights the potential for anomalies, as captured by the anomaly detection approach, to encompass both emphysematous and non-emphysematous regions. Additionally, it may also explain why, in contrast to PRM^{fSAD} and PRM^{Emph}, the patient-wise anomaly score significantly differs across all GOLD stages ($p<0.01$), except between never-smoker controls and GOLD 0.

Our cluster analysis and visualization of PRM-derived regions have illuminated distinct patterns within the dataset. By employing principal component analysis (PCA) for dimensionality reduction, we successfully identified four stable clusters that exhibited variations corresponding to different stages of the disease. Notably, Cluster 1 appears to embody regions that persist consistently across all GOLD stages, indicative of common characteristics shared across disease phenotypes. Cluster 2, on the other hand, emerges as a distinct subset that represents healthy regions and unaffected by the disease, potentially explaining its volume decline as disease severity increases.

The most intriguing finding lies in Cluster 3, where regions manifest features closely aligned with PRM-derived fSAD (Figure 7), which suggests an association with early-stage COPD characteristics. It is conceivable that Cluster 1 represents areas that have never experienced the impact of COPD, while Cluster 3 may signify regions embarking on the initial stages of disease progression. This interpretation finds support in the progressively increasing representation of Cluster 3 across GOLD classes, hinting at its role in the evolution of the disease from milder to more advanced stages. Although the correlations observed between PFTs and PRM-derived Emphysema and fSAD classes are superior to Cluster 3, Cluster 3 is the only able to differentiate never-smoker controls and GOLD 0 subjects. Amidst these notable clusters, Cluster 4 adds a layer of complexity. While present in a smaller subset of patients ($n = 24$), notably it covers over 50% of the lung in 22 individuals, showing distinct clinical features like high gas trapping ($31.3 \pm 23.4\%$) and extensive smoking history (33 ± 12 years). Still, the decision to retain it was supported by robust clustering metrics detailed in the Supplementary materials.

The non-constant bias observed in the Bland–Altman analysis suggests that the agreement between PRM-derived volumes and cluster volumes is not uniform across the entire range of measurements. This non-uniformity may stem from inherent differences in how PRM and the clustering method characterize and quantify lung regions. The regression-based equations provide a nuanced understanding of this non-constant bias, indicating that the discrepancy between methods is influenced by the magnitude of the measurements. These findings may imply that the clusters, despite their ability to capture distinct patterns related to COPD stages, may not uniformly agree with PRM-derived volumes across all levels of disease severity. The varying agreement observed in different parts of the measurement range could be attributed to the complex and heterogeneous nature of COPD, where different phenotypes and disease manifestations may impact the agreement between methods differently. Moreover, it's important to acknowledge an unexpected discrepancy observed in the distribution of clusters between the left and right lungs. Visual inspection of the original inspiratory and expiratory CT scans did not consistently support the distinct left/right lung asymmetry detected by the clustering algorithm. Notably, we observed an apparent concentration of clusters 3 and 4 in the left lung and clusters 1 and 2 in the right lung. This left/right lung asymmetry prompts caution in the interpretation of cluster-specific findings. While our primary objectives center around the comparative analysis between PRM and anomaly detection, this unexpected observation underscores the need for transparency regarding potential sources of bias in the clustering component. This information adds a layer of transparency to our study, acknowledging that the agreement between PRM and clustering might be influenced by factors that vary across different regions of the lung or disease states. Importantly, it's worth noting that the anomaly score, which consistently demonstrated significant correlations with PRM-derived measurements and clinical data, exhibited a more uniform agreement, providing a robust and complementary perspective on COPD phenotyping. Our findings also indicate that the anomaly detection method provides a more robust and complementary perspective on COPD phenotyping compared to Clustering. This is demonstrated by statistically significant higher correlations between anomaly scores and clinical variables, along with larger effect sizes. Anomaly detection's ability to identify deviations from the distribution of "normal" samples makes it better suited for capturing subtle variations indicative of disease pathology, contrasting

with Clustering's reliance on proximity-based grouping, while based on the exact same latent features.

Furthermore, while our findings indicate a strong potential for anomaly detection in COPD phenotyping, further research is needed to establish its generalizability across diverse populations. Additionally, our current exploration does not delve into distinguishing COPD from other comorbities, such as lung cancer, warranting careful consideration and validation in future investigations.

In conclusion, we introduce the anomaly detection approach as a novel perspective in COPD phenotyping, extending the lens beyond threshold-based analysis and methodologies focused on distinct imaging features. By identifying anomalies spanning a spectrum of disease features, including those beyond functional small-airway disease and emphysema, we unveil a more comprehensive viewpoint, offering an additional layer of information that goes beyond traditional PRM phenotyping. The resulting insights offer a window into the intricate distribution of the disease. This transparency in AI diagnostics empowers clinicians to personalize interventions while supplementing established methods. By integrating this innovative approach into clinical practice, we lay the groundwork for refined diagnostics and tailored interventions, ultimately leading to enhanced patient outcomes.

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

Ethics approval conducted in COPDGene. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. TN: Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. CL: Investigation, Methodology, Writing – original draft, Writing – review & editing, Validation. TW: Investigation, Methodology, Writing – original draft, Writing – review & editing. VW: Data curation, Formal analysis, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MN: Supervision, Writing – original draft, Writing – review & editing. PJ: Supervision, Writing – original draft, Writing – review & editing. OS: Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. CH: Investigation, Supervision, Writing – original draft, Writing – review & editing. OW: Investigation, Supervision, Writing – original draft, Writing – review & editing. JB: Investigation, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. H-UK: Writing

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1360706/full#supplementary-material

References

1. Chen S, Kuhn M, Prettner K, Yu F, Yang T, Bärnighausen T, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020–50: a health-augmented macroeconomic modelling study. *Lancet Glob Health*. (2023) 11:e1183–93. doi: 10.1016/S2214-109X(23)00217-6

2. Halpin DMG, Criner GJ, Papi A, Singh D, Anzueto A, Martinez FJ, et al. Global initiative for the diagnosis, management, and prevention of chronic obstructive lung disease. The 2020 GOLD science committee report on COVID-19 and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2021) 203:24–36. doi: 10.1164/rccm.202009-3533SO

3. Adeloye D, Song P, Zhu Y, Campbell H, Sheikh A, Rudan I, et al. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. *Lancet Respir Med*. (2022) 10:447–58. doi: 10.1016/S2213-2600(21)00511-7

4. Bhatt SP, Balte PP, Schwartz JE, Cassano PA, Couper D, Jacobs DR Jr, et al. Discriminative accuracy of FEV1:FVC thresholds for COPD-related hospitalization and mortality. *JAMA*. (2019) 321:2438–47. doi: 10.1001/jama.2019.7233

5. Diab N, Gershon AS, Sin DD, Tan WC, Bourbeau J, Boulet LP, et al. Underdiagnosis and Overdiagnosis of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2018) 198:1130–9. doi: 10.1164/rccm.201804-0621CI

6. Gevenois PA, de Vuyst P, de Maertelaer V, Zanen J, Jacobovitz D, Cosio MG, et al. Comparison of computed density and microscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med*. (1996) 154:187–92. doi: 10.1164/ajrccm.154.1.8680679

7. Schroeder JD, McKenzie AS, Zach JA, Wilson CG, Curran-Everett D, Stinson DS, et al. Relationships between airflow obstruction and quantitative CT measurements of emphysema, air trapping, and Airways in Subjects with and without Chronic Obstructive Pulmonary Disease. *Am J Roentgenol*. (2013) 201:W460–70. doi: 10.2214/AJR.12.10102

8. Bhatt SP, Washko GR, Hoffman EA, Newell JD Jr, Bodduluri S, Diaz AA, et al. Imaging advances in chronic obstructive pulmonary disease. Insights from the genetic epidemiology of chronic obstructive pulmonary disease (COPDGene) study. *Am J Respir Crit Care Med*. (2019) 199:286–301. doi: 10.1164/rccm.201807-1351SO

9. Galbán CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, et al. Computed tomography–based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. *Nat Med*. (2012) 18:1711–5. doi: 10.1038/nm.2971

10. González G, Ash SY, Vegas-Sánchez-Ferrero G, Onieva Onieva J, Rahaghi FN, Ross JC, et al. Disease staging and prognosis in smokers using deep learning in chest computed tomography. *Am J Respir Crit Care Med*. (2018) 197:193–203. doi: 10.1164/rccm.201705-0860OC

11. Tang LYW, Coxson HO, Lam S, Leipsic J, Tam RC, Sin DD. Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. *The Lancet Digital Health*. (2020) 2:e259–67. doi: 10.1016/S2589-7500(20)30064-9

12. Singla S, Gong M, Riley C, Sciurba F, Batmanghelich K. Improving clinical disease subtyping and future events prediction through a chest CT-based deep learning approach. *Med Phys*. (2021) 48:1168–81. doi: 10.1002/mp.14673

13. Sun J, Liao X, Yan Y, Zhang X, Sun J, Tan W, et al. Detection and staging of chronic obstructive pulmonary disease using a computed tomography–based weakly supervised deep learning approach. *Eur Radiol*. (2022) 32:5319–29. doi: 10.1007/s00330-022-08632-7

14. Park H, Yun J, Lee SM, Hwang HJ, Seo JB, Jung YJ, et al. Deep learning–based approach to predict pulmonary function at chest CT. *Radiology*. (2023) 307:e221488. doi: 10.1148/radiol.221488

15. Almeida SD, Lüth CT, Norajitra T, Wald T, Nolden M, Jäger PF, et al. cOOpD: reformulating COPD classification on chest CT scans as anomaly detection using contrastive representations In: H Greenspan et al, editors. *Medical image computing and computer assisted intervention – MICCAI 2023*, vol. *14224*. Cham: Springer Nature Switzerland (2023)

16. Li F, Choi J, Zou C, Newell JD Jr, Comellas AP, Lee CH, et al. Latent traits of lung tissue patterns in former smokers derived by dual channel deep learning in computed tomography images. *Sci Rep*. (2021) 11:4916. doi: 10.1038/s41598-021-84547-5

17. Almeida SD, Norajitra T, Lüth CT, Wald T, Weru V, Nolden M, et al. Prediction of disease severity in COPD: a deep learning approach for anomaly-based quantitative assessment of chest CT. *Eur Radiol*. (2023). doi: 10.1007/s00330-023-10540-3

18. van der Maaten L, Hinton G. Visualizing data using t-SNE. *JMLR*. (2008) 9:2579–605.

19. Poličar PG, Stražar M, Zupan B. openTSNE: A modular Python library for t-SNE dimensionality reduction and embedding. *bioRxiv*. (2019). doi: 10.1101/731877

20. Lorenzo-Seva U. Horn's parallel analysis for selecting the number of dimensions in correspondence analysis. *Methodology*. (2011) 7:96–102. doi: 10.1027/1614-2241/a000027

21. Brock G, Pihur V, Datta S, Datta S. clValid: an *R* package for cluster validation. *J Stat Soft*. (2008) 25:4. doi: 10.18637/jss.v025.i04

22. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg*. (2018) 126:1763–8. doi: 10.1213/ANE.0000000000002864

23. Diedenhofen B, Musch J. Cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS One*. (2015) 10:e0121945. doi: 10.1371/journal.pone.0121945

24. Zou GY. Toward using confidence intervals to compare correlations. *Psychol Methods*. (2007) 12:399–413. doi: 10.1037/1082-989X.12.4.399

25. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. (1999) 8:135–60. doi: 10.1177/096228029900800204

26. Cohen J. *Statistical power analysis for the behavioral sciences*. Amsterdam: Elsevier (1977).

27. Bhatt SP, Soler X, Wang X, Murray S, Anzueto AR, Beaty TH, et al. Association between functional small airway disease and FEV$_1$ decline in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. (2016) 194:178–84. doi: 10.1164/rccm.201511-2219OC

28. Boes JL, Hoff BA, Bule M, Johnson TD, Rehemtulla A, Chamberlain R, et al. Parametric response mapping monitors temporal changes on lung CT scans in the subpopulations and intermediate outcome measures in COPD study (SPIROMICS). *Acad Radiol*. (2015) 22:186–94. doi: 10.1016/j.acra.2014.08.015

29. McDonough JE, Yuan R, Suzuki M, Seyednejad N, Elliott WM, Sanchez PG, et al. Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. *N Engl J Med*. (2011) 365:1567–75. doi: 10.1056/NEJMoa1106955

30. Hogg JC, Macklem PT, Thurlbeck WM. Site and nature of airway obstruction in chronic obstructive lung disease. *N Engl J Med*. (1968) 278:1355–60. doi: 10.1056/NEJM196806202782501

31. Chukowry PS, Spittle DA, Turner A. Small airways disease, biomarkers and COPD: where are we? *COPD*. (2021) 16:351–65. doi: 10.2147/COPD.S280157

32. Hwang HJ, Seo JB, Lee SM, Kim N, Yi J, Lee JS, et al. New method for combined quantitative assessment of air-trapping and emphysema on chest computed tomography in chronic obstructive pulmonary disease: comparison with parametric response mapping. *Korean J Radiol*. (2021) 22:1719–29. doi: 10.3348/kjr.2021.0033

| Frontiers in Medicine

Spirometry test values can be estimated from a single chest radiograph

Akifumi Yoshida[1]*, Chiharu Kai[1,2], Hitoshi Futamura[3], Kunihiko Oochi[4], Satoshi Kondo[5], Ikumi Sato[2,6] and Satoshi Kasai[1]*

[1]Department of Radiological Technology, Faculty of Medical Technology, Niigata University of Health and Welfare, Niigata, Japan, [2]Major in Health and Welfare, Graduate School of Niigata University of Health and Welfare, Niigata, Japan, [3]Konica Minolta, Inc., Tokyo, Japan, [4]Kyoto Industrial Health Association, Kyoto, Japan, [5]Graduate School of Engineering, Muroran Institute of Technology, Muroran, Japan, [6]Department of Nursing, Faculty of Nursing, Niigata University of Health and Welfare, Niigata, Japan

Introduction: Physical measurements of expiratory flow volume and speed can be obtained using spirometry. These measurements have been used for the diagnosis and risk assessment of chronic obstructive pulmonary disease and play a crucial role in delivering early care. However, spirometry is not performed frequently in routine clinical practice, thereby hindering the early detection of pulmonary function impairment. Chest radiographs (CXRs), though acquired frequently, are not used to measure pulmonary functional information. This study aimed to evaluate whether spirometry parameters can be estimated accurately from single frontal CXR without image findings using deep learning.

Methods: Forced vital capacity (FVC), forced expiratory volume in 1 s (FEV_1), and FEV_1/FVC as spirometry measurements as well as the corresponding chest radiographs of 11,837 participants were used in this study. The data were randomly allocated to the training, validation, and evaluation datasets at an 8:1:1 ratio. A deep learning network was pretrained using ImageNet. The input and output information were CXRs and spirometry test values, respectively. The training and evaluation of the deep learning network were performed separately for each parameter. The mean absolute error rate (MAPE) and Pearson's correlation coefficient (r) were used as the evaluation indices.

Results: The MAPEs between the spirometry measurements and AI estimates for FVC, FEV_1 and FEV_1/FVC were 7.59% ($r = 0.910$), 9.06% ($r = 0.879$) and 5.21% ($r = 0.522$), respectively. A strong positive correlation was observed between the measured and predicted indices of FVC and FEV_1. The average accuracy of >90% was obtained in each estimation of spirometry indices. Bland−Altman analysis revealed good agreement between the estimated and measured values for FVC and FEV_1.

Discussion: Frontal CXRs contain information related to pulmonary function, and AI estimation performed using frontal CXRs without image findings could accurately estimate spirometry values. The network proposed for estimating pulmonary function in this study could serve as a recommendation for performing spirometry or as an alternative method, suggesting its utility.

1 Introduction

Imaging tests and pulmonary function tests (PFTs) are two important examination modalities that are fundamental to respiratory medicine. Imaging tests are used to diagnose abnormalities based on the anatomy and morphology of the respiratory tract, whereas PFTs are used to evaluate the physiological functions of the respiratory tract as quantitative values. Spirometry is a relatively simple method for measuring the ventilatory performance and is performed in routine practice and as part of medical examinations. Spirometry quantitatively measures the pulmonary capacity and velocity by determining the pressure and flow rate. The results are interpreted based on the symptoms and other clinical findings. Forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV_1) can be measured using spirometry. These indices can be evaluated relative to the decline in pulmonary function by calculating the ratio of the measured values (% FVC and % FEV_1) to the representative values corresponding to the individual's age, height, and sex. Post-bronchodilator FEV_1/FVC <0.7 indicates obstructive ventilatory defects and is used as a strong diagnostic criterion (1–4). Thus, FVC, FEV_1 and FEV_1/FVC are important clinical assessment indices (5, 6). They allow for earlier detection of diseases that affect pulmonary function, such as chronic obstructive pulmonary disease (COPD) and asthma, than imaging tests. Spirometry remains the gold standard for diagnosing ventilatory defects (2). It can detect asymptomatic cases with obstructive ventilatory defects as well as cases of impaired pulmonary function, even in the absence of obstructive ventilatory defects (7–11). Conversely, spirometry is usually performed in symptomatic patients (12), low uptake compared to that in chest radiography is the major problem of spirometry in preventive medicine. Moreover, participants must cooperate during the test and breathe with effort to obtain accurate results. Low throughput is an additional issue. Throughput is further limited in cases that require infection control measures. Thus, spirometry must be encouraged, and alternative tests with good throughput must be developed to overcome the challenges in performing PFTs during clinical examinations.

Imaging tests are associated with high throughput and a relatively high screening uptake rate. Chest radiographs (CXRs) remain the first choice of imaging test for cardiopulmonary screening and are commonly acquired during routine primary care, including health checkups. The CXR can visually identify morphological abnormalities in the lungs and other thoracic regions and can detect various diseases, for example, pneumonia and lung cancer. If the CXR shows abnormal findings related to pulmonary function, such as emphysema in COPD, this can be detected without spirometry. However, it is difficult to detect lesions that cause abnormal pulmonary function at an early stage with CXR, and, therefore, it is generally not used to assess pulmonary function. Thus, spirometry and CXR are complementary and have advantages and disadvantages. If cases with functional abnormalities can be detected in CXR without detectable image findings, it may lead to the creation of health-promoting opportunities for patients. Hence, it would be clinically useful if pulmonary function could be accurately obtained from the CXR.

Previous studies have estimated pulmonary function using the shape of the rib cage on CXRs acquired during static imaging (13–16). Similarly, studies have investigated the relationship between image characteristics and pulmonary function on dynamic chest X-ray radiographs (DCRs) acquired during dynamic imaging (17, 18). Pulmonary function has been estimated using image

characteristics measured from landmarks in the images and regression models or equations; however, the accuracy of the estimated values was limited as the correlation between image features and lung function was not high. Furthermore, it requires manual measurement of image characteristics, a labor-intensive task, and may lead to errors. Machine learning has resulted in breakthroughs in medical image analysis in recent years, and several studies have used general image recognition models in medical image analysis and the estimation of functional parameters and other information from images (19). Sogancioglu et al. (20) reported the use of artificial intelligence (AI) for the estimation of the lung volume from pseudo-CXRs calculated from CT images. However, the estimated lung volumes were calculated from CT image data and not pulmonary function values. Schroeder et al. (21) reported the estimation of the % predFEV$_1$ and FEV$_1$/FVC as PFT values from bidirectional CXR pairs using deep learning. The study used two-view CXRs including imaging findings for estimation, not frontal CXRs alone. It was not clear whether pulmonary function impairment could be estimated from CXRs without imaging findings. Health checkups are performed routinely under the national system in Japan, and almost all adults undergo CXR screening. However, CXR screening is not always performed bidirectionally. It is important to determine whether accurate pulmonary function values can be obtained from frontal CXR images to develop an AI system for estimating pulmonary function from CXRs that can be used during medical examinations worldwide, including in developing countries.

Therefore, this study aimed to estimate the spirometry measurements from single frontal CXR without image findings using a general image recognition model and evaluate the precision of the estimation.

2 Materials and methods

This study was conducted after receiving approval for the use of medical data obtained during medical examinations from the Institutional Review Board of the Niigata University of Health and Welfare and the data-providing institutions (Approval number: 18952-221124).

2.1 Materials

2.1.1 Data

Frontal CXRs acquired at a single institution in Japan for 2019 were used in this study. The CXR images in 8-bit PNG format were used. Figure 1 shows a representative CXR. The FVC, FEV_1 and FEV_1/FVC values obtained via forced vital capacity testing were used as the pre-bronchodilator spirometry data, as described in multiple COPD studies (22–24). Figure 2 presents the inclusion and exclusion criteria for the CXR and PFT data. The dataset used in this study are cases with no image findings noted in the radiology reports of the screening CXR. Cases with any abnormal findings such as lung opacities, lung cancer or other pulmonary disease, pleural lesions, cardiovascular lesions, musculoskeletal lesions, tracheal abnormalities, postoperative and supported devices were excluded. The CXRs in the dataset does not include any image findings noted,

including inactive findings. The CXR data and corresponding PFT data were extracted from only one sample per participant. A total of 11,837 data samples, including the corresponding heights, sexes, and ages, were included in the PFT data; there were no missing data values. Table 1 presents the demographic characteristics of the datasets. A total of 9,469, 1,184, and 1,184 samples were used for the training, validation, and test of the deep learning network to ensure that the data ratio was maintained at 8:1:1.

FIGURE 1
Sample chest radiographs used in this study. The original images were down-sampled and zero-padded to a 512×512 matrix with the aspect ratio preserved. Additionally, they were resampled to 224×224 and used as input.

2.1.2 Experimental environment

MATLAB 2022a (MathWorks, Inc.) was used to implement the framework for performing the deep learning operations. Image processing and deep learning computations were performed using MATLAB in this study.

2.2 Methods

2.2.1 Network training and evaluation

In addition to the pre-training data from the ImageNet classification task, ResNet-18, ResNet-50, ResNet-101, DenseNet-201, and Inception-ResNet-V2, which are publicly available in the MATLAB add-in library, were used as the initial weights ([25], [26]). The fully connected layers closest to the output layer of each network were replaced with a new layer with an output class of one. The training conditions were as follows: optimization method, Adam; loss function root mean square error; batch size, 32–256 (variable); initial learning rate, 1×10^{-5}; maximum number of epochs, 50; image data augmentation, $\pm 5°$ random rotation/random horizontal flip/$\pm 5\%$ random scaling. The batch size was varied for each network type and then optimized. The network weights were updated using the training dataset, and the network performance at each epoch was displayed using the validation dataset. The weights in the epoch with the lowest loss for the validation dataset were saved to complete the learning. Network training and estimation were performed separately for FVC, FEV_1 and FEV_1/FVC.

2.2.2 Evaluation

CXRs from the test dataset and the FVC or FEV_1 estimations were the input and output of the network, respectively. The mean average percentage error (MAPE) and Pearson's correlation coefficient (r) between the reference measured values and network-estimated values were used as the evaluation indices. Bland–Altman analysis ([27]) was
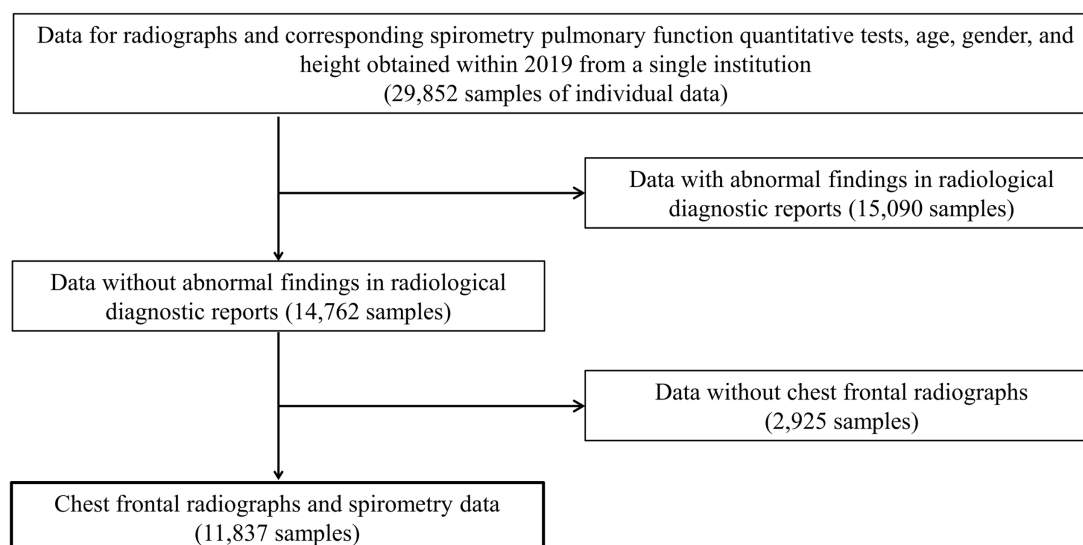
FIGURE 2
The inclusion and exclusion criteria for data acquisition. Only frontal chest radiographs and spirometry data obtained at a single institution with no abnormal findings on diagnostic reports and no history of undergoing radiography and spirometry on the same day were used.

TABLE 1 Demographic characteristics and pulmonary function indices of datasets.

		Training	Validation	Test	Overall
Participant	n	9,469	1,184	1,184	11,837
Sex	Female, n (%)	3,325 (35.1%)	425 (35.9%)	422 (35.6%)	4,172 (35.2%)
	Male, n (%)	6,144 (64.9%)	759 (64.1%)	762 (64.4%)	7,665 (64.8%)
Age (year)		50 [43–58]	50 [43–58]	50 [43–58]	50 [43–58]
Height (cm)		167.3 [160.5–172.8]	166.8 [160.2–172.6]	167.3 [160.1–173.0]	167.3 [160.4–172.8]
FVC (mL)		3,600 [2,980–4,210]	3,680 [3,038–4,320]	3,610 [2,960–4,180]	3,600 [2,980–4,210]
FEV1 (mL)		2,880 [2,410–3,400]	3,585 [2,978–4,240]	2,900 [2,398–3,413]	2,880 [2,410–3,400]
FEV1/FVC (%)		81.1 [77.0–84.8]	81 [76.8–84.8]	81.1 [77.0–84.7]	81.1 [77.0–84.8]
predFVC (mL)		3,890 [3,010–4,310]	3,880 [2,980–4,320]	3,885 [3,000–4,330]	3,890 [3,010–4,310]
predFEV1 (mL)		3,260 [2,530–3,700]	3,250 [2,498–3,710]	3,245 [2,520–3,710]	3,260 [2,520–3,700]
% FVC		97.8 [89.9–105.6]	97.8 [89.3–106.2]	97.3 [89.8–105.4]	97.6 [89.5–105.8]
% FEV1		97.7 [89.5–105.8]	93.8 [85.0–101.9]	93.4 [85.9–101.3]	93.6 [85.4–101.9]
% FEV1 category, n (%)	>80%	10,281 (86.9%)	1,020 (86.1%)	1,027 (86.7%)	10,281 (86.9%)
	50%–79%	1,528 (12.9%)	161 (13.6%)	154 (13%)	1,528 (12.9%)
	30%–49%	24 (0.2%)	1 (0.1%)	3 (0.3%)	24 (0.2%)
	<30%	4 (0%)	2 (0.2%)	0 (0%)	4 (0%)

Age, height, FVC, FEV$_1$, FEV$_1$/FVC, predFVC, predFEV1, % FVC, % FEV$_1$ are shown as median [25–75th percentile values]. FVC, forced vital capacity; FEV$_1$, forced expiratory volume in 1 s; predFVC, predicted FVC; pred FEV$_1$, predicted FEV$_1$; % FVC, percent-predicted FVC; % FEV$_1$, percent-predicted FEV$_1$.

TABLE 2 Comparison of estimation performance of the network for each pulmonary function indices.

Network	FVC		FEV$_1$		FEV$_1$/FVC	
	MAPE	r	MAPE	r	MAPE	r
ResNet-18	7.995	0.903	9.442	0.865	5.344	0.492
ResNet-50	8.083	0.907	9.063	0.871	5.885	0.465
ResNet-101	8.206	0.909	9.328	0.866	5.210	0.522
DenseNet-201	8.101	0.909	9.249	0.876	5.565	0.466
Inception-ResNet-V2	7.585	0.910	9.055	0.879	5.674	0.512

FVC, forced vital capacity; FEV$_1$, forced expiratory volume in 1 s.

performed using the reference measured value and the error between the estimated value and the measured value. The estimated value and the measured value were considered to be variables that could be treated equally if >95% of the evaluation data were included in the limits of agreement (LOA) at mean ± 1.96 SD.

3 Results

Table 2 presents the results of FVC, FEV$_1$ and FEV$_1$/FVC estimations for each network. FVC and FEV$_1$ estimates showed strong positive correlations with both networks. Inception-ResNet-V2, which had the largest number of parameters, achieved the best MAPE and correlation coefficients for FVC and FEV$_1$. The MAPE and correlation coefficients for FVC estimation were 7.585–8.246 and 0.903–0.910, respectively. The MAPE and correlation coefficients for FEV$_1$ estimation were 9.055–9.442 and 0.865–0.879, respectively. The MAPE and correlation coefficients for FVC estimation were superior to those of FEV$_1$ estimation, regardless of the network type used. Figure 3 presents

the results of the comparison between the FVC estimation results of the Inception-ResNet-V2 network, which yielded the lowest MAPE and the highest correlation coefficient, and the reference. Figure 4 presents the results of the comparison between the FEV$_1$ estimation results and the reference. The 95% confidence interval for the mean error rate of FVC estimation (Figure 3B) ranged between −1.741% and −0.615% in the Bland–Altman plot. The slope of the coefficient for the determination of the % error-reference of an approximately straight line, $R^2 = 0.106$, was not significant. The agreement between the estimated and measured FVC values was confirmed, as 96.1% of the data were included within the LOA. The 95% confidence interval for the mean percentage error of FEV$_1$ estimation ranged between 0.606% and 2.164% in the Bland–Altman plot (Figure 4B). The slope of the coefficient for the determination for the % error-reference of an approximately straight line, $R^2 = 0.157$, was not significant. The agreement between the FEV$_1$ estimates and measured values was confirmed, as 97.6% of the data were included within the LOA. Figure 5 presents the results of the deep learning network with the best correlation coefficient and MAPE for estimating FEV$_1$/FVC. The MAPE was acceptable at 5.20%, whereas the

FIGURE 3
FVC estimation results using Inception-ResNet-V2. **(A)** Comparison of measured and estimated values. **(B)** Bland–Altman-like plot presenting the measured value-estimated error rate relationship. The correlation coefficient and error rate were the best among the networks used, with 96.1% of the data within the limits of agreement (mean $\pm$ 1.96 SD), confirming agreement between spirometry measurements and AI estimation using chest radiography. FVC, forced vital capacity.

FIGURE 4
FEV_1 estimation results using Inception-ResNet-V2. **(A)** Comparison of measured and estimated values. **(B)** Bland–Altman-like plot representing the measured value-estimated error rate relationship. The correlation coefficient and error rate were the best among the networks used, with 97.6% of the data within the limits of agreement (mean $\pm$ 1.96 SD), confirming agreement between spirometry measurements and AI estimation using chest radiography. FEV_1, forced expiratory volume in 1 s.

correlation was moderate at $r = 0.522$. The correlation between FEV_1/FVC estimates and measured values was weaker than those observed for the estimation of FVC and FEV_1. The 95% confidence interval for the mean error rate of FVC estimation (Figure 5B) ranged between -221213.6% and 15.7% in the Bland–Altman plot. The slope of the coefficient for the determination for the % error-reference of an approximately straight line, $R^2 = 0.759$, was significant. The agreement between the FEV_1/FVC estimates and measured values was confirmed, as 96.8% of the data were included within the LOA.

4 Discussion

A typical deep learning network was used to estimate the FVC, FEV_1 and FEV_1/FVC values from a frontal CXR via spirometry in this study. Strong positive correlations were observed between the estimated FVC and FEV_1 values and the corresponding measured values. The MAPE was low (<10%) for FVC, FEV_1 and FEV_1/FVC estimations. The Bland–Altman analysis revealed good agreement between the estimated and measured values for FVC and FEV_1. Thus,
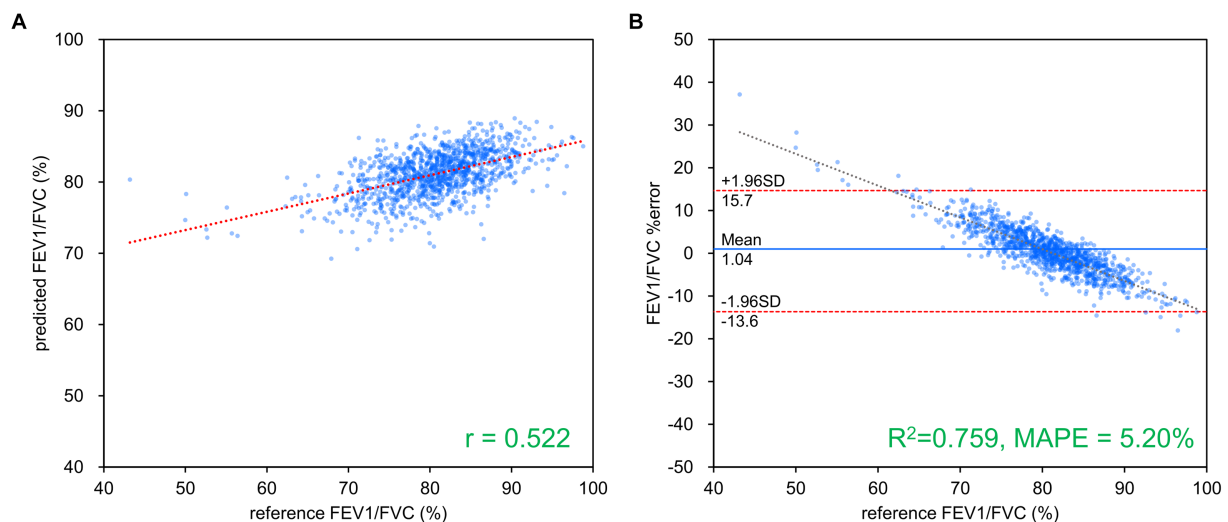
FIGURE 5
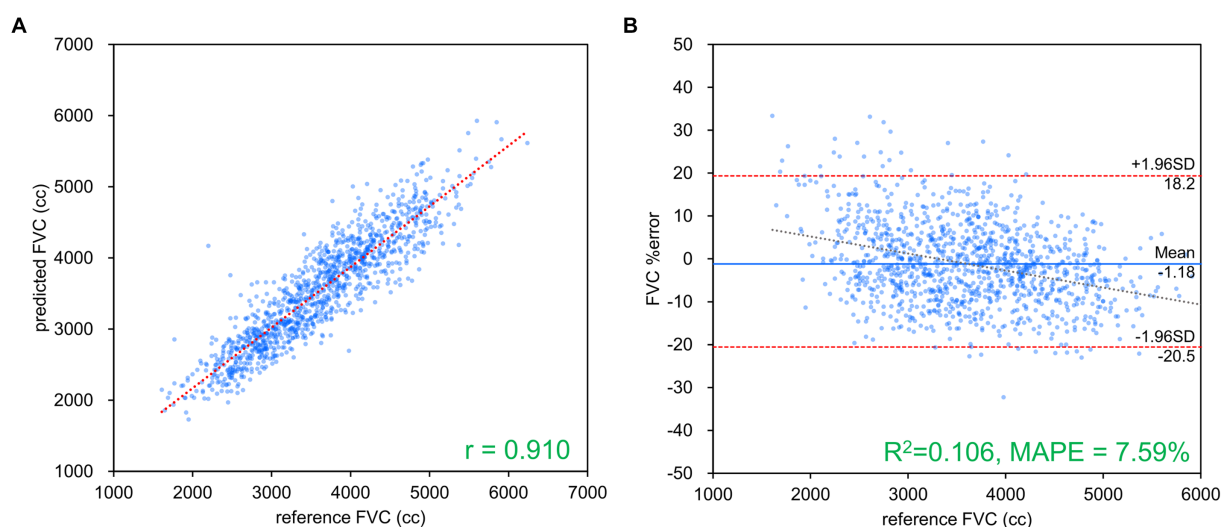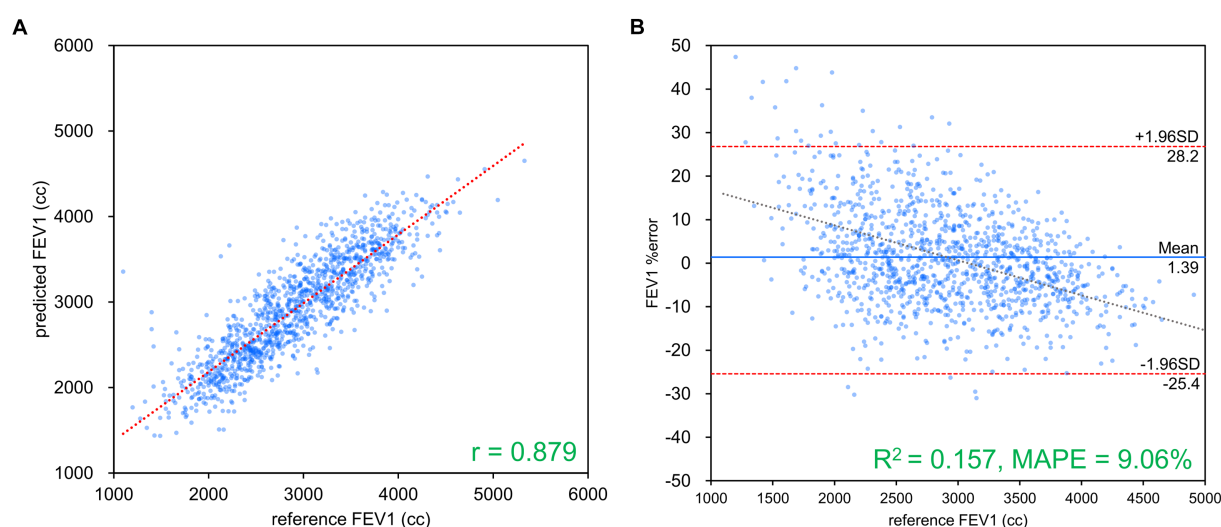FEV_1/FVC estimation results using ResNet-101. **(A)** Comparison of measured and estimated values. **(B)** Bland–Altman-like plot representing the measured value-estimated error rate relationship. The correlation between the estimated and measured values was moderate, while the error rate was low at about 5%. FVC, forced vital capacity; FEV_1, forced expiratory volume in 1 s.

the findings of this study indicate that frontal CXRs contain information related to pulmonary function and that AI estimation performed using frontal CXRs can estimate spirometry-measured values FVC and FEV_1 with high accuracy.

The pulmonary function parameters to be estimated in this study were FVC and FEV_1, which are expiratory volumes exhaled during forced breathing with no time limit. FVC is the total expiratory volume exhaled during forced breathing without any time limit, whereas FEV_1 is the expiratory volume exhaled during the first second of forced breathing. Thus, FEV_1 can be considered a part of FVC, where FEV_1 is the flow velocity. FEV_1, a highly sensitive indicator of decreased ventilatory capacity, is decreased in patients with obstructive ventilatory defects owing to air trapping caused by damaged alveoli, which increases the peripheral airway resistance and limits the expiratory volume that can be exhaled in a short period of time (28–30). This decrease in FEV_1 is particularly significant in patients with progressive COPD; however, it can also be observed in the pre-COPD stage and early stages of COPD, wherein the decrease in ventilation capacity is less evident (31). Specific findings are observed on the CXRs of patients with severe COPD; however, such findings are not observed in patients with early-stage COPD. Therefore, it is reasonable to assume that the accuracy of FEV_1 estimation is relatively inferior to that of FVC estimation, an index that varies more frequently among patients. The correlation of the estimated FEV_1/FVC and those of measurements was weaker than those observed for the case of estimation of FVC and FEV_1. This may be attributed to the individual variability of FVC and FEV_1, which makes the FEV_1/FVC value a more complex predictor.

Subgroups were created based on the age, height, sex, % FVC, and % FEV_1 related to the estimation error to increase the robustness of the performance of the AI estimation method used in this study. Age, height, and sex are the information used to determine the % FVC and % FEV_1 in spirometry. The % FVC and % FEV_1 are relative to the predicted FVC and FEV_1 values, respectively, which are standard values for the same age, height, and sex expressed as percentages. Thus, %

FVC and % FEV_1 are indicators of a participant's pulmonary function relative to the standard population. Each subgroup, except for the subgroup created on the basis of sex, was divided into categories, and the error rates for each category were compared. The categories for each subgroup were as follows: age category, <30 years, 30–49 years, 50–59 years, 60–69 years, and >70 years; height category, <150 cm, 150–160 cm, 160–170 cm, 170–180 cm, and >180 cm; sex category, male and female; % FVC category, <70, 70–80, 80–90, 90–100, 100–110, 110–120, and >120, % FEV_1 category <70, 70–80, 80–90, 90–100, 100–110, 110–120, and >120. Differences in the distributions of error rates among categories were tested using the Kruskal–Wallis method (significance level $p < 0.05$) and multiple comparisons. Figure 6 presents the distributions of error rates in the FVC estimation according to the subgroup and category. The distribution of error rates tended to widen with increasing age in the age category (Figure 6A); however, multiple comparisons performed using the Kruskal–Wallis test revealed no statistically significant differences among the categories. There were no trends or significant differences in height, sex, or % FVC subgroups (Figure 6B-D). Significant differences were observed in the distribution between the categories with low % FEV_1 and the other categories in the % FEV_1 subgroup ($p < 0.001$). Figure 7 presents the distributions of error rates in the FEV_1 estimation according to the subgroup and category. No significant differences were observed between the categories in terms of age, height, or sex. Thus, the findings suggest that robust performance was obtained without error bias for age, height, and sex. Significant differences were observed in the errors between categories with low % FVC values and several other categories in the % FVC subgroup. Multiple comparisons revealed the relationship between each % FEV_1 category and the FEV_1 estimation error rate (Figure 7E). Significantly different mean ranks were observed between all categories except between categories 100–110 and 110–120 and between 110–120 and >120 in the % FEV_1 subgroup ($p < 0.001$). Figure 8 presents the distributions of error rates in the FEV_1/FVC estimation according to the subgroup and category. No significant differences were observed between the categories in

FIGURE 6
Relationship between the subgroups and FVC estimation error rates in the evaluation data. **(A)** Percentage error by age category. **(B)** Percentage error by height category. **(C)** Percentage error by gender. **(D)** Percentage error per % FVC category. **(E)** Percentage error by % FEV$_1$ category. The higher the age category and the lower the % FEV$_1$ category, the larger the variance of the percentage error tended to be. FVC, forced vital capacity; FEV$_1$, forced expiratory volume in 1 s.

FIGURE 7
Relationship between the subgroups and FEV$_1$ estimation error rates in the evaluation data. **(A)** Percentage error by age category. **(B)** Percentage error by height category. **(C)** Percentage error by gender. **(D)** Percentage error per % FVC category. **(E)** Percentage error by % FEV$_1$ category. The variance of the percent error tended to be larger for the higher age categories and for the lower % FVC and % FEV$_1$ categories. FVC, forced vital capacity; FEV$_1$, forced expiratory volume in 1 s.

terms of height and sex. Thus, the findings suggest that a robust performance was obtained without error bias for height and sex. Multiple comparisons revealed the relationship between each % FEV$_1$ category and the FEV$_1$/FVC estimation error rate (Figure 8E). Significantly different mean ranks were observed between categories between categories 30–39 years and 50–59 years in the age subgroup ($p < 0.01$) and between categories between categories 80–90 and >120 in the % FVC subgroup. Additionally, in the % FEV$_1$ subgroup, significantly different mean ranks were observed between all categories ($p < 0.05$), except between categories <70 and 70–80; 70–80 and 80–90; 80–90 and 100–110; 100–110 and 110–120 and >120; 110–120 and >120. Table 3 presents the number of data points and error rates for

each subgroup and category. The median error tended to be more positively biased for categories with a lower % FEV$_1$ in the % FEV$_1$ subgroup. It is suspected that the low % FVC and low % FEV$_1$ categories had small samples and that the characteristics of % FVC and % FEV$_1$ might not have been learned sufficiently. However, the error rate did not increase significantly for the other age and height categories with a lesser amount of data. Therefore, the results of this study do not exclude the possibility that the relationship between lower % FVC and % FEV$_1$ and imaging features has not been sufficiently trained by network. Future studies should increase the number of samples with low % FVC and % FEV$_1$ during training and validate the robustness of the % FVC and % FEV$_1$ subgroups.
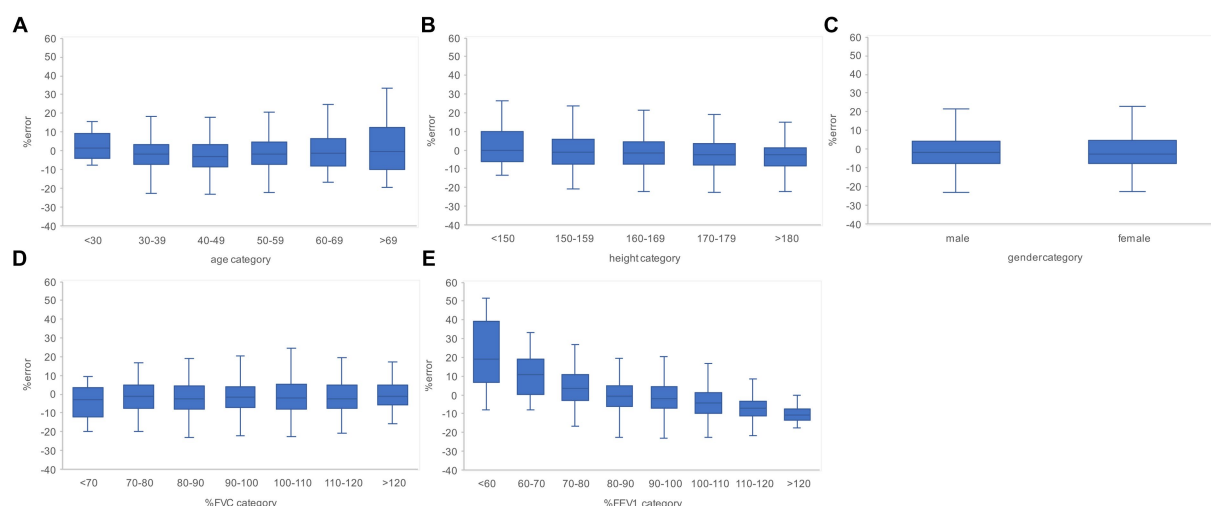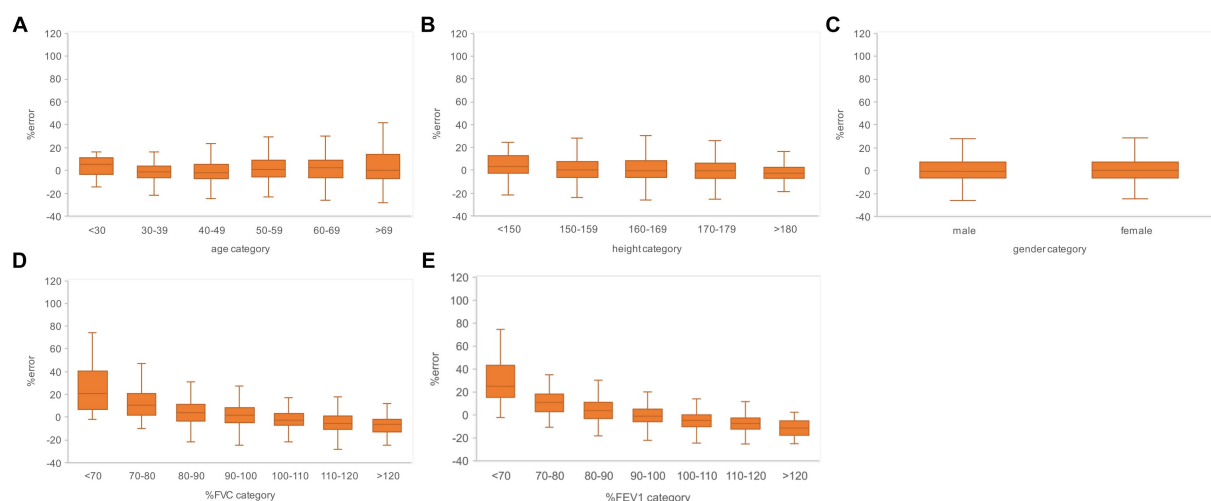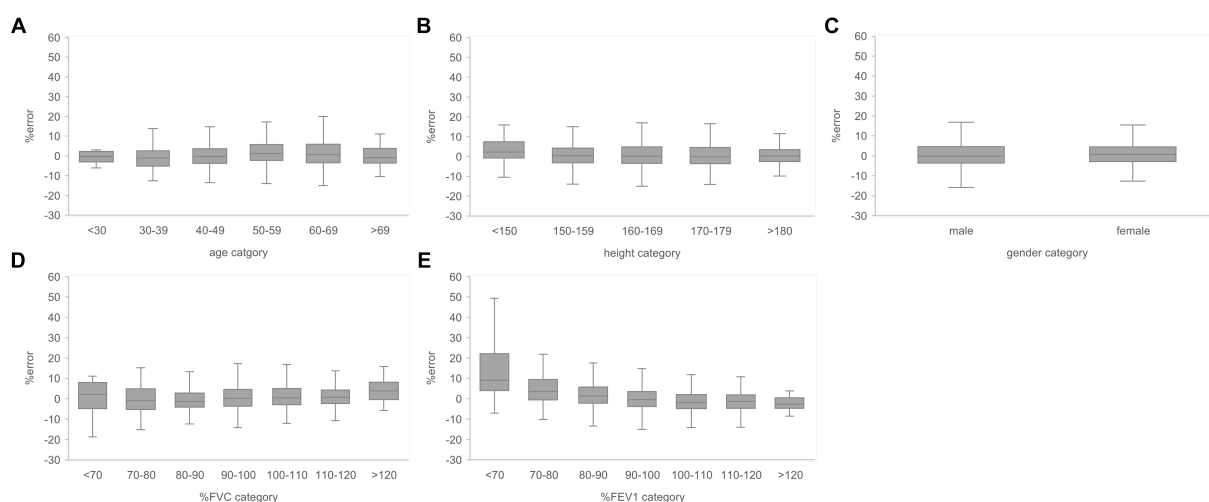
FIGURE 8
Relationship between the subgroups and FEV1/FVC estimation error rates in the evaluation data. **(A)** Percentage error by age category. **(B)** Percentage error by height category. **(C)** Percentage error by gender. **(D)** Percentage error per % FVC category. **(E)** Percentage error by % FEV1 category. The variance of the percent error tended to be larger for the lower % FEV1 categories. FVC, forced vital capacity; FEV$_1$, forced expiratory volume in 1s.

Previous studies investigating the relationship between CXRs and pulmonary function manually extracted image characteristics from dynamic CXRs to investigated their correlations with pulmonary function. Hino et al. (17) and Hida et al. (18) investigated the correlation between image characteristics and pulmonary function values on dynamic CXR. In the study by Hino et al. (17), the highest correlation coefficient was between lung field area and FEV$_1$ at the maximal inspiratory position with effort breathing in DCR was the highest (right $r = 0.59$, left $r = 0.62$). The study by Hida et al. (18) also showed the highest correlation between whole excursion of the diaphragm and FEV$_1$ in DCR, although the correlation was weak (right $r = 0.27$, left $r = 0.38$). While previous studies have reported a moderate correlation between FEV$_1$ and lung field area based on DCR image measurements, this study used deep learning networks to automatically extract and select image characteristics from static CXRs and revealed a strong positive correlation ($r = 0.879$) between CXRs and FEV$_1$. The findings of this study suggest that pulmonary function can be estimated accurately from static images using deep learning networks, resulting in a significant improvement in accuracy. In a previous study in which pulmonary function was estimated from CXR using machine learning, Schroeder et al. (21) estimated FEV$_1$/FVC using bidirectional CXR pair and obtained $R^2 = 0.415$ (conversion $r = 0.644$), which is a moderately positive correlation. In this study, only frontal CXR was used to estimate FEV1/FVC. An $R^2 = 0.272$ ($r = 0.522$) was obtained, indicating a moderate positive correlation. The absence of lateral CXR in this study is expected to have resulted in the deep learning network extracting less information compared to if bidirectional CXR pairs were utilized, leading to lower estimation performance.

Among pathologies with obstructive ventilation defects, COPD is the most common chronic respiratory disease worldwide, with approximately 174 million affected individuals (32). COPD is an irreversible pathology; thus, it is important to detect and initiate treatment prior to its progression. However, the symptoms of COPD only become apparent as the disease reaches advanced stages. Moreover, it is difficult to detect COPD early using CXRs. Therefore, detecting and initiating treatment at the earliest possible stage for patients with COPD who are asymptomatic has become an important public health issue worldwide. In this study, the FVC and FEV$_1$ values measured using spirometry could be estimated with an average accuracy of >90% using only frontal CXRs, which are the most commonly acquired images in imaging tests, in this study. The method used in this study provides spirometry estimates without any additional burden to the CXR examinee. In the future, if the robustness of the estimation performance to the characteristics of the data is sufficiently verified, estimation of pulmonary function using CXR could be used as an adjunct to spirometry in individuals with low estimated pulmonary function or as an alternative to pulmonary function measurement. Chest radiography (screening CXR) is a low-cost and relatively widespread cancer screening method that can be used as an alternative for the COPD risk assessment. The findings of this study suggest that FVC and FEV$_1$ could be estimated with an average accuracy of >90% and >87% for participants with % FEV$_1$ of >80% and >70%, respectively. Thus, the network developed in this study could be used as an alternative for COPD risk assessment in patients with mildly impaired pulmonary function and for the control of the pre-COPD group.

This study has some limitations. Only cases with no abnormal findings in the CXR report were used to eliminate the influence of abnormal findings on the estimation of pulmonary function by image features of abnormal findings. Another reason is that it is significant for use in estimating pulmonary function is CXR without abnormal findings related to abnormalities in pulmonary function. However, the available training data can be expected to increase and a higher network performance can be achieved if the pulmonary function can be estimated accurately, even in cases with abnormal findings. The results of this study did not exclude the possibility of inferior estimation performance by deep learning for

TABLE 3 Number of test data and percentage error of AI estimation according to the age, height, sex, % FVC and % FEV$_1$.

Subgroup	Category	n	Absolute percentage error (mean $\pm$ SD)		
			FVC	FEV$_1$	FEV$_1$/FVC
Age (year)	<30	12	6.43 ± 4.98	7.96 ± 5.30	2.96 ± 3.30
	30–39	135	6.86 ± 5.56	7.25 ± 6.18	4.82 ± 3.63
	40–49	420	7.11 ± 8.44	7.84 ± 9.76	4.70 ± 6.45
	50–59	382	7.50 ± 7.40	9.57 ± 13.12	5.23 ± 4.96
	60–69	181	8.43 ± 6.87	11.07 ± 13.03	6.58 ± 8.64
	>69	54	11.00 ± 7.38	12.76 ± 12.36	5.74 ± 6.07
Height (cm)	<150	31	7.97 ± 7.00	10.21 ± 9.6	5.37 ± 4.36
	150–159	264	7.94 ± 6.48	8.94 ± 8.13	4.67 ± 3.72
	160–169	437	7.66 ± 7.23	9.86 ± 13.60	5.83 ± 7.22
	170–179	399	7.33 ± 5.45	8.37 ± 7.13	4.94 ± 4.05
	>180	53	6.76 ± 5.86	7.33 ± 9.57	4.64 ± 5.53
Sex	Male	762	7.46 ± 6.63	9.10 ± 11.46	5.35 ± 5.83
	Female	422	7.79 ± 6.07	8.96 ± 7.93	4.94 ± 4.73
% FVC	<70	16	7.38 ± 5.88	38.24 ± 52.84	9.77 ± 15.57
	70–80	59	7.29 ± 9.18	13.12 ± 12.57	5.63 ± 6.98
	80–90	227	7.41 ± 9.21	9.36 ± 11.44	5.3 ± 7.18
	90–100	386	7.33 ± 10.15	8.9 ± 12.54	5.42 ± 8.76
	100–110	316	7.9 ± 9.64	6.96 ± 8.69	4.85 ± 6.03
	110–120	143	8.02 ± 10.99	8.63 ± 8.96	4.58 ± 5.98
	>120	37	7.3 ± 10.35	9.13 ± 8.22	5.12 ± 5.29
% FEV1	<70	40	16.3 ± 14.61	35.36 ± 24.07	16.14 ± 17.98
	70–80	117	8.29 ± 9.84	12.37 ± 10.46	6.72 ± 7.46
	80–90	292	7.01 ± 9.1	8.43 ± 9.93	4.98 ± 6.28
	90–100	394	6.71 ± 8.24	6.69 ± 8.36	4.64 ± 5.86
	100–110	223	7.43 ± 7.92	7.25 ± 7.69	4.41 ± 5.33
	110–120	94	8.22 ± 6.7	9.1 ± 8.03	3.99 ± 4.88
	>120	24	9.73 ± 5.02	11.82 ± 7.1	3.73 ± 4.01

FVC, forced vital capacity; FEV$_1$, forced expiratory volume in 1 s; % FVC, percent-predicted FVC; % FEV$_1$, percent-predicted FEV$_1$.

cases with lower % FVC and % FEV$_1$. To validate and further generalize the findings of this study, it will be necessary to train a larger number of samples with low % FVC and % FEV$_1$ and to perform external validation using data from another facility. Only ImageNet-pretrained networks publicly available in MATLAB and general deep-learning networks were used in this study. Depending on the samples and networks used, a larger network scale had greater correlation coefficient and MAPE. Thus, it is possible that larger deep learning networks can be used to develop pulmonary function estimation networks with higher performance.

5 Conclusion

Pulmonary function values measured using spirometry were estimated from the corresponding frontal CXRs using a general deep learning network. FVC, FEV$_1$ and FEV$_1$/FVC were estimated with an

average accuracy of >90%. The pulmonary function estimation network developed in this study may be a useful method for pulmonary function screening or a potential substitute for spirometry.

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 humans were approved by Ethics Committee of Niigata University of Health and Welfare. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived

the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because this study was conducted using only anonymously processed information provided by Konica Minolta, Inc.

Author contributions

AY: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. CK: Writing – review & editing. HF: Resources, Writing – review & editing. KO: Resources, Writing – review & editing. SKo: Writing – review & editing. IS: Writing – review & editing. SKa: Conceptualization, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

HF was employed by Konica Minolta, Inc.

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.

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.

References

1. Buist AS, McBurnie MA, Vollmer WM, Gillespie S, Burney P, Mannino DM, et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet*. (2007) 370:741–50. doi: 10.1016/S0140-6736(07)61377-4

2. Agustí A, Celli BR, Criner GJ, Halpin D, Anzueto A, Barnes P, et al. Global initiative for chronic obstructive lung disease 2023 report: GOLD executive summary. *Eur Respir J*. (2023) 61:2300239. doi: 10.1183/13993003.00239-2023

3. Fouka E, Papaioannou AI, Hillas G, Steiropoulos P. Asthma-COPD overlap syndrome: recent insights and unanswered questions. *J Pers Med*. (2022) 12:708. doi: 10.3390/jpm12050708

4. Celli B, Fabbri L, Criner G, Martinez FJ, Mannino D, Vogelmeier C, et al. Definition and nomenclature of chronic obstructive pulmonary disease: time for its revision. *Am J Respir Crit Care Med*. (2022) 206:1317–25. doi: 10.1164/rccm.202204-0671PP

5. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med*. (2004) 350:1005–12. doi: 10.1056/NEJMoa021322

6. Hurst JR, Anzueto A, Vestbo J. Susceptibility to exacerbation in COPD. *Lancet Respir Med*. (2017) 5:e29. doi: 10.1016/S2213-2600(17)30307-7

7. Lindberg A, Jonsson AC, Rönmark E, Lundgren R, Larsson LG, Lundbäck B. Ten-year cumulative incidence of COPD and risk factors for incident disease in a symptomatic cohort. *Chest*. (2005) 127:1544–52. doi: 10.1378/chest.127.5.1544

8. Kalhan R, Dransfield MT, Colangelo LA, Cuttica MJ, Jacobs DR Jr, Thyagarajan B, et al. Respiratory symptoms in young adults and future lung disease. The CARDIA lung study. *Am J Respir Crit Care Med*. (2018) 197:1616–24. doi: 10.1164/rccm.201710-2108OC

9. Allinson JP, Hardy R, Donaldson GC, Shaheen SO, Kuh D, Wedzicha JA. The presence of chronic mucus hypersecretion across adult life in relation to chronic obstructive pulmonary disease development. *Am J Respir Crit Care Med*. (2016) 193:662–72. doi: 10.1164/rccm.201511-2210OC

10. Guerra S, Sherrill DL, Venker C, Ceccato CM, Halonen M, Martinez FD. Chronic bronchitis before age 50 years predicts incident airflow limitation and mortality risk. *Thorax*. (2009) 2009:894–900. doi: 10.1136/thx.2008.110619

11. Lange P, Celli B, Agustí A. Lung-function trajectories and chronic obstructive pulmonary disease. *N Engl J Med*. (2015) 373:1575. doi: 10.1056/NEJMc1510089

12. MacIntyre NR, Selecky PA. Is there a role for screening spirometry? *Respir Care*. (2010) 55:35–42.

13. Barnhard HJ, Pierce JA, Joyce JW, Bates JH. Roentgenographic determination of total lung capacity. A new method evaluated in health, emphysema and congestive heart failure. *Am J Med*. (1960) 28:51–60. doi: 10.1016/0002-9343(60)90222-9

14. Loyd HM, String ST, DuBois AB. Radiographic and plethysmographic determination of total lung capacity. *Radiology*. (1966) 86:7–14. doi: 10.1148/86.1.7

15. Pratt PC, Klugh GA. A method for the determination of total lung capacity from posteroanterior and lateral chest roentgenograms. *Am Rev Respir Dis*. (1967) 96:548–52. doi: 10.1164/arrd.1967.96.3.548

16. Rodenstein DO, Sopwith T, Denison DM, Stanescu DC. Reevaluation of the radiographic method for measurement of total lung capacity. *Bull Eur Physiopathol Respir*. (1985) 21:521–5.

17. Hino T, Hata A, Hida T, Yamada Y, Ueyama M, Araki T, et al. Projected lung areas using dynamic X-ray (DXR). *Eur J Radiol Open*. (2020) 7:100263. doi: 10.1016/j.ejro.2020.100263

18. Hida T, Yamada Y, Ueyama M, Araki T, Nishino M, Kurosaki A, et al. Time-resolved quantitative evaluation of diaphragmatic motion during forced breathing in a health screening cohort in a standing position: dynamic chest phrenicography. *Eur J Radiol*. (2019) 113:59–65. doi: 10.1016/j.ejrad.2019.01.034

19. Çallı E, Sogancioglu E, van Ginneken B, van Leeuwen KG, Murphy K. Deep learning for chest X-ray analysis: a survey. *Med Image Anal*. (2021) 72:102125. doi: 10.1016/j.media.2021.102125

20. Sogancioglu E, Murphy K, Scholten TE, Boulogne LH, Prokop M, van Ginneken B. Automated estimation of total lung volume using chest radiographs and deep learning. *Med Phys*. (2022) 49:4466–77. doi: 10.1002/mp.15655

21. Schroeder JD, Bigolin Lanfredi R, Li T, Chan J, Vachet C, Paine R, et al. Prediction of obstructive lung disease from chest radiographs via deep learning trained on pulmonary function data. *Int J Chron Obstruct Pulmon Dis*. (2020) 15:3455–66. doi: 10.2147/COPD.S279850

22. Tseng H, Henry TS, Veeraraghavan S, Mittal PK, Little BP. Pulmonary function tests for the radiologist. *Radiographics*. (2017) 37:1037–58. doi: 10.1148/rg.2017160174

23. Tashkin DP, Wang H, Halpin D, Kleerup EC, Connett J, Li N, et al. Comparison of the variability of the annual rates of change in FEV1 determined from serial measurements of the pre-versus post-bronchodilator FEV1 over 5 years in mild to moderate COPD: results of the lung health study. *Respir Res*. (2012) 13:70. doi: 10.1186/1465-9921-13-70

24. Mannino DM, Diaz-Guzman E, Buist S. Pre- and post-bronchodilator lung function as predictors of mortality in the lung health study. *Respir Res*. (2011) 12:136. doi: 10.1186/1465-9921-12-136

25. ImageNet. Available at: http://www.image-net.org (Accessed October 27, 2023).

26. MathWorks. Pretrained deep neural networks. Available at: https://mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html (Accessed 27, October 2023).

27. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. (1986) 1:307–10. doi: 10.1016/j.ijnurstu.2009.10.001

28. Yuan R, Hogg JC, Pare PD, Sin DD, Wong JC, Nakano Y, et al. Prediction of the rate of decline in FEV_1 in smokers using quantitative computed tomography. *Thorax*. (2009) 64:944–9. doi: 10.1136/thx.2008.112433

29. Mohamed Hoesein FA, de Hoop B, Zanen P, Gietema H, Kruitwagen CL, van Ginneken B, et al. CT-quantified emphysema in male heavy smokers: association with lung function decline. *Thorax*. (2011) 66:782–7. doi: 10.1136/thx.2010.145995

30. Vestbo J, Edwards LD, Scanlon PD, Yates JC, Agusti A, Bakke P, et al. Changes in forced expiratory volume in 1 second over time in COPD. *N Engl J Med*. (2011) 365:1184–92. doi: 10.1056/NEJMoa1105482

31. Han MK, Agusti A, Celli BR, Criner GJ, Halpin DMG, Roche N, et al. From GOLD 0 to pre-COPD. *Am J Respir Crit Care Med*. (2021) 203:414–23. doi: 10.1164/rccm.202008-3328PP

32. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet*. (2018) 392:1789–858. doi: 10.1016/S0140-6736(18)32279-7

Functional gas exchange measures on ¹²⁹Xe MRI and spectroscopy are associated with age, sex, and BMI in healthy subjects

David Mummy[1], Shuo Zhang[1], Aryil Bechtel[1], Junlan Lu[2],
Joseph Mammarappallil[1], Suphachart Leewiwatwong[3],
Anna Costelle[2], Aparna Swaminathan[4] and
Bastiaan Driehuys[1,2,3]*

[1]Department of Radiology, Duke University, Durham, NC, United States, [2]Medical Physics Graduate
Program, Duke University, Durham, NC, United States, [3]Department of Biomedical Engineering, Duke
University, Durham, NC, United States, [4]Department of Medicine, Duke University, Durham, NC,
United States

Introduction: Hyperpolarized ¹²⁹Xe MRI and spectroscopy is a rapidly growing technique for assessing lung function, with applications in a wide range of obstructive, restrictive, and pulmonary vascular disease. However, normal variations in ¹²⁹Xe measures of gas exchange across healthy subjects are not well characterized, presenting an obstacle to differentiating disease processes from the consequences of expected physiological heterogeneity. Here, we use multivariate models to evaluate the role of age, sex, and BMI in a range of commonly used ¹²⁹Xe measures of gas exchange.

Materials and methods: Healthy subjects ($N = 40$, 16F, age 44.3 ± 17.8 yrs., min-max $22-87$ years) with no history of cardiopulmonary disease underwent ¹²⁹Xe gas exchange MRI and spectroscopy. We used multivariate linear models to assess the associations of age, sex, and body mass index (BMI) with the RBC:Membrane (RBC:M), membrane to gas (Mem:Gas), and red blood cell to gas (RBC:Gas) ratios, as well as measurements of RBC oscillation amplitude and RBC chemical shift.

Results: Age, sex, and BMI were all significant covariates in the RBC:M model. Each additional 10 years of age was associated with a 0.05 decrease in RBC:M ($p < 0.001$), each additional 10 points of BMI was associated with a decrease of 0.07 ($p = 0.02$), and males were associated with a 0.17 higher RBC:M than females ($p < 0.001$). For Mem:Gas, male sex was associated with a decrease and BMI was associated with an increase. For RBC:Gas, age was associated with a decrease and male sex was associated with an increase. RBC oscillation amplitude increased with age and RBC chemical shift was not associated with any of the three covariates.

Discussion: ¹²⁹Xe MRI and spectroscopy measurements in healthy subjects, particularly the widely used RBC:M measurement, exhibit heterogeneity associated in part with variations in subject age, sex, and BMI. Elucidating the contributions of these and other factors to ¹²⁹Xe gas exchange measurements is a critical component for differentiating disease processes from expected variation in healthy subjects. Notably, the Mem:Gas and RBC chemical shift

appear to be stable with aging, suggesting that unexplained deviations in these metrics may be signs of underlying abnormalities.

1 Introduction

[129]Xe gas exchange MRI and spectroscopy are emerging as techniques with clinical potential for evaluating disease status, progression, and therapy response in a range of cardiopulmonary diseases, including asthma (1–3), COPD (4, 5), COVID-19 (6, 7), interstitial lung diseases (8–11), and pulmonary hypertension (12–14). Inhaled [129]Xe gas freely diffuses from the lung airspaces, through the alveolar-capillary interstitial membrane, and into the capillary bloodstream, where it transiently binds with red blood cells (RBCs). Three-dimensional images of [129]Xe signal within each of these separate compartments (ventilation, interstitial membrane, and RBC) can then be individually reconstructed, enabling each of their contributions to overall gas exchange function to be quantified (15). The addition of whole-lung static and time-resolved magnetic resonance spectroscopy (MRS) provides insights into blood oxygenation and hemodynamics (13).

These imaging and spectroscopic techniques yield a range of measurements that reflect various aspects of cardiopulmonary function. On imaging, measurements of membrane and RBC signals are normalized to the spatially corresponding gas signal on a voxel-wise basis (termed the Mem:Gas and RBC:Gas ratios respectively). The resulting 3D gas, membrane, and RBC images are often quantified by the percentage of reduced or absent signal relative to a previously established reference population (16), and elevated membrane signal assessed using this method has been observed as a feature of interstitial lung diseases (9, 10). On spectroscopy, the widely used metric of the RBC to membrane signal ratio (RBC:M) reflects overall gas exchange efficiency (17). Additional emerging metrics of interest on spectroscopy are the RBC chemical shift, which may reflect the degree of delayed or impaired capillary blood oxygenation (18), and the amplitude of the dynamic RBC signal, which oscillates with the cardiac cycle (19) and is known to vary in a range of cardiopulmonary diseases (13).

The maturation of the field of [129]Xe MRI/MRS and its ongoing transition to clinical use requires the development of suitable reference standards to (1) understand and interpret the findings for an individual patient, (2) understand the expected variation of these metrics in real-world populations, and thereby (3) differentiate potentially pathological findings from the expected heterogeneity across healthy subjects. Common [129]Xe metrics like those cited above, which are used in research studies of patient populations, are often compared with samples of healthy subjects that are already on hand. However, those subjects may or may not be age-, sex-, or otherwise adequately matched with the population of interest. This raises the question of whether observed differences are due to disease or simply due to differences in basic subject characteristics. Specifically,

variations in [129]Xe gas exchange function in healthy subjects due to age, sex, and BMI are not well understood.

Recent work by Plummer et al. has sought to address this gap by modeling age-related changes in [129]Xe image intensity histograms, with the goal of creating a tailored reference histogram for a given set of patient attributes (20). Here, we take an alternative approach (21) akin to that adopted by Collier et al. (22) and use multivariate linear models to evaluate age-related changes in [129]Xe MRI and spectroscopy, including the RBC:M, RBC:Gas, Mem:Gas, RBC chemical shift, and RBC amplitude oscillations. Finally, we present a prototype of a "percent predicted" model for the RBC:M as an example approach for creating reference tables analogous to those used in conventional clinical pulmonary function tests (PFTs).

2 Materials and methods

2.1 Participants

Healthy subjects with no history of pulmonary disease, less than 5 pack-years smoking history, and no smoking history within the last 5 years were compiled from existing HIPAA-compliant [129]Xe MRI studies at our institution (Pro00025110, Pro00059856, Pro00060259, Pro00106763, and Pro00107570). All participants underwent an informed consent process prior to recruitment and written consent at imaging. Hyperpolarized [129]Xe was produced and administered under an investigational new drug approval (109,490). It was recently granted approval by the U.S. Food and Drug Administration for a ventilation indication (23).

2.2 [129]Xe hyperpolarization and imaging acquisition

Hyperpolarized [129]Xe was produced by continuous-flow spin-exchange optical pumping and cryogenic accumulation with commercially available hyperpolarizer systems (Models 9820 and 9810, Polarean) and dispensed into a Tedlar dose delivery bag (Jensen Inert Products). [129]Xe MRI and spectroscopy were performed using a quadrature transmit-receive flexible vest coil (Clinical MR Solutions) on a 3 T scanner (Siemens, Trio/Prisma) during two separate breath holds. Participants were instructed to breathe out normally prior to inhaling the [129]Xe dose, with the aim of beginning the inhale at approximately functional residual capacity (FRC). For calibration and spectroscopy, participants received a [129]Xe dose equivalent (DE) of 81.9 ± 21.6 mL (mean $\pm$ sd), where the DE is the hypothetical volume of pure 100% hyperpolarized [129]Xe that would produce the same net magnetization as the dose being sampled; this required an overall

xenon volume of 420.6 ± 90.0 mL. For imaging, a larger DE of 177.6 ± 29.1 mL was used, corresponding to a xenon volume of 674.2 ± 76.3 mL. For all doses, a nonpolarized ^{129}Xe blend was added to increase the total delivery bag volume to approximately 20% of the subject's forced vital capacity (FVC) per the Xenon Consortium protocol (24), or an untailored volume of 1 L (approximately one-third of the study population), with overall bag volumes of 943 ± 165 mL (min 500 mL, max 1,250 mL). Throughout imaging, heart rate and oxygen saturation were continuously monitored (Model 7500, Nonin).

Calibration and spectroscopy were performed during the first 10-s breath hold by acquiring ~500 ^{129}Xe free-induction decays at 15-msec intervals (echo time, 0.45 msec; target flip angle, 20°; dwell time, 19 μsec; 512 points) (24). During the second 15-s breath hold, three-dimensional images were acquired using an interleaved radial acquisition of gas- and dissolved-phase (i.e., membrane uptake and RBC transfer) signals. Dissolved-phase images used a repetition time of 15 msec and a 20° flip angle [equivalent to an effective repetition time of 249 ms with a 90° flip angle (25)], matching the calibration scan, and an echo time of 0.45–0.50 ms chosen to target a 90° phase angle separation between the signals from the two compartments. The dissolved phase was excited using an RF pulse that was offset 218 ppm from the gas-phase peak.

2.3 Spectroscopy processing and analysis

Free-induction decays were acquired and fit to three peaks (gas, membrane uptake, and RBC transfer) in the time domain, with each peak described by amplitude, chemical shift, phase, and either one (gas and RBC) or two (membrane uptake) full widths at half maximum (19). The use of the two full widths for the membrane signal is due to its characterization as a Voigt line shape rather than the Lorentzian shape of the other two compartments. Static spectroscopy indices included the RBC:M and RBC chemical shift. In addition, the dynamic (time-resolved) spectroscopy measurement of amplitude of the RBC signal oscillations ("RBC amplitude oscillations") was analyzed. These oscillations are driven by the cardiac cycle and were quantified by their peak-to-peak value relative to the mean (26).

2.4 Image reconstruction and analysis

The two dissolved-phase compartments were decomposed using the one-point Dixon method (15), using knowledge of the participant-specific RBC:M derived from spectroscopy. Three-dimensional images of the membrane and RBC compartments were individually reconstructed with a nominal isotropic resolution of 6.3 mm, and then divided by the gas-phase intensities from the same acquisition on a voxel-by-voxel basis to create maps relative to the gas signal. The resulting images were quantified using previously established methods (16). Only the ventilation image was corrected for coil transmit inhomogeneity using the N4ITK software tool (27). This correction was not needed for the Mem:Gas and RBC:Gas images since for such ratio maps, first-order coil inhomogeneity effects cancel out. The Mem:Gas and RBC:Gas images were characterized by calculating the mean of the individual voxel-wise ratios.

2.5 Quality control

We applied a qualitative rubric based on visual inspection of the images in conjunction with SNR, the presence of artifacts, adequate field of view, and so on. This quality control process was developed for a recent clinical trial (28) and a manuscript is in preparation describing our methods in detail.

2.6 Statistical analysis

Correlations between age and each individual metric were assessed using Spearman's correlation. The RBC:M, Mem:Gas, and RBC:Gas ratios, as well as the RBC chemical shift and RBC amplitude oscillations, were each assessed individually using multivariate linear models, with age, sex, and BMI as covariates.

Finally, we created a prototype of a "percent predicted" model of RBC:Mem by using the parameter estimates from the corresponding multivariate linear model to predict the expected value for each individual patient based on their age, sex, and BMI and then dividing the observed value by this predicted value to yield the "RBC:M percent predicted" value (RBC:M%), analogous to the percent predicted values employed in PFTs. The resulting distribution of percent predicted values was assessed for normality with the Shapiro–Wilk test.

All statistical calculations were performed using R version 4.2.1.

3 Results

We recruited 40 total subjects (16F, age 44.3 ± 17.8 yrs., min-max 22–87 years) imaged between June 2018 and June 2022. Demographic, ^{129}Xe gas exchange, and clinical data for this population are shown in Table 1. Subsets of these healthy subjects have also been described in previous studies (9, 29, 30).

Representative grayscale images from subjects in the top fifth percentile of RBC:M, at the median value, and in the bottom fifth percentile are shown in Figure 1. Note that the age of the subjects

TABLE 1 Demographic, pulmonary function test, and xenon MRI results for the study population, divided into three age groups.

	Age group		
	<30 (N = 13)	30–50 (N = 10)	>50 (N = 17)
Sex (F)	6 (46%)	6 (60%)	4 (24%)
Age	25.7 ± 2.6	37.8 ± 6.5	62.2 ± 9.6
BMI	25.4 ± 3.1	30.4 ± 5.9	25.0 ± 3.2
FVC%	102.6 ± 9.1	98.0 ± 20.3	109.4 ± 15.8
FEV1%	98.2 ± 12.8	99.4 ± 21.2	107.3 ± 16.2
FEV1/FVC%	98.0 ± 9.5	101.0 ± 8.9	98.7 ± 8.2
Mem:Gas × 100	0.71 [0.58, 0.83]	0.83 [0.75, 0.89]	0.66 [0.59, 0.68]
RBC:Gas × 100	0.45 [0.33, 0.46]	0.46 [0.33, 0.49]	0.32 [0.28, 0.38]
RBC:M	0.61 [0.51, 0.64]	0.49 [0.44, 0.55]	0.48 [0.44, 0.53]
RBC shift (ppm)	218.2 [218.1, 218.6]	218.2 [217.8, 218.3]	218.3 [218.2, 218.6]
RBC amplitude oscillations (%)	10.7 [9.8, 11.7]	10.4 [9.4, 11.7]	11.6 [9.9, 13.2]

Results are either mean ± standard deviation or median [first quartile, third quartile]. FEV1, forced expiratory volume in 1 s percent predicted; FVC%, forced vital capacity percent predicted; RBC:M, red blood cell to membrane ratio.
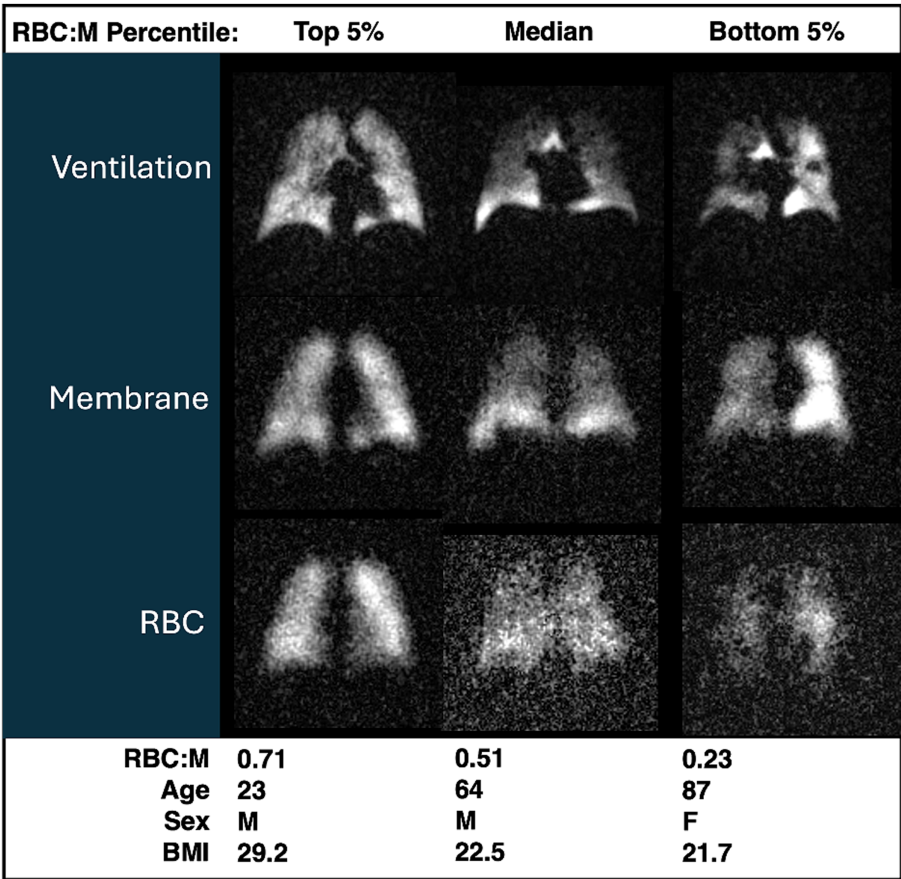
FIGURE 1
Grayscale images of ventilation, membrane, RBC from three participants with RBC:M values in the top fifth percentile (left), at the median (center), and in the bottom fifth percentile (right). Age, sex, and BMI are presented for each subject. Note that with increasing age, both RBC:M and the RBC signal decrease, while the membrane signal is relatively preserved. RBC, red blood cell; RBC:M, RBC to membrane ratio.

increases with decreasing RBC:M, and the membrane signal is relatively preserved while the RBC signal shows a marked decrease with decreasing RBC:M. Representative samples of color-binned gas exchange imaging from three age groups are shown in Figure 2. One subject exhibits completely unremarkable findings; another subject exhibits modestly increased Mem:Gas, and another subject (age 68) shows marked RBC defects. On spectroscopy, while the RBC:M values exhibit a range of findings from normal to poor in these three subjects, the RBC amplitude oscillations are overall normal, and the RBC chemical shift shows minimal variation.

Overall, as shown in Figure 3, the RBC:M ratio was inversely correlated with age (Spearman's $\rho = -0.40$, $p = 0.01$), as was the RBC:Gas ($\rho = -0.43$, $p = 0.006$), while the Mem:Gas was not correlated. As shown in Figure 4, the RBC chemical shift was not correlated with age, but RBC amplitude oscillations were correlated ($\rho = 0.33$, $p = 0.03$).

Results from all six of the multivariate models are shown in Table 2. In the model of RBC:M, all three parameters of age, sex, and BMI were significant. Specifically, each additional 10 years of age was associated with a 0.05 decrease in RBC:M ($p < 0.001$), male sex was associated with an increase of 0.17 ($p < 0.001$) vs. female, and each additional 10 points of BMI was associated with a decrease of 0.07 ($p = 0.02$). In the analogous model of Mem:Gas, males were associated with a reduction of 0.1×10^{-2} relative to females ($p = 0.02$), and every

10 points of BMI was associated with an increase of 0.1×10^{-2} ($p = 0.03$); Mem:Gas was not significantly affected by age. In the model of RBC:Gas, every 10 years of age was associated with a decrease of 0.03×10^{-2} ($p < 0.001$), and it was 0.07×10^{-2} higher for males than females ($p = 0.02$), while BMI was not a significant parameter. In the model of RBC amplitude oscillations, age was the only significant parameter ($p < 0.01$), with every 10 years of age associated with a 0.5% absolute increase in RBC amplitude oscillations. In the model of RBC chemical shift, none of the three parameters were significant.

The histogram of RBC:M percent predicted values based on the multivariate model of RBC:M as applied to our study population is shown in Figure 5, and the corresponding distribution of percent predicted values in this study population was not significantly different from normal ($p = 0.45$).

4 Discussion

This study suggests that age, sex, and BMI may all play roles of varying but critical significance in most measurements of ^{129}Xe gas exchange. Thus, interpretations of findings in individual patients should be informed by knowledge of the specific effects of these characteristics. Notably, none of the attributes we studied appeared
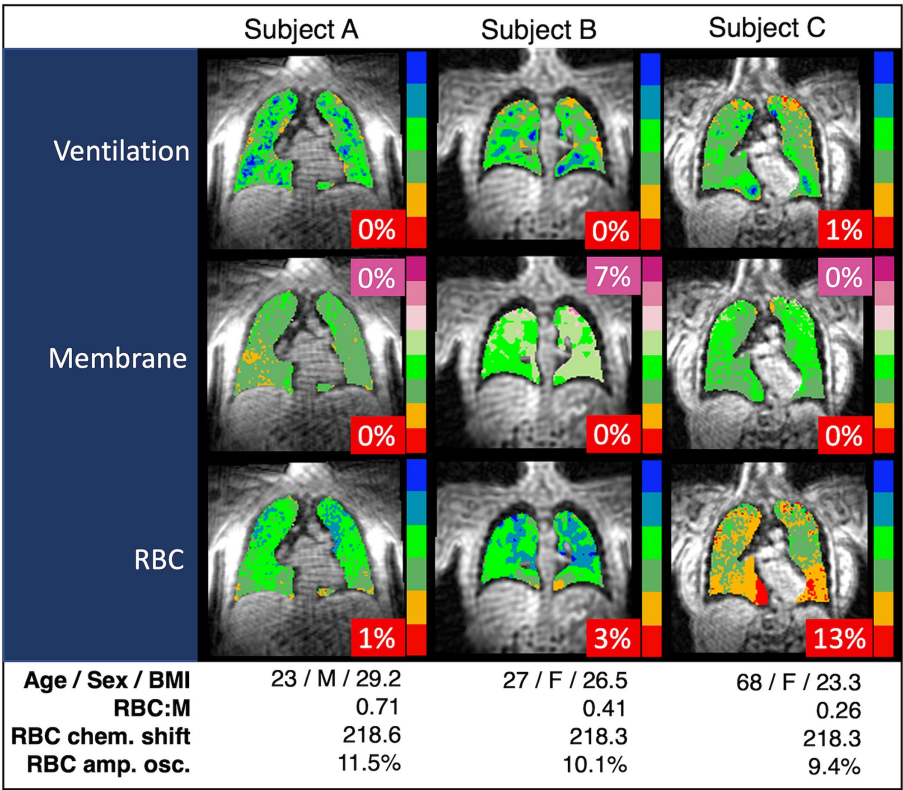
FIGURE 2

Spectroscopy measurements and color-binned ¹²⁹Xe gas exchange images from three representative subjects. Subject A appears to have exemplary lung function with no abnormal values. Subject B has the increased membrane signal (purple, 7%) exhibited in some of the subjects with higher BMI, but otherwise exhibits normal lung function. Subject C has the RBC defects (red, 13%) and decreased RBC:M associated with age. In our binning method, "defects" are determined based on a healthy reference cohort used as a standard comparison for all subjects, and then interpreted on an individual basis using a similar approach to that presented in this work. Note that chemical shift is consistent across all three subjects. RBC:M, red blood cell to membrane ratio.
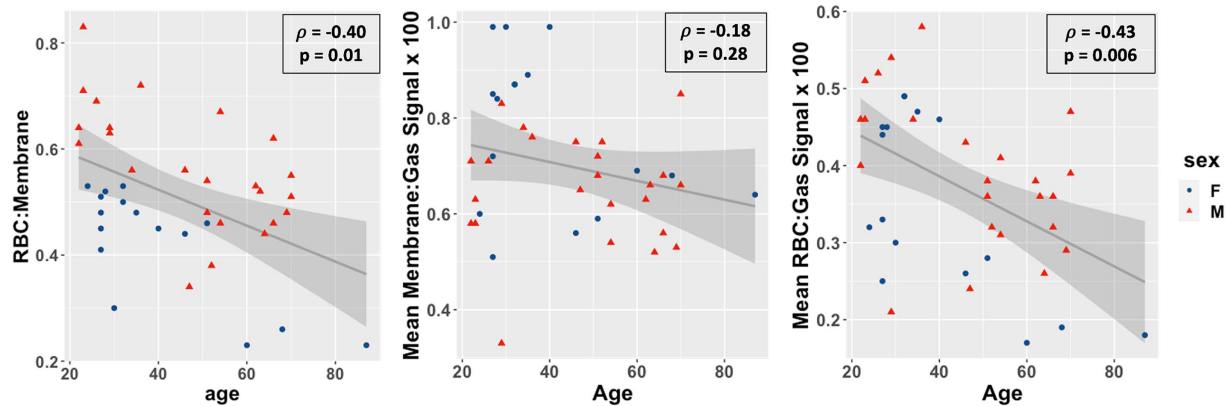
FIGURE 3

Scatterplots of the red blood cell to membrane ratio (RBC:M, left), mean membrane:gas signal (middle), and mean RBC:gas signal RBC amplitude oscillations (right) vs. age in our study population. Age was negatively correlated with RBC:M ($r = -0.46$, $p = 0.003$), not correlated with mean membrane signal ($r = -0.19$, $p = 0.2$), and negatively correlated with mean RBC signal ($r = -0.43$, $p = 0.005$). This result suggests that a decrease in RBC is the primary driver of the age-related decrease in RBC:M, rather than any meaningful changes in membrane signal. Note also the significant sex-related disparity in the RBC:M scatterplot, possibly driven in part by sex-related differences in hemoglobin.

to significantly affect RBC chemical shift, suggesting that deviations in this metric may be a consistent marker of an underlying disease process.

We observed a clear age-related decrease in the RBC:M ratio, as also observed by Plummer et al. (20) and strikingly similar to the preliminary results reported by Collier et al. (22), who estimated an

FIGURE 4
Scatterplots of the RBC chemical shift (left) and RBC amplitude oscillations (right) vs. age in our study population. Age was not correlated with RBC chemical shift ($\rho = 0.15$, $p = 0.36$) and positively correlated with RBC amplitude oscillations ($r = 0.33$, $p = 0.03$). RBC, red blood cell.
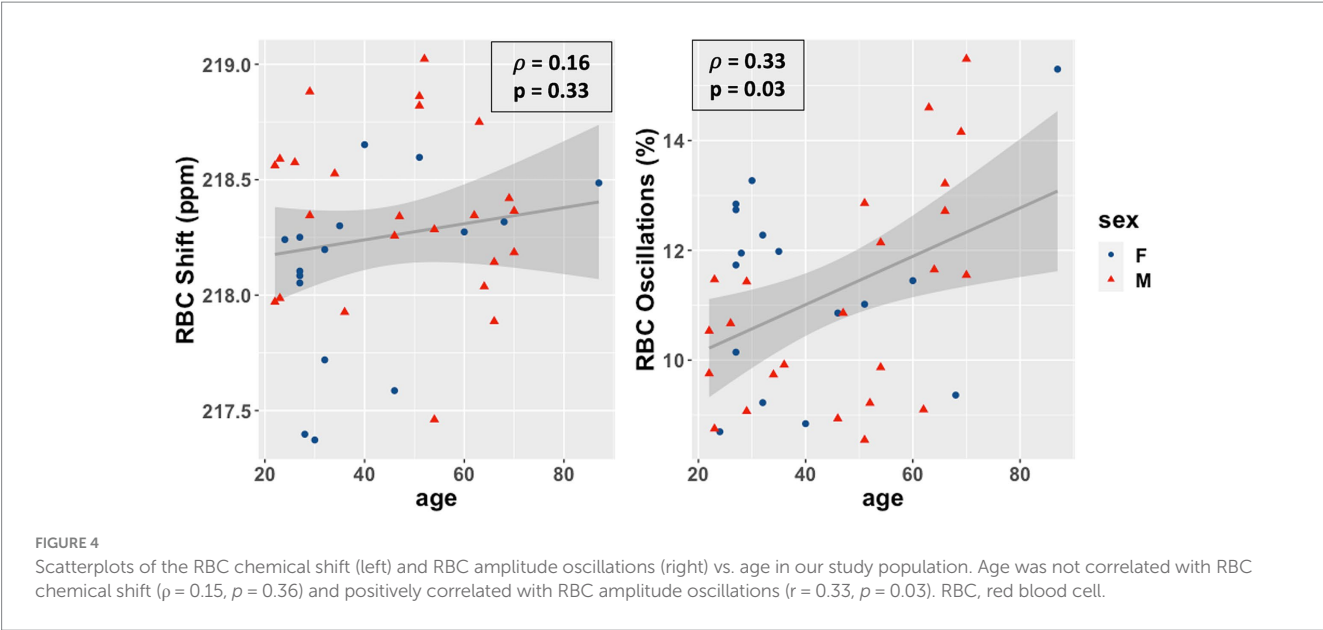
TABLE 2 Multivariate model parameters in linear models of RBC:M, Mem:Gas, RBC:Gas, RBC oscillations, and the RBC chemical shift, with parameters with $p < 0.05$ shown in bold with an asterisk, with the exception of the intercept (italics), which was significant in all models.

Outcome	Covariate	Estimate	*p*-value
RBC:M	*Intercept*	*0.79*	*<0.001*
	Age	**−0.005**	**<0.001***
	Male sex	**0.17**	**<0.001***
	BMI	**−0.007**	**0.02***
Mean RBC:Gas ×100	*Intercept*	*0.5*	*<0.001*
	Age	**−0.003**	**<0.001***
	Male sex	**0.07**	**0.02***
	BMI	−0.0008	0.8
Mean Mem:Gas × 100	*Intercept*	*0.52*	*<0.001*
	Age	−0.0009	0.45
	Male sex	**−0.1**	**0.02***
	BMI	**0.01**	**0.03***
RBC amplitude oscillations (%)	*Intercept*	*7.6*	*<0.001*
	Age	**0.05**	**0.004***
	Male sex	−0.59	0.3
	BMI	0.07	0.3
RBC chemical shift (ppm)	*Intercept*	*218.4*	*<0.001*
	Age	0.002	0.7
	Male sex	0.24	0.07
	BMI	−0.014	0.3

Note the association of BMI with RBC:M and Mem:Gas, and the absence of any significant association between RBC chemical shift and age, sex, or BMI. RBC:M, red blood cell to membrane ratio.

FIGURE 5
Histogram of RBC:M percent predicted (RBC:M%) values in our population with superposed curve fit. The distribution is not significantly different from normal ($p = 0.45$). Predicted values were determined by plugging age, sex, and BMI for each subject into the multivariate linear model in Table 2. Observed values for each subject were then divided by the predicted value to yield the %Pred value. The fifth percentile of the histogram is at an RBC:M% value of 78%, relatively close to the 80% percent predicted value generally used as a cutoff for "normal" in pulmonary function testing. RBC:M, red blood cell to membrane ratio.

age coefficient of −0.004/yr. vs. our finding of −0.005/yr. As the name implies, the RBC:M depends on both the RBC and membrane signals, which obscures the contribution of these individual compartments; a decreased RBC:M could result from changes in the relative proportion of the two signals whether or not they individually increase or decrease. In this population, however, the stability of the membrane signal with age, coupled with the decrease in mean RBC signal, suggests that, rather than membrane diffusion limitation, vascular changes are the primary driver of age-related decreases in RBC:M. Both Plummer and Collier also observed an age-related decrease in Mem:Gas and RBC:Gas in adults, but differing methodologies across studies make these results more difficult to compare directly. Particularly, Plummer et al. included pediatric subjects in their study, and used a more complex curve-fitting
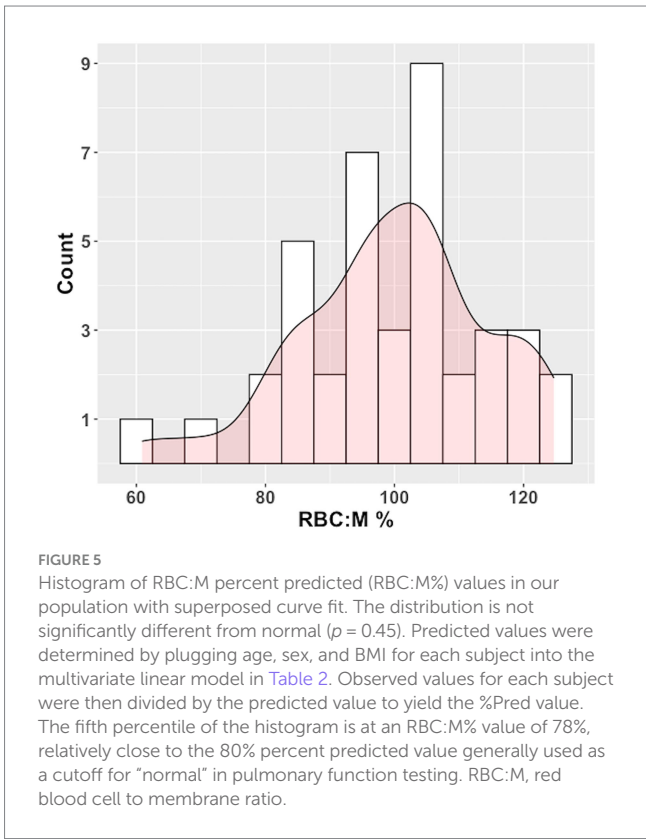
technique to account for the non-linear behavior from child to adulthood.

The increased RBC amplitude oscillations with age may be caused in part by reduced pulmonary vascular volume, a known consequence of normal aging (31). This type of characterization may be useful to incorporate into models of pulmonary hypertension (19) or other studies that use xenon hemodynamics as an indicator of cardiovascular abnormalities. Ongoing work in our lab is advancing this work to understand the relationship between these measurements and cardiac output, and to derive fundamental cardiovascular parameters such as pulmonary vascular resistance from dynamic spectroscopy. It is also possible to derive 3D images of these temporal dynamics in the RBC signal using advanced reconstruction techniques, further refining the ability to characterize the spatial characteristics of dynamic cardiopulmonary function (12, 32).

The stability of the RBC chemical shift measurement with age, sex, and BMI is also noteworthy. This measurement is believed to reflect the efficiency with which capillary blood becomes fully oxygen-saturated as it passes through the gas exchanging region of the alveolus (9, 18, 19). This stability is somewhat non-intuitive, given that RBC defects increase with age, and prior work has shown in patients that such defects are associated with a reduced chemical shift (9, 19). This lower RBC shift has generally been attributed to reduced transit time through the gas exchange regions. Given that our work has shown RBC transfer defects to increase with aging, while RBC shift remains stable, this would suggest that cardiac output must decrease commensurately to the loss of capillary blood volume. While further work is needed to understand this behavior, the cardiac index is indeed known to decrease with age (33). Nonetheless, the fact that neither age, sex, nor BMI was significantly associated with RBC chemical shift changes would suggest that abnormal findings in this measurement may be an indicator of an underlying disease process independent of any these attributes.

The root cause of the association between BMI and membrane signal (both on its own in Mem:Gas and ostensibly in RBC:M, since we did not observe an association between BMI and the RBC signal) is not yet clear. However, a recent study in mice has shown that obesity is associated with a thickened interstitium and thus a thickened air-blood barrier (34). It is also possible that the increased membrane signal is due in part to a purely physical increased density of the lung due to the pressure of additional body weight, especially when the subject is in a supine position and tightly wrapped in a vest coil. Obesity has been associated with reduced FRC (35), and lung inflation volume alone is also known to affect ^{129}Xe gas exchange measurements (36). Further work in this area is needed to isolate and test the effect of BMI on membrane signal separate from other potential factors such as lung inflation volume.

In our prototype of a "percent predicted" model of RBC:M, intended to be analogous to the familiar percent predicted tables used in pulmonary function testing, the percent predicted values were normally distributed, with a 5th percentile value of 78%, very close to the 80% threshold generally used as a lower limit of normal for PFTs (although caution must be used when employing this type of fixed threshold to avoid over- or under-diagnosis) (37). This suggests the feasibility of this approach for creating age-normalized reference values to isolate true gas exchange abnormalities from the expected effects of aging both in individual subjects and across disease cohorts.

We acknowledge limitations to this work, particularly the need for a larger population evenly distributed across a wide range of age, sex, BMI, and other demographic characteristics to truly develop a reference cohort and expand our model to include additional variables and possible interaction terms. Future studies should also correct for hemoglobin, as this is known to have an effect on ^{129}Xe gas exchange measurements and thus may be a confounder driving the apparent association with sex and RBC:M in our model due to the well-established sex-based differences in hemoglobin (29). However, hemoglobin values were only collected for a subset of our study population, thus limiting our ability to implement that correction in this work. In addition, our spectroscopic measurements may not be directly comparable with those of other groups, given that, rather than conventional phased Lorentzian fitting, we employ Voigt fitting of the membrane signal in the time domain to better represent the shape of the peak. We also note that despite upgrading our scanner during the study, we did not observe a discernable difference in SNR or signal quantification. We further note that different flip angles are applied to each compartment, and the effect of this relative flip angle difference in the context of constant replenishment of the dissolved-phase ^{129}Xe signal is not well understood. However, in our experience, these differences are dwarfed by physiological and/or pathological variation and would require improved simulation, modeling, and experimentation to identify and tease out. Finally, we note that ^{129}Xe MRI is currently a specialized and expensive technique, although with ongoing wider adoption, costs are expected to come down as economies of scale become possible. These limitations notwithstanding, this work represents a necessary first step toward establishing the feasibility of this approach to understanding age-related changes of widely-used xenon MRI gas exchange metrics and the development of generalized reference values.

In conclusion, hyperpolarized Xenon MRI reveals associations between age, sex, and BMI, particularly in the RBC:M ratio but also to varying degrees in the Mem:Gas, RBC:Gas, and RBC hemodynamic oscillations. These results suggest a clear need to incorporate the effects of these attributes when interpreting these metrics in individual patients or comparing across patients or populations. In contrast, none of these attributes appeared to affect the RBC chemical shift, and thus variations in this metric from the healthy reference may be a reliable marker of abnormal function. Further work in this area will determine the role of potential model covariates in a larger study population – particularly the potential role of BMI in the membrane signal – and refine our reference values and percent predicted models, with a goal of ultimately advancing clinical assessment and individual patient management using ^{129}Xe MRI.

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 humans were approved by Duke University Health System Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements.

The participants provided their written informed consent to participate in this study.

Author contributions

DM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. SZ: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – review & editing. AB: Data curation, Formal analysis, Software, Writing – review & editing. JL: Formal analysis, Software, Writing – review & editing. JM: Funding acquisition, Project administration, Supervision, Writing – review & editing. SL: Formal analysis, Software, Writing – review & editing. AC: Formal analysis, Methodology, Software, Writing – review & editing. AS: Conceptualization, Investigation, Methodology, Writing – review & editing. BD: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

BD is founder and chief technology officer for Polarean Imaging. DM is a consultant for Polarean Imaging. Polarean Imaging manufactures the hyperpolarization equipment used in this study but had no involvement whatsoever in the development of this study or writing of the manuscript.

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.

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.

References

1. Svenningsen S, Kjarsgaard M, Haider E, Venegas C, Konyer N, Friedlander Y, et al. Effects of Dupilumab on mucus plugging and ventilation defects in patients with moderate-to-severe asthma: a randomized, double-blind, placebo-controlled trial. *Am J Respir Crit Care Med*. (2023) 208:995–7. doi: 10.1164/rccm.202306-1102LE

2. McIntosh MJ, Kooner HK, Eddy RL, Jeimy S, Licskai C, Mackenzie CA, et al. Asthma control, airway mucus, and 129Xe MRI ventilation after a single benralizumab dose. *Chest*. (2022) 162:520–33. doi: 10.1016/j.chest.2022.03.003

3. Hall CS, Quirk JD, Goss CW, Lew D, Kozlowski J, Thomen RP, et al. Single-session bronchial Thermoplasty guided by 129Xe magnetic resonance imaging. A pilot randomized controlled clinical trial. *Am J Respir Crit Care Med*. (2020) 202:524–34. doi: 10.1164/rccm.201905-1021OC

4. Mummy DG, Coleman EM, Wang Z, Bier EA, Lu J, Driehuys B, et al. Regional gas exchange measured by 129Xe MRI before and after combination bronchodilators treatment in chronic obstructive pulmonary disease. *J Magn Reson Imaging*. (2021) 54:964–74. doi: 10.1002/jmri.27662

5. Myc L, Qing K, He M, Tustison N, Lin Z, Manichaikul AW, et al. Characterisation of gas exchange in COPD with dissolved-phase hyperpolarised xenon-129 MRI. *Thorax*. (2021) 76:178–81. doi: 10.1136/thoraxjnl-2020-214924

6. Matheson AM, McIntosh MJ, Kooner HK, Lee J, Desaigoudar V, Bier E, et al. Persistent 129Xe MRI pulmonary and CT vascular abnormalities in symptomatic individuals with post-acute COVID-19 syndrome. *Radiology*. (2022) 305:466–76. doi: 10.1148/radiol.220492

7. Grist JT, Collier GJ, Walters H, Kim M, Chen M, Abu Eid G, et al. Lung abnormalities detected with hyperpolarized 129Xe MRI in patients with long COVID. *Radiology*. (2022) 305:709–17. doi: 10.1148/radiol.220069

8. Hahn AD, Carey KJ, Barton GP, Torres LA, Kammerman J, Cadman RV, et al. Functional xenon-129 magnetic resonance imaging response to antifibrotic treatment in idiopathic pulmonary fibrosis. *ERJ Open Res*. (2023) 9:00080–2023. doi: 10.1183/23120541.00080-2023

9. Mummy DG, Bier EA, Wang Z, Korzekwinski J, Morrison L, Barkauskas C, et al. Hyperpolarized 129Xe MRI and spectroscopy of gas-exchange abnormalities in nonspecific interstitial pneumonia. *Radiology*. (2021) 301:211–20. doi: 10.1148/radiol.2021204149

10. Wang JM, Robertson SH, Wang Z, He M, Virgincar RS, Schrank GM, et al. Using hyperpolarized 129Xe MRI to quantify regional gas transfer in idiopathic pulmonary fibrosis. *Thorax*. (2018) 73:21–8. doi: 10.1136/thoraxjnl-2017-210070

11. Rankine LJ, Wang Z, Wang JM, He M, McAdams HP, Mammarappallil J, et al. 129Xenon gas exchange magnetic resonance imaging as a potential prognostic marker for progression of idiopathic pulmonary fibrosis. *Ann Am Thorac Soc*. (2020) 17:121–5. doi: 10.1513/AnnalsATS.201905-413RL

12. Niedbalski PJ, Bier EA, Wang Z, Willmering MM, Driehuys B, Cleveland ZI. Mapping cardiopulmonary dynamics within the microvasculature of the lungs using dissolved 129Xe MRI. *J Appl Physiol*. (2020) 129:218–29. doi: 10.1152/japplphysiol.00186.2020

13. Wang Z, Bier EA, Swaminathan A, Parikh K, Nouls J, He M, et al. Diverse cardiopulmonary diseases are associated with distinct xenon magnetic resonance imaging signatures. *Eur Respir J*. (2019) 54:1900831. doi: 10.1183/13993003.00831-2019

14. Bier EA, Alenezi F, Lu J, Wang Z, Mammarappallil JG, O'Sullivan-Murphy B, et al. Noninvasive diagnosis of pulmonary hypertension with hyperpolarised 129Xe magnetic resonance imaging and spectroscopy. *ERJ Open Res*. (2022) 8:00035–2022. doi: 10.1183/23120541.00035-2022

15. Kaushik SS, Robertson SH, Freeman MS, He M, Kelly KT, Roos JE, et al. Single-breath clinical imaging of hyperpolarized 129xe in the airspaces, barrier, and red blood cells using an interleaved 3D radial 1-point Dixon acquisition. *Magn Reson Med*. (2016) 75:1434–43. doi: 10.1002/mrm.25675

16. Wang Z, Robertson SH, Wang J, He M, Virgincar RS, Schrank GM, et al. Quantitative analysis of hyperpolarized 129Xe gas transfer MRI. *Med Phys*. (2017) 44:2415–28. doi: 10.1002/mp.12264

17. Weatherley ND, Stewart NJ, Chan H-F, Austin M, Smith LJ, Collier G, et al. Hyperpolarised xenon magnetic resonance spectroscopy for the longitudinal assessment of changes in gas diffusion in IPF. *Thorax*. (2019) 74:500–2. doi: 10.1136/thoraxjnl-2018-211851

18. Norquay G, Leung G, Stewart NJ, Wolber J, Wild JM. 129Xe chemical shift in human blood and pulmonary blood oxygenation measurement in humans using hyperpolarized 129Xe NMR. *Magn Reson Med*. (2017) 77:1399–408. doi: 10.1002/mrm.26225

19. Bier EA, Robertson SH, Schrank GM, Rackley C, Mammarappallil JG, Rajagopal S, et al. A protocol for quantifying cardiogenic oscillations in dynamic 129Xe gas exchange spectroscopy: the effects of idiopathic pulmonary fibrosis. *NMR Biomed*. (2019) 32:e4029. doi: 10.1002/nbm.4029

20. Plummer JW, Willmering MM, Cleveland ZI, Towe C, Woods JC, Walkup LL. Childhood to adulthood: accounting for age dependence in healthy-reference distributions in 129Xe gas-exchange MRI. *Magn Reson Med*. (2023) 89:1117–33. doi: 10.1002/mrm.29501

21. Mummy D, Swaminathan A, Bier E, Yarnall K, Bechtel A, Lu J, et al. Hyperpolarized 129Xe MRI and spectroscopy in healthy control subjects reveals age-related changes in measurements of pulmonary gas exchange. In: *Proceedings of the Joint Annual Meeting ISMRM-ESMRMB 2022 and ISMRT Annual Meeting, London, England, UK*. (2022).

22. Collier GJ, Chan H-F, Stewart NJ, Norquay G, Smith LJ, Rao M, et al. Age and lung volume dependence of dissolved xenon-129 imaging parameters. In: *Proceedings of the Joint Annual Meeting ISMRM-ESMRMB 2022 and ISMRT Annual Meeting, London, England, UK*. (2022).

23. FDA. XENOVIEW Prescribing Information. (2022). Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/214375s000lbl.pdf

24. Niedbalski PJ, Hall CS, Castro M, Eddy RL, Rayment JH, Svenningsen S, et al. Protocols for multi-site trials using hyperpolarized 129Xe MRI for imaging of ventilation, alveolar-airspace size, and gas exchange: a position paper from the 129Xe MRI clinical trials consortium. *Magn Reson Med*. (2021) 86:2966–86. doi: 10.1002/mrm.28985

25. Ruppert K, Amzajerdian F, Hamedani H, Xin Y, Loza L, Achekzai T, et al. Assessment of flip angle–TR equivalence for standardized dissolved-phase imaging of the lung with hyperpolarized 129Xe MRI. *Magn Reson Med*. (2019) 81:1784–94. doi: 10.1002/mrm.27538

26. Costelle A, Lu J, Bechtel A, Bier E, Kabir S, Mammarappallil J, et al., Mummy D DB. Quantifying cardiogenic oscillations of hyperpolarized 129Xe gas exchange MR spectra in a healthy reference cohort. In: *International Society for Magnetic Resonance in Medicine International Conference*. (2023).

27. 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

28. Bell LC, Coimbra AF, Zou Y, Dimond C, Mata J, Muglar J, et al. Multisite implementation of 129Xe-MRI assessment of COPD disease progression: XeCITE early results. *Eur Respir Soc*. (2023) 62:PA2283. doi: 10.1183/13993003.congress-2023.PA2283

29. Bechtel A, Lu J, Mummy D, Bier E, Leewiwatwong S, Mugler J III, et al. Establishing a hemoglobin adjustment for 129Xe gas exchange MRI and MRS. *Magn Reson Med*. (2023) 90:1555–68. doi: 10.1002/mrm.29712

30. Lu J, Wang Z, Bier E, Leewiwatwong S, Mummy D, Driehuys B. Bias field correction in hyperpolarized 129Xe ventilation MRI using templates derived by RF-depolarization mapping. *Magn Reson Med*. (2022) 88:802–16. doi: 10.1002/mrm.29254

31. Hermann EA, Motahari A, Hoffman EA, Allen N, Bertoni AG, Bluemke DA, et al. Pulmonary blood volume among older adults in the community: the MESA lung study. *Circ Cardiovasc Imaging*. (2022) 15:e014380. doi: 10.1161/CIRCIMAGING.122.014380

32. Lu J, Alenezi F, Bier E, Leewiwatwong S, Mummy D, Kabir S, et al. Optimized quantitative mapping of cardiopulmonary oscillations using hyperpolarized 129Xe gas exchange MRI: digital phantoms and clinical evaluation in CTEPH. *Magn Reson Med*. (2024) 91:1541–55. doi: 10.1002/mrm.29965

33. Carlsson M, Andersson R, Markenroth Bloch K, Steding-Ehrenborg K, Mosén H, Stahlberg F, et al. Cardiac output and cardiac index measured with cardiovascular magnetic resonance in healthy subjects, elite athletes and patients with congestive heart failure. *J Cardiovasc Magn Reson*. (2012) 14:50–7. doi: 10.1186/1532-429X-14-51

34. Pankoke S, Schweitzer T, Bikker R, Pich A, Pfarrer C, Mühlfeld C, et al. Obesity impacts hypoxia adaptation of the lung. *Am J Physiol Cell Mol Physiol*. (2023) 325:L352–9. doi: 10.1152/ajplung.00125.2023

35. Jones RL, Nzekwu M-MU. The effects of body mass index on lung volumes. *Chest*. (2006) 130:827–33. doi: 10.1378/chest.130.3.827

36. Garrison WJ, Qing K, He M, Zhao L, Tustison NJ, Patrie JT, et al. Lung volume dependence and repeatability of hyperpolarized 129Xe MRI gas uptake metrics in healthy volunteers and participants with COPD. *Radiol Cardiothorac Imaging*. (2023) 5:e220096. doi: 10.1148/ryct.220096

37. Miller MR, Quanjer PH, Swanney MP, Ruppel G, Enright PL. Interpreting lung function data using 80% predicted and fixed thresholds misclassifies more than 20% of patients. *Chest*. (2011) 139:52–9. doi: 10.1378/chest.10-0189

frontiers | Frontiers in Medicine

*CORRESPONDENCE
Felix Doellinger
✉ felix.doellinger@charite.de
Marcus A. Mall
✉ marcus.mall@charite.de

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

‡These authors have contributed equally to
this work and share last authorship

Contrast agent-free functional magnetic resonance imaging with matrix pencil decomposition to quantify abnormalities in lung perfusion and ventilation in patients with cystic fibrosis

Felix Doellinger[1]*[†], Grzegorz Bauman[2,3][†], Jobst Roehmel[4,5,6],
Mirjam Stahl[4,5,6], Helena Posch[1], Ingo G. Steffen[7],
Orso Pusterla[2,3], Oliver Bieri[2,3], Mark O. Wielpütz[8,9,10][‡] and
Marcus A. Mall[4,5,6]*[‡]

[1]Department of Radiology, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany, [2]Division of Radiological Physics, Department of Radiology, University of Basel Hospital, Basel, Switzerland, [3]Department of Biomedical Engineering, University of Basel, Basel, Switzerland, [4]Department of Pediatric Respiratory Medicine, Immunology and Critical Care Medicine, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany, [5]Berlin Institute of Health at Charité-Universitätsmedizin Berlin, Berlin, Germany, [6]German Center for Lung Research (DZL), Associated Partner Site, Berlin, Germany, [7]Department of Pediatric Hematology and Oncology, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany, [8]Department of Diagnostic and Interventional Radiology, University Hospital of Heidelberg, Heidelberg, Germany, [9]Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), Heidelberg, Germany, [10]Department of Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik at University Hospital of Heidelberg, Heidelberg, Germany

Background: Previous studies showed that contrast-enhanced (CE) morpho-functional magnetic resonance imaging (MRI) detects abnormalities in lung morphology and perfusion in patients with cystic fibrosis (CF). Novel matrix pencil decomposition MRI (MP-MRI) enables quantification of lung perfusion and ventilation without intravenous contrast agent administration.

Objectives: To compare MP-MRI with established morpho-functional MRI and spirometry in patients with CF.

Methods: Thirty-nine clinically stable patients with CF (mean age 21.6 ± 10.7 years, range $8-45$ years) prospectively underwent morpho-functional MRI including CE perfusion MRI, MP-MRI and spirometry. Two blinded chest radiologists assessed morpho-functional MRI and MP-MRI employing the validated chest MRI score. In addition, MP-MRI data were processed by automated software calculating perfusion defect percentage (QDP) and ventilation defect percentage (VDP).

Results: MP perfusion score and QDP correlated strongly with the CE perfusion score (both $r = 0.81$; $p < 0.01$). MP ventilation score and VDP showed strong inverse correlations with percent predicted FEV1 ($r = -0.75$ and $r = -0.83$; $p < 0.01$). The comparison of visual and automated parameters showed that both MP perfusion score and QDP, and MP ventilation score and VDP were strongly correlated ($r = 0.74$ and $r = 0.78$; both $p < 0.01$). Further, the MP perfusion score and MP ventilation score, as well as QDP and VDP were strongly correlated ($r = 0.88$ and $r = 0.86$; both $p < 0.01$).

Conclusion: MP-MRI detects abnormalities in lung perfusion and ventilation in patients with CF without intravenous or inhaled contrast agent application, and correlates strongly with the well-established CE perfusion MRI score and spirometry. Automated analysis of MP-MRI may serve as quantitative noninvasive outcome measure for diagnostic monitoring and clinical trials.

Introduction

Cross-sectional imaging enabling the detection and localization of morphological and functional abnormalities of the lungs plays an important role in diagnostic monitoring of patients with cystic fibrosis (CF) (1, 2). The ability to localize and quantify pulmonary changes, ideally by noninvasive and effort-independent investigation, facilitates individualized therapeutic approaches. Given the onset of lung disease in early childhood, and the predicted further increase in life expectancy of patients with CF in the era of CFTR-directed therapeutics (3–5), avoidance of potential unwanted side effects of imaging techniques is increasingly important. In this context, despite its broad availability, the life-long use of chest computed tomography (CT) is limited by the risks associated with the cumulative dose of radiation exposure (6, 7). Over the past 15 years, magnetic resonance imaging (MRI) has emerged as an alternative to detect lung abnormalities in CF. (8–11) Advantages of MRI compared to CT are the absence of potentially harmful ionizing radiation, as well as the ability to perform simultaneous functional imaging studies for a combined assessment of lung morphology and functional parameters such as lung perfusion (12–16).

Morpho-functional MRI is an established technique that combines spatially resolved sequences to detect morphological changes and temporally resolved first-pass lung perfusion imaging with i.v. application of a contrast agent to assess abnormalities in lung perfusion. For semi-quantitative assessment of CF lung disease a morpho-functional chest MRI score was established that is based on a visual scoring system consisting of five morphological subscores including bronchiectasis/wall thickening and airway mucus plugging, and a perfusion score (17). This MRI score was shown to be robust and reproducible in single-and multicenter observational studies and to be sensitive to detect lung abnormalities across a broad age range of clinically stable patients with CF from infancy to adulthood (10, 11, 18), acute changes associated with pulmonary exacerbations (9), and response to therapeutic interventions including antibiotic therapy for pulmonary exacerbation, as well as cystic fibrosis transmembrane conductance regulator (CFTR) modulator therapies (13, 19, 20). However, first-pass perfusion imaging uses gadolinium-containing intravenous contrast agents that are generally safe and well tolerated, but also have some disadvantages. Gadolinium-based contrast agents can cause intolerance reactions, and although not observed with current macrocyclic contrast agents, previously used linear contrast agents were found to cause cerebral gadolinium deposits (21–23). Therefore, avoidance of intravenous contrast agents is desirable for long-term patient safety. Additional advantages of contrast agent-free

MRI include shortening of the examination time, cost savings, no need for venipuncture that can be particularly stressful for children, and no contamination of the environment and groundwater by excreted contrast medium. Therefore, contrast agent-free MRI provides an important and patient-relevant technical advance for sensitive monitoring of CF lung disease.

Several techniques exist to derive information about lung perfusion from contrast agent-free MRI (24–29). Fourier decomposition MRI (FD-MRI) (30) was the basis for several derivatives. Matrix pencil decomposition MRI (MP-MRI) (31, 32) and phase resolved functional lung MRI (PREFUL-MRI) (29) have been shown to simultaneously assess lung ventilation and perfusion in CF patients (16, 33, 34). Previously, simultaneous assessment of lung ventilation and perfusion was only possible with nuclear medicine techniques with the use of radioisotopes, whereas inhalation of hyperpolarized gases only allows studies of ventilation (12, 35–37). Another advantage of MP-MRI is the possibility to evaluate data not only visually, but also by automated software analysis. This may open a possibility to replace contrast-enhanced (CE) perfusion MRI with a contrast agent-free technique.

The aim of this study was to investigate the correlation between visually scored and automatically calculated MP-MRI parameters for lung perfusion and ventilation with the morpho-functional MRI score as well as lung function testing. We therefore conducted a prospective observational study in 39 patients with CF who underwent morpho-functional MRI, including contrast-enhanced perfusion MRI, and MP-MRI as part of the same examination and assessment of lung function by spirometry.

Materials and methods

Study design and participants

This prospective observational study was approved by the Ethics Committee of Charité-Universitätsmedizin Berlin (EA2/170/20) and informed written consent was obtained from all patients or their legal guardians. During a two-year period, a total of 47 CF patients aged eight years and older in stable clinical condition were prospectively recruited for this study. In seven patients, the patients or their legal guardians refused intravenous contrast injection and in one case CE perfusion MRI was not evaluable due to a previously unknown thrombosis of the subclavian vein (which caused delayed inflow of the contrast agent). This led to exclusion of eight patients from the analysis and a study population of 39 patients.

TABLE 1 Characteristics of study population.

Characteristic	Mean ± SD
Number of patients	39
Age, in years (range)	21.6 ± 10.7 (8-45)
Male/female (%)	18/21 (46.2/53.8)
Weight, in kg	52.8 ± 16.9
Height, in cm	161.0 ± 16.6
BMI (kg/m²)	19.7 ± 3.6
CFTR genotype	
F508del/F508del (%)	19 (48.7)
F808del/other (%)	14 (35.9)
Other/other (%)	6 (15.4)
CFTR modulator therapy	
Lum-Iva (%)	8 (20.5)
Tez-Iva (%)	2 (5.1)
Elx-Tez-Iva (%)	3 (7.7)
None	26 (66.7)
Pancreatic insufficiency (%)	34 (87.2)
CF-related diabetes (%)	6 (15.4)
Chronic *Pseudomonas aeruginosa* (%)	9 (23.1)
Chronic NTM (%)	0 (0)

BMI, body mass index; CFTR, cystic fibrosis transmembrane conductance regulator; Elx, elexacaftor; Iva, ivacaftor; Lum, lumacaftor; Tez, tezacaftor; NTM, nontuberculous mycobacteria. Data are presented as mean ± SD, if not indicated otherwise.

MRI and spirometry were performed as simultaneously as possible, which meant the same day for 20 of 39 patients. The maximum interval between the two examinations was 14 days (mean: 1 day). Patient characteristics are summarized in Table 1.

Magnetic resonance imaging

All imaging was acquired with the same 1.5 Tesla whole-body MR scanner (MAGNETOM Aera, Siemens Healthineers AG, Erlangen, Germany). MRI, including morphological sequences, MP-MRI and CE perfusion imaging, was always performed in the same examination and in the same order. Various versions of the examination protocol for mucoobstructive diseases exist at our center, in which technical parameters such as the field of view have been adapted to different age groups. The total examination time for the complete protocol averaged 15–20 min.

Morpho-functional MRI was performed as described previously (10, 11, 17), using a protocol comprising a balanced steady-state free-precession sequence (TrueFISP, Siemens) in axial, coronal and sagittal planes, T1-weighted spoiled gradient echo sequences (VIBE, Siemens) in axial and coronal planes before and after i.v. contrast administration and a fat-saturated T2-weighted sequence with rotating phase encoding (BLADE, Siemens) in axial and coronal planes (10, 11, 13, 15, 38). First-pass lung perfusion images were acquired with a spoiled gradient-echo sequence (time resolved angiography with stochastic trajectories, TWIST, Siemens) in coronal plane. During acquisition of this sequence a contrast bolus was injected via a peripheral catheter or central line at a rate of 2–4 mL/s and the first passage of the contrast bolus through the pulmonary circulation was captured within a dynamic series of up to 30 volume datasets of the whole lung, each with an acquisition time of 1.5 s. Two macrocycle-structured gadolinium-based contrast agents approved for this indication, Gadobutrol (Gadovist©, Bayer, Germany) and Gadoterate meglumine (Dotarem©, Guerbet, France), were used equally in the study. The amount of contrast medium was adjusted to the weight of our patients and was 1.5–8 mL gadobutrol, respectively, 2.3–16 mL gadoterate meglumine. As gadobutrol is twice as concentrated (1.0 mmoL/mL) as gadoterate meglumine (0.5 mmoL/mL), the total injection volume was only half as high in each case. Perfusion image datasets were post-processed by subtracting the baseline images (acquired prior to the contrast agent administration) from those with maximal contrast enhancement in the lung parenchyma.

Contrast agent-free functional imaging was performed as previously described using the MP-MRI technique (30–32, 39, 40). MP-MRI is a derivative of Fourier decomposition MRI and a non-invasive and contrast agent-free technique for the assessment of pulmonary ventilation and perfusion (32). MP-MRI relies on dynamic free-breathing ultra-fast balanced steady-state free precession (uf-bSSFP) acquisitions of chest images and provides regional ventilation and perfusion information from a single acquisition series (39). The uf-bSSFP pulse sequence uses an optimized excitation pulses and gradient switching patterns of a conventional Cartesian bSSFP imaging scheme accompanied by partial echo readouts and ramp sampling techniques. As a result, echo time and repetition time are shortened, which improves signal in the lung parenchyma and reduces motion as well as off-resonance artifacts. In order to cover the whole chest volume, imaging is performed using between 8–10 slices. Each slice consists of 160 coronal images acquired in approximately 50 s. Spectral information extracted using MP decomposition comprising respiratory and cardiac frequencies, amplitudes and phases is used to generate spatially-resolved fractional ventilation and perfusion maps. The technique requires no further patient compliance such as prolonged breath-holding. The time-resolved uf-bSSFP image series are processed by elastic image registration for compensation of respiratory motion (40). The registration algorithm preserves signal magnitude in each image but automatically aligns lung structures such as vessels, thoracic wall and airways. Subsequently, the lung regions on the registered images are automatically segmented using an artificial neural network (31). The motion-corrected and segmented images are further analyzed voxel-wise using matrix pencil decomposition to extract the amplitudes of the respiratory and cardiac signal modulations in the lung parenchyma, and were used to generate fractional ventilation and perfusion images. The procedures for image registration, lung segmentation using an artificial neural network and matrix pencil decomposition were combined into an image processing pipeline for MP-MRI. The software was written in Python and C++ for Linux operating system using CUDA Toolkit 12.0 (NVIDIA Corp., Santa Clara, CA), ITK library 5.2 (41) and Armadillo library 12.0 (NICTA, Brisbane, Australia). The image postprocessing and analysis was performed on a workstation equipped with 2x Epyc 7502 CPU (AMD Inc. Santa Clara, CA) and Quatro RTX 8000 GPU (NVIDIA Corp.). The data processing took approximately 10 min per subject. MP-MRI was always performed before contrast agent injection to avoid bias caused by the contrast agent.

Visual image assessment

In the first step of data analysis, two experienced thoracic radiologists with 13 years (MOW) and 6 years (FD) of experience in chest MRI of CF patients were blinded to clinical and demographic characteristics and applied the principles of the validated morpho-functional MRI score (17) to morphological MRI sequences, subtraction images of contrast-enhanced first-pass perfusion imaging and to color-coded maps computed from MP-MRI data.

In the morpho-functional MRI score, five morphologic subscores and a perfusion score assessed on subtracted perfusion maps from contrast-enhanced first-pass perfusion imaging are assessed for each lobe and the lingula separately. The five morphologic subscores for (I) bronchiectasis and wall thickening, (II) mucus plugging, (III) sacculations and abscesses, (IV) consolidations, and (V) pleural reactions (including effusions) sum up to the morphology score and the global score results from the sum of morphology score and perfusion score (17). Figure 1 shows CF-related MRI findings resulting in the morpho-functional chest MRI score. The extent of pulmonary morphologic and perfusion abnormalities is graded separately for each lobe as follows: 0 (no abnormality), 1 (<50% of the lobe), or 2 (≥50% of lobe affected). Because this system considers the lingula as a separate, sixth lobe, maximum scores of 12 can be obtained for each subscore, 60 for the morphology score, and 72 for the global score.

The color-coded MP-MRI perfusion and ventilation maps were scored in analogy to the MRI perfusion score, but blinded to morpho-functional MRI (including CE perfusion MRI) and the results of automated software analysis. The principle of the morpho-functional score was easily transferred to the color-coded visualizations of the MP-MRI perfusion and ventilation analyses, which were also scored with 0–1–2 points for each lobe separately.

The scores obtained by the two readers were very consistent and data were analysed using arithmetic means.

Automated assessment of MP-MRI data

In the second step of data analysis, fractional ventilation and perfusion MP-MRI data were automatically analyzed using a dedicated in-house developed software (31). The distributions of voxel values from ventilation and perfusion images obtained in the segmented lung regions (S) were used to calculate a threshold value (T) used to identify regions with ventilation and perfusion impairment. Lung regions with values below the threshold value were classified as regions with functional impairment. The threshold values were calculated for each slice independently. The extent of perfusion and ventilation defects detected by MP-MRI was expressed as the percentage of hypoperfused lung volume (perfusion defect percentage, QDP) and hypoventilated lung volume (ventilation defect percentage, VDP) to the total lung volume. In addition, an experimental parameter "ventilation/perfusion overlap" (VQO) was calculated, describing the spatial overlap of ventilation and perfusion. The parameter is expressed as the percentage of simultaneously hypoperfused and hypoventilated lung volume. VQO can be interpreted as a measure of the hypoxic pulmonary vasoconstriction in which the localized hypoxia leads to vasoconstriction with reduced perfusion in the affected bronchovascular territory.

Instead of using a median-based threshold as in previous studies with MP-MRI, we employed a method which estimates the threshold from the mean difference between maximum and minimum intensity projections applied on the ventilation and perfusion images. The main advantage of this method is its improved robustness to signal inhomogeneities caused by coil sensitivities. Threshold values (T) used for the identification of impaired ventilation and perfusion lung regions were calculated is the following way:

- Maximum and minimum intensity projections were performed for voxels inside the two-dimensional segmented lung region (S) along each dimension:

$$\mathrm{MIP}_x\left(y\right) = \max\left(S\left(x,y\right)\right), \mathrm{MIP}_y\left(x\right) = \max\left(S\left(x,y\right)\right)$$

and

$$\mathrm{MinIP}_x\left(y\right) = \min\left(S\left(x,y\right)\right), \mathrm{MinIP}_y\left(x\right) = \min\left(S\left(x,y\right)\right)$$

FIGURE 1
CF-related MRI findings representing the subscores of the morpho-functional chest MRI score. **(A–E)** Show image sections of a high-resolution T2-weighted sequence acquired in transversal plane. The arrows indicate bronchiectasis/wall thickening **(A)**, mucus plugging **(B)**, sacculations **(C)**, consolidation **(D)** and an inflammatory pleural thickening **(E)**, which is part of the score as a "special finding." **(F)** Shows a post-processed subtraction image from the contrast-enhanced perfusion MRI acquired in coronal plane. The arrows mark a large perfusion deficit in the left upper lobe and a small perfusion deficit in the middle lobe. By definition, each subscore in each lobe is assessed individually and 0−2 points are scored per lobe depending on whether the corresponding pathology is not present (0), involves less than 50% (1) or more than 50% (2) of the lobe.

- The resulting one-dimensional projections were then interpolated to the same maximum number of elements for calculation of the square-root of the Hadamard product:

$$\mathrm{MIP}_{xy} = \sqrt{\mathrm{MIP}_x \cdot \mathrm{MIP}_y}, \mathrm{MinIP}_{xy} = \sqrt{\mathrm{MinIP}_x \cdot \mathrm{MinIP}_y}$$

- The threshold T is calculated as the mean value of the resulting maximum and minimum projection difference:

$$T = \overline{\mathrm{MIP}_{xy} - \mathrm{MinIP}_{xy}}$$

Visual and software analysis of MP-MRI data was possible in all patients. MP-MRI postprocessing was fully automated and required no manual interaction.

Spirometry

Spirometry was performed according to standards endorsed by the American Thoracic Society and European Respiratory Society (42, 43). Percent predicted values were calculated using the Global Lung Initiative (GLI) reference values (44). Forced expiratory volume in one second (FEV1), mid-expiratory flow at 25% of vital capacity (MEF25) and forced vital capacity (FVC) were included in the analyses. Percent predicted results were based on equations of the global lung function initiative (45).

Statistical analysis

Statistical analyses were performed with R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria). According to histograms and Kolmogorov–Smirnov tests, a normal distribution of metric data was not assumed. Data are presented as mean ± SD, median, interquartile range (IQR; 25th–75th percentile), and range (min–max). Spearman's rank correlation coefficient r was used to describe the association between metric parameters. The correlation coefficient r measures how close the relationship between two parameters is to linearity. As two parameters can correlate both positively and negatively, r can be a minimum of −1.0 and a maximum of 1.0 (46). According to *Rowntree* and *Karlik* the correlations are assumed to be negligible ($r = 0.0$–0.2), weak/low ($r = 0.2$–0.4), moderate ($r = 0.4$–0.7), strong (0.7–0.9), and very strong (0.9–1.0) (46, 47). A p-value <0.05 was considered statistically significant.

Results

Characteristics of study population

Thirty-nine patients with CF aged 8 to 45 years with a mean age of 21.6 ± 10.7 years and a mean percent predicted forced expiratory volume in 1 s (ppFEV1) of 76 ± 28% were enrolled in this study. Our study population showed an almost balanced sex ratio and the distribution of CFTR genotypes and proportion of exocrine pancreatic insufficiency was representative for CF patients in Central Europe (Table 1). Nine out of 39

patients were chronically infected with *Pseudomonas aeruginosa*, and one-third of the patients was on CFTR modulator therapy at the time of the study MRI (Table 1). MRI detected signs of CF lung disease in all patients (Figure 2 and Table 2). As previously reported (10, 11), the subscore bronchiectasis/wall thickening was the most important contributor to the MRI morphology score, followed by the mucus plugging subscore, whereas the contribution of the other morphologic subscores (i.e., abscesses/sacculations; consolidation; special findings) was low in our cohort of clinically stable patients (Tables 2, 3). The MRI global score was 24.5 ± 9.9 and the CE perfusion score was 6.9 ± 2.8. Visual scores of the experimental MP-MRI sequences were 7.2 ± 1.7 for the MP perfusion score and 6.7 ± 1.8 for the MP ventilation score (Table 3).

MP-MRI enables assessment of lung perfusion and correlates with contrast-enhanced perfusion MRI

First, we compared assessment of lung perfusion by MP-MRI to CE first-pass MR angiography as the established gold standard for studies of lung perfusion by MRI (Figure 2). Visual evaluation of the MP perfusion score showed a strong correlation with the established CE perfusion score ($r = 0.81$; $p < 0.01$) (Figure 3). Similarly, QDP derived from automated analysis of MP-MRI showed a strong correlation with the CE perfusion score ($r = 0.81$; $p < 0.01$). Comparison of the two MP-MRI parameters of lung perfusion showed a strong correlation between the visual MP perfusion score and automated QDP ($r = 0.74$; $p < 0.01$) (Figure 4). MP perfusion score and QDP also showed strong correlations with MRI global score, MRI morphology score, and bronchiectasis/wall thickening subscore as well as the CE perfusion score (Figure 4 and Table 4).

MP-MRI enables assessment of pulmonary ventilation and correlates with spirometry

Next, we determined the relationship between MP-MRI ventilation and spirometry. The visual MP ventilation score showed strong inverse correlations with ppFEV1 ($r = −0.75$; $p < 0.01$) and other spirometry outcomes (Figure 5 and Table 4). Further, VDP determined from automated MP-MRI analysis showed a strong inverse correlation with ppFEV1 ($r = −0.83$; $p < 0.01$) and other spirometry outcomes (Figure 5 and Table 4). The visual MP ventilation score and automated VDP were strongly correlated with each other ($r = 0.78$; $p < 0.01$) (Figure 5). Both MP ventilation parameters, the MP ventilation score and VDP, showed strong correlations with the MRI global score, MRI morphology score and bronchiectasis/wall thickening subscore as well as the CE perfusion score (Figure 4 and Table 4).

Relationship between lung perfusion and ventilation determined by MP-MRI

Finally, we determined the relationship between lung perfusion and ventilation determined by MP-MRI in CF lung disease. The visual MP perfusion score showed a strong correlation with the MP ventilation score ($r = 0.88$; $p < 0.01$) (Figure 6). Similarly, the automated QDP showed a strong correlation with VDP ($r = 0.86$; $p < 0.01$) (Figure 6). Finally, a crossover comparison showed intermediate to strong correlations
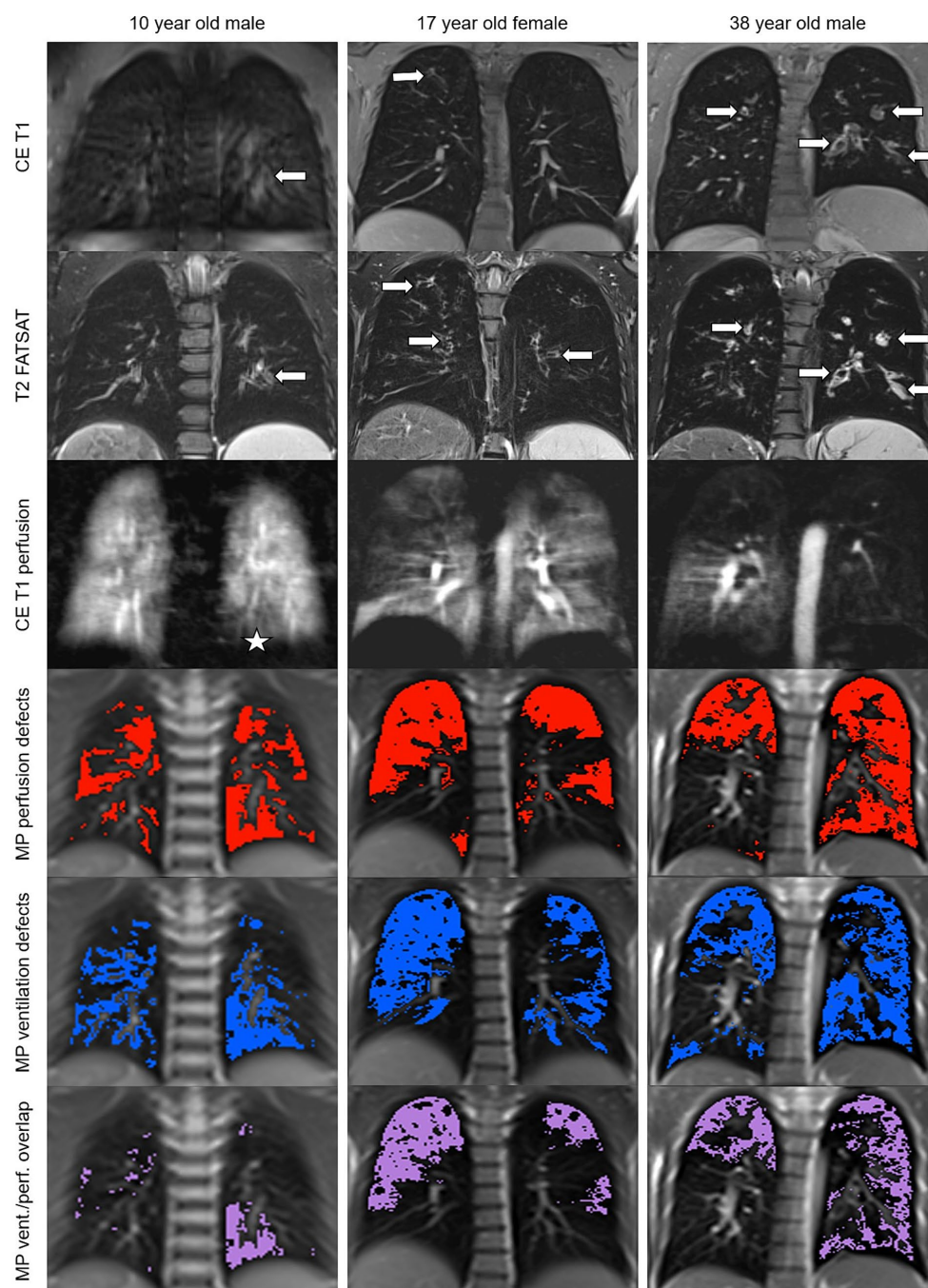
FIGURE 2
Changes in lung morphology, perfusion and ventilation detected by morpho-functional MRI and MP-MRI in patients with CF. Representative images of a 10 year old school child, a 17 year old adolescent and a 38 year old adult with CF. Contrast-enhanced T1-weighted and fat-saturated T2-weighted MRI shows dilated and/or wall-thickened bronchi (arrows). Contrast-enhanced perfusion MRI and color-coded visualizations of MP-MRI data show areas of decreased lung perfusion and ventilation as well as the overlap between the perfusion and ventilation defects that match with the location of the above morphological changes (asterisk in the youngest patient, obvious in the other two).

between the visual MP perfusion score and VDP, and between the visual MP ventilation score and QDP (Figure 6). The experimental parameter ventilation/perfusion overlap (VQO) showed comparably strong correlations with the spirometry results and the visual scores as the more established parameters QDP and VDP from previous studies (Table 4). Taken together, these results demonstrate the capability of MP-MRI for simultaneous quantitative assessment of abnormalities in lung ventilation and perfusion by MRI in our cohort of CF patients.

Discussion

In a cohort of 39 CF patients with a broad range of lung disease severity, we show that lung perfusion determined by contrast agent-free MP-MRI correlates strongly with CE perfusion MRI (2, 10, 11, 28, 48). Second, we show that MP-MRI enables simultaneous quantitative assessment of lung perfusion and ventilation, and that lung ventilation determined by MP-MRI shows a strong correlation with spirometry.

TABLE 2 Prevalence and magnitude of changes in lung structure and perfusion detected by chest MRI in patients with CF in the total study population and grouped by age.

Age group, years		Total	≤9	10–19	20–29	30–39	≥40
Patients, n		39	4	16	9	7	3
Mean age, years		21.6	8.5	13.8	24.8	33.6	42.7
Mean BMI		19.7	14.4	19.2	21.1	21.6	21.6
Global	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	24.5 (9.9)	21.1 (6.3)	16.9 (6.9)	29.4 (8.3)	34.6 (5.2)	31.2 (3.1)
Morphology	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	17.6 (7.4)	15.6 (5.5)	12.0 (5.6)	21.0 (5.7)	25.0 (4.2)	22.3 (2.8)
Bronchiectasis/wall thickening	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	8.7 (2.6)	8.1 (2.1)	6.6 (1.6)	9.7 (2.3)	11.4 (1.1)	10.8 (1.1)
Mucus plugging	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	4.9 (2.5)	4.4 (0.7)	3.1 (1.5)	6.8 (2.5)	5.9 (2.3)	7.2 (1.5)
Abscesses/sacculations	Prevalence	13 (33.3)	0 (0)	3 (18.8)	3 (33.3)	5 (71.4)	2 (66.7)
	Score	0.4 (0.9)	0.0 (0.0)	0.2 (0.4)	0.6 (1.4)	1.0 (0.8)	0.3 (0.2)
Consolidation	Prevalence	25 (64.1)	1 (25.0)	7 (43.8)	7 (77.8)	7 (85.7)	3 (100)
	Score	1.0 (1.3)	0.3 (0.4)	0.7 (1.1)	0.9 (0.7)	2.5 (1.6)	1.0 (0.4)
Special findings	Prevalence	32 (82.1)	3 (75.0)	10 (62.5)	9 (100)	7 (100)	3 (100)
	Score	2.6 (2.1)	2.9 (2.4)	1.5 (1.6)	3.1 (1.7)	4.3 (1.8)	3.0 (1.5)
CE perfusion	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	6.9 (2.8)	5.5 (1.2)	4.8 (1.8)	8.4 (2.8)	9.6 (1.3)	8.8 (1.0)
MP perfusion	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	7.2 (1.7)	7.0 (0.9)	6.0 (1.2)	7.9 (1.7)	8.7 (1.2)	7.5 (0.0)
MP ventilation	Prevalence	39 (100)	4 (100)	16 (100)	9 (100)	7 (100)	3 (100)
	Score	6.7 (1.8)	6.8 (0.3)	5.5 (1.4)	7.2 (1.8)	8.6 (1.3)	6.7 (0.6)

Prevalence data are presented as n of patients (percentage) and absolute MRI scores as mean (SD). BMI, body mass index; CE, contrast-enhanced; MP-MRI, matrix pencil decomposition MRI; MRI, magnetic resonance imaging.

Third, our MP-MRI-derived metrics of lung perfusion and ventilation demonstrate a strong cross-correlation in patients with CF. Finally, we show that automated analysis of MP-MRI computing QDP and VDP strongly correlates with reader-based visual scoring of MP-MRI and may therefore be used for unbiased quantitative analysis of changes in lung perfusion and ventilation. Collectively, our data support the use of MP-MRI for non-invasive quantitative studies of abnormalities in lung perfusion and ventilation in patients with CF.

To the best of our knowledge, this is the first study that directly compared contrast agent-free functional MP-MRI, established contrast-enhanced morpho-functional MRI, and spirometry for detection of abnormalities in the lungs of patients with CF. It is also the first study to use MP-MRI in adult CF patients. However, this is not the first study ever to use contrast-agent-free functional MP-MRI in patients with cystic fibrosis. In a study from 2018, 23 CF patients aged 6–18 years and a comparison group of healthy children underwent MP-MRI on two consecutive days (49). The automated post-processing of the MRI data was carried out with a previous version of the software used in our study. QDP and VDP were almost identical in both examinations, which proved a very good short-term reproducibility of the MP-MRI measurements. In addition, QDP and VDP were significantly higher in the CF patients than in the healthy controls. As in our study, a strong correlation was found between

FEV1 and VDP ($r = 0.65$), but this was even stronger in our study ($r = 0.83$). In another, more recent study, 24 children with CF were examined with MP-MRI and spirometry at one-year intervals in order to demonstrate the therapeutic effect of CFTR modulator therapy (16). Morphologic lung changes were assessed visually using the chest MRI score as in our study and MP-MRI data were quantified using a previous version of our software. The positive therapeutic effect was demonstrated by the fact that QDP and VDP decreased significantly under modulator therapy. However, both MP-MRI studies on CF patients listed here were performed without contrast administration, which prevented a comparison of the MP-MRI perfusion analyses with the gold standard of MRI perfusion imaging.

Abnormalities in lung perfusion are a hallmark of CF that are associated with lung disease severity in pediatric and adult patients (10, 11, 13, 17). The MP-MRI technique used for contrast agent-free assessment of lung perfusion in our study is a further development of FD-MRI initially introduced more than a decade ago (30). The strong correlation between perfusion signals obtained by MP-MRI and CE MRI, i.e., the current standard of reference for MRI-based perfusion studies, support the future use of MP-MRI for contrast agent-free quantification of perfusion defects associated with CF lung disease. Another derivative of FD-MRI is perfusion-weighted phase-resolved functional lung (PREFUL) MRI (29), that has also been used

TABLE 3 Results of spirometry, morpho-functional MRI and MP-MRI.

	Parameters	Mean	SD	Median	Q25	Q75	Min	Max	Range
PFT	ppFEV1	76.1	28.3	78.0	51.6	100.3	28.1	127.5	99.4
	ppMEF25	47.8	36.7	37.3	18.4	72.7	6.6	139.3	132.7
	ppFVC	87.7	21.1	89.0	73.8	105.7	42.4	132.7	90.3
Morpho-functional MRI	Global score	24.5	9.9	27.5	14.8	31.2	7.5	42.0	34.5
	Morphology score	17.6	7.4	20.0	10.0	23.5	5.0	32.5	27.5
	Bronchiectasis/wall thickening	8.7	2.6	9.0	6.0	11.2	4.0	12.0	8.0
	Mucus plugging	4.9	2.5	4.5	3.0	6.0	0.5	11.0	10.5
	Abscesses/sacculations	0.4	0.9	0.0	0.0	0.5	0.0	4.5	4.5
	Consolidation	1.0	1.3	0.5	0.0	1.2	0.0	4.0	4.0
	Special findings	2.6	2.1	2.0	1.0	4.2	0.0	7.0	7.0
	CE perfusion score	6.9	2.8	6.5	5.0	9.0	0.5	11.5	11.0
MP-MRI	MP perfusion score	7.2	1.7	7.0	6.0	8.5	3.5	10.5	7.0
	MP ventilation score	6.7	1.8	6.5	5.5	7.2	2.5	10.5	8.0
	QDP	22.8	19.7	45.3	29.0	60.9	1.4	74.0	72.6
	VDP	44.5	15.3	19.5	10.9	35.2	0.3	53.7	53.4
	VQO	38.7	18.5	39.0	23.3	54.9	4.0	69.8	65.8

CE, contrast-enhanced; MP, matrix pencil decomposition; MRI, magnetic resonance imaging; PFT, pulmonary function test (spirometry); ppFEV1, percent predicted forced expiratory volume in 1 s; ppFVC, percent predicted forced vital capacity; ppMEF25, percent predicted maximum expiratory flow at 25 percent; Q25, 25th quartile; Q75, 75th quartile; QDP, ventilation defect percentage; SD, standard deviation; VDP, perfusion defect percentage; VQO, ventilation/perfusion overlap.

FIGURE 3
Correlation between lung perfusion determined by MP-MRI and contrast-enhanced perfusion MRI. (A) Correlation between visual MP perfusion score and visual CE perfusion score. (B) Correlation between automated QDP and visual CE perfusion score. (C) Correlation between automated QDP and visual MP perfusion score. Spearman's rank correlation coefficient r and p-values are provided for each correlation.
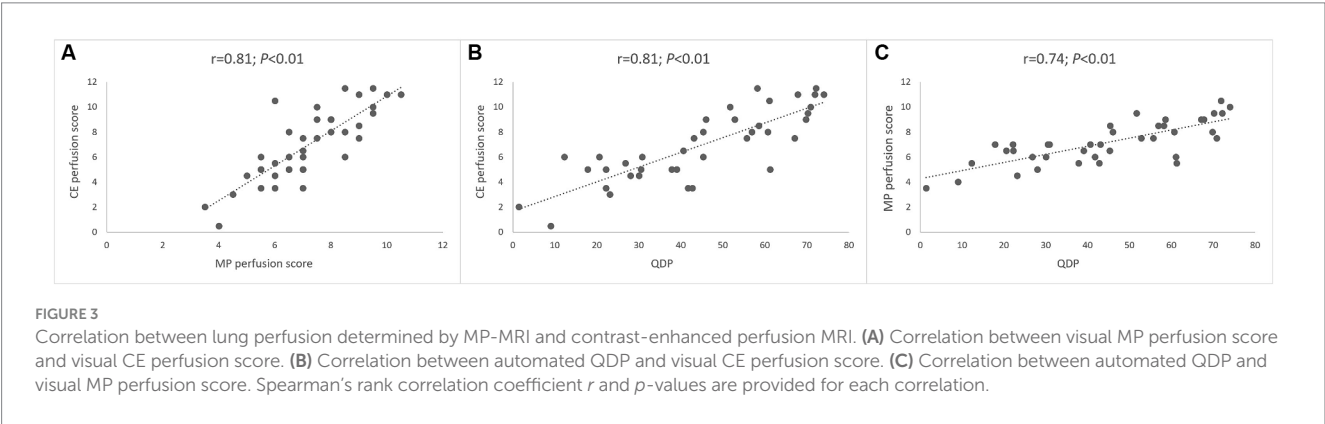
successfully for contrast agent-free assessment of lung perfusion in CF, chronic obstructive pulmonary disease (COPD) and chronic thrombembolic pulmonary hypertension (CTEPH) (34, 50–52). Similar to the visual perfusion scores we used in the present study, binary perfusion defect maps were generated and semi-quantitatively evaluated to compare the perfusion signals obtained with PREFUL-MRI to CE perfusion MRI and single-photon emission computed tomography SPECT (50). While a good overall agreement between the perfusion maps across the three different techniques was reported, this decreased significantly near the heart and diaphragm in the maps derived from PREFUL-MRI data, probably due to susceptibility to motion artifacts (50). Such a decrease in agreement between FD-based perfusion signals and CE perfusion MRI in the aforementioned artifact-prone regions was not observed for the MP-MRI perfusion maps generated in our study (Figure 2), suggesting

a potential advantage for assessment of abnormal perfusion in these lung regions.

Another advantage of MRI is that it can be used for the assessment of lung ventilation (2, 27, 32, 49, 51, 52). So far, the most common approach for MRI-based ventilation imaging has been the use of hyperpolarized gases such as ^{3}He, ^{129}Xe and ^{19}F as inhaled contrast agents that enable visualization of regional ventilation defects in patients with CF. (53–56) However, with a very limited number of sites equipped to perform hyperpolarized gas MRI worldwide, access to this technique remains limited, also limiting its use for multi-center studies. FD-MRI-based techniques such as MP-MRI allow assessment of pulmonary ventilation in addition to pulmonary perfusion and can be installed on standard 1.5 T MRI scanners. This facilitates their use for diagnostic monitoring as well as outcome measurement in multicenter
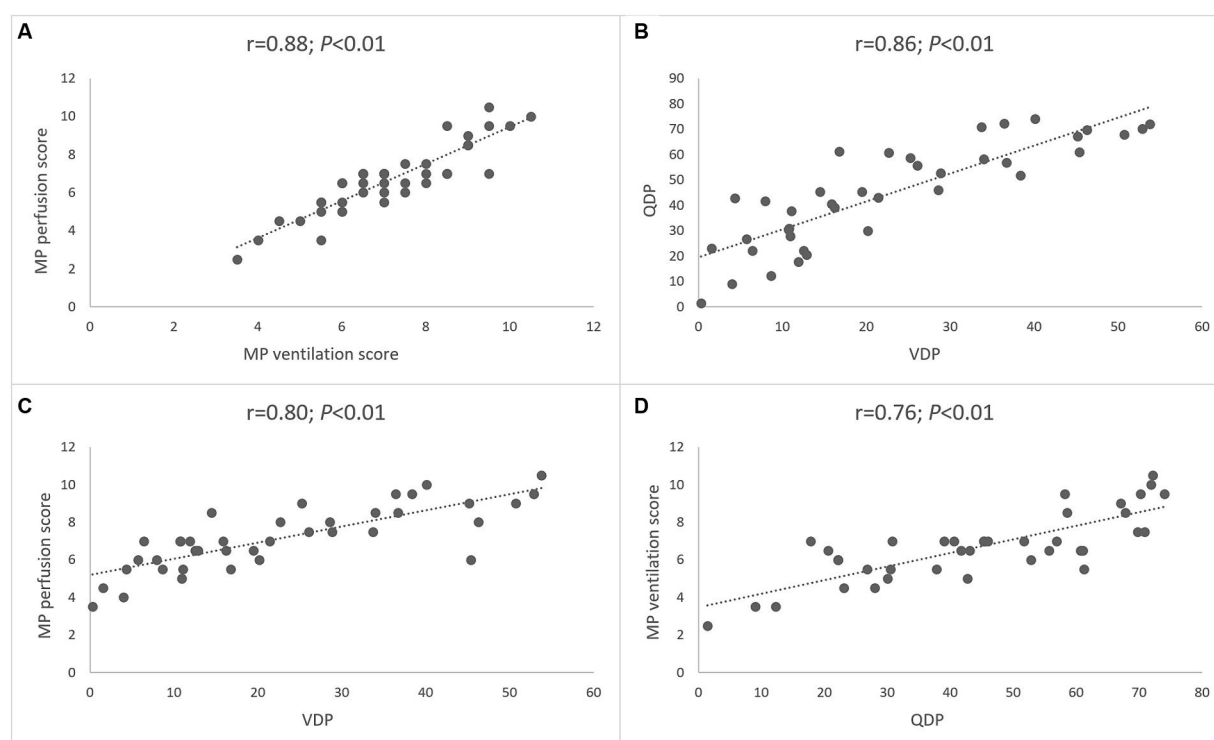
FIGURE 4
Correlation between morpho-functional MRI scores and MP-MRI parameters. The first and second row shows correlations of the MRI global score and visual MP perfusion score **(A)**, automated perfusion defect percentage **(B)**, visual MP ventilation score **(C)** and automated ventilation defect percentage **(D)**. The third and fourth row shows correlations of the MRI morphology score and visual MP perfusion score **(E)**, automated perfusion defect percentage **(F)**, visual MP ventilation score **(G)** and automated ventilation defect percentage **(H)**. Spearman's rank correlation coefficient *r* and *p*-values are provided for each correlation.

studies, as previously shown for contrast-enhanced morpho-functional MRI (10, 13, 18, 20). For application at higher field strength (i.e., 3 T) data acquisition with spoiled gradient echo (SPGR) based techniques or transient SPGR pulse sequence can be better suited as the mitigation of the off-resonance artifacts with bSSFP imaging used by MP-MRI is more demanding at high magnetic field strength (57). However, the clinical experience with MP-MRI at 3 T is currently limited.

The strong correlation between ventilation determined by MP-MRI and spirometry observed in our study supports a potential role of MP-MRI for quantitative, but also regional assessment of lung function in patients with CF. Similarly, PREFUL-MRI was used successfully for assessment of lung ventilation in patients with chronic pulmonary diseases including CF. (14, 51, 52, 58) The ultra-fast balanced steady-state free precession sequence of the MP-MRI technique has the advantage that lung perfusion and ventilation can be assessed simultaneously, leading to an examination time of 6–7 min for complete acquisition of MP-MRI data in free breathing, whereas slightly longer examination time has been reported in studies using PREFUL-MRI (14). On the other hand, there are obvious advantages of PREFUL-MRI over MP-MRI. The most important is the significantly better portability of this contrast-agent-free functional MRI technique to MRI scanners from other manufacturers and with other field strengths, since PREFUL-MRI is now a commercially available protocol and MP-MRI uses a custom uf-bSSFP pulse sequence—although MP-MRI is now part of the routine at our center. However, a direct comparison of

MP-MRI with PREFUL-MRI for contrast agent-free functional lung imaging is still pending. The strong correlation between lung ventilation and perfusion found in our study is consistent with results from a previous experimental study that compared FD-MRI, CE perfusion MRI, and ^{3}He MRI in a porcine model of artificial bronchus occlusion (35), as well as previous studies comparing the lung clearance index derived from the multiple-breath washout technique with contrast-enhanced perfusion derived from morpho-functional MRI in patients with CF. (10) Collectively, these results support that regional ventilation impairment due to airway mucus plugging leads to hypoxic pulmonary vasoconstriction as an important mechanism of impaired lung perfusion in CF.

MP-MRI uses data acquired during tidal breathing and employs image registration to compensate for the respiratory motion. Main outcome parameters characterizing the functional impairment, namely VDP and QDP, are generally robust with respect to respiratory patterns. The distribution of the perfusion signal used for calculation of QDP is spectrally isolated from other signal modulations caused by respiratory motion. The fractional ventilation depends on the breathing amplitude and the density change of lung parenchyma. VDP as a relative measure of ventilation inhomogeneity does not depend on the absolute values of fractional ventilation. However, low density changes in the lung tissue caused by shallow breathing can more difficult to detect due to the inherently low signal of the lung and corruption by noise.

This study has limitations. Our cross-sectional study in clinically stable patients does not provide information on the causal and

TABLE 4 Correlations of spirometry, morpho-functional MRI and MP-MRI (Spearman's rank correlation coefficient *r*, all *p*-values <0.01).

		PFT			Morpho-functional MRI								MP-MRI				
		ppFEV1	ppMEF 25	ppFVC	Global score	Morphology score	Bronchiectasis/ wall thickening	Mucus plugging	Abscesses/ sacculations	Consolidation	Special findings	CE perfusion score	MP perfusion score	MP ventilation score	QDP	VDP	VQO
PFT	ppFEV1		0.95	0.92	−0.76	−0.70	−0.71	−0.56	−0.42	−0.39	−0.54	−0.79	−0.69	−0.75	−0.78	−0.83	−0.84
	ppMEF 25	0.95		0.83	−0.78	−0.72	−0.74	−0.58	−0.51	−0.44	−0.57	−0.80	−0.68	−0.70	−0.75	−0.82	−0.74
	ppFVC	0.92	0.83		−0.63	−0.54	−0.58	−0.44	−0.25	−0.24	−0.38	−0.68	−0.62	−0.67	−0.73	−0.74	−0.77
Morpho- functional MRI	Global score	−0.76	−0.78	−0.63		0.98	0.88	0.83	0.59	0.71	0.81	0.93	0.78	0.76	0.80	0.83	0.80
	Morphology score	−0.70	−0.72	−0.54	0.98		0.86	0.83	0.61	0.77	0.84	0.85	0.72	0.68	0.74	0.73	0.74
	Bronchiectasis/ wall thickening	−0.71	−0.74	−0.58	0.88	0.86		0.78	0.47	0.58	0.65	0.83	0.80	0.70	0.73	0.79	0.76
	Mucus plugging	−0.56	−0.58	−0.44	0.83	0.83	0.78		0.31	0.57	0.58	0.74	0.68	0.58	0.64	0.70	0.66
	Abscesses/ sacculations	−0.42	−0.51	−0.25	0.59	0.61	0.47	0.31		0.65	0.59	0.52	0.33	0.37	0.33	0.32	0.34
	Consolidation	−0.39	−0.44	−0.24	0.71	0.77	0.58	0.57	0.65		0.67	0.55	0.37	0.26	0.41	0.39	0.36
	Special findings	−0.54	−0.57	−0.38	0.81	0.84	0.65	0.58	0.59	0.67		0.74	0.55	0.57	0.57	0.51	0.53
	CE perfusion score	−0.79	−0.80	−0.68	0.93	0.85	0.83	0.74	0.52	0.55	0.74		0.81	0.80	0.81	0.89	0.87
MP-MRI	MP perfusion score	−0.69	−0.68	−0.62	0.78	0.72	0.80	0.68	0.33	0.37	0.55	0.81		0.88	0.74	0.80	0.81
	MP ventilation score	−0.75	−0.70	−0.67	0.76	0.68	0.70	0.58	0.37	0.26	0.57	0.80	0.88		0.76	0.78	0.79
	QDP	−0.78	−0.75	−0.65	0.80	0.74	0.73	0.64	0.33	0.41	0.57	0.81	0.74	0.76		0.86	0.94
	VDP	−0.83	−0.82	−0.74	0.83	0.73	0.79	0.70	0.32	0.39	0.51	0.89	0.80	0.78	0.86		0.95
	VQO	−0.84	−0.74	−0.77	0.80	0.74	0.76	0.66	0.34	0.36	0.53	0.87	0.81	0.79	0.94	0.95	

CE, contrast-enhanced; FVC, forced vital capacity; MEF 25, maximum expiratory flow at 25 percent; MP, matrix pencil decomposition; MRI, magnetic resonance imaging; PFT, pulmonary function test (spirometry); ppFEV1, percent predicted forced expiratory volume in 1 s; QDP, ventilation defect percentage; VDP, perfusion defect percentage; VQO, ventilation/perfusion overlap.
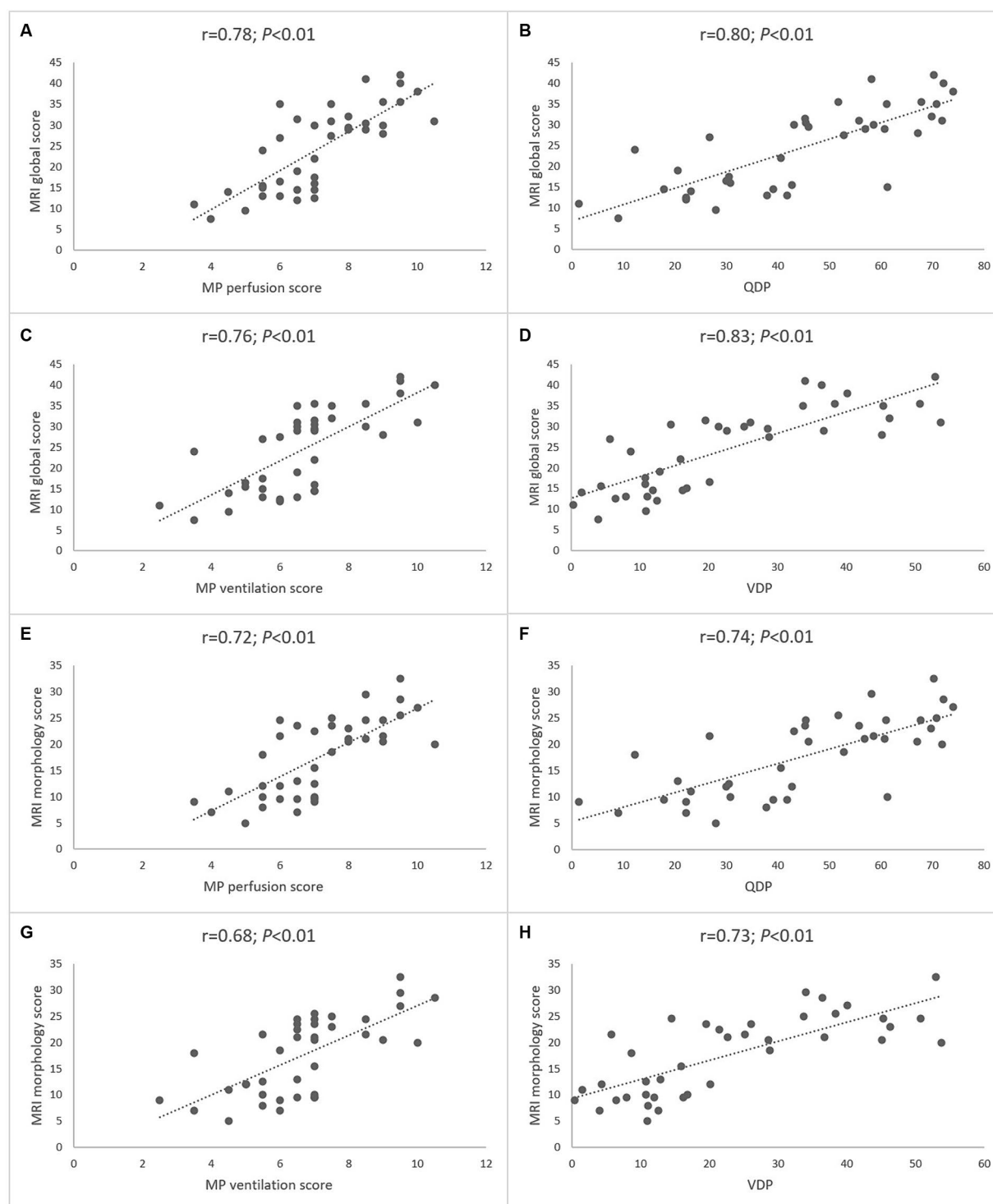
FIGURE 5
Correlation between lung ventilation determined by MP-MRI and spirometry. **(A)** Correlation between visual MP ventilation score and ppFEV1.
(B) Correlation between automated VDP and ppFEV1. **(C)** Correlation between automated VDP and visual MP ventilation score. Spearman's rank
correlation coefficient r and p-values are provided for each correlation.

long-term relationship between impairments in lung ventilation and perfusion detected by MP-MRI in CF lung disease. Therefore, longitudinal studies ideally including infants and preschool children will be required to address pertinent questions such as changes during acute pulmonary exacerbations and reversibility of abnormalities and

response to therapy including CFTR modulator therapy that targets the underlying cause of the disease at different stages of CF lung disease. These longitudinal studies will be facilitated by the automated quantitative analysis of MP-MRI ventilation and perfusion data presented in our study. As the software available for this study was not
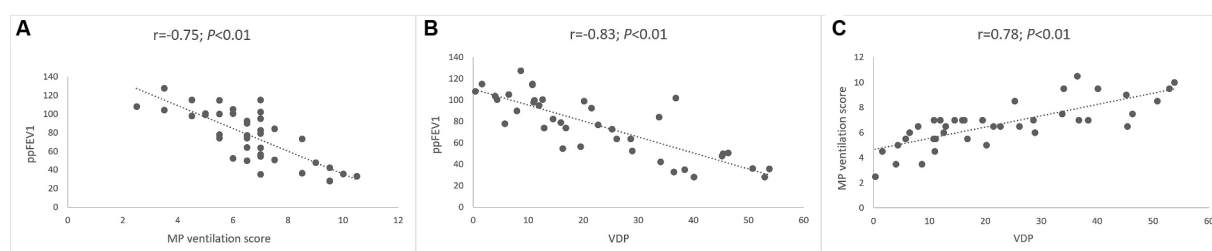
FIGURE 6
Relationship between lung perfusion and ventilation determined by MP-MRI. **(A)** Correlation between visual MP ventilation score and MP perfusion score. **(B)** Correlation between automated VDP and QDP. **(C)** Correlation between automated VDP and visual MP perfusion score. **(D)** Correlation between automated QDP and visual MP ventilation score. Spearman's rank correlation coefficient *r* and *p*-values are provided for each correlation.

yet able to calculate regional differences, the automated parameters QDP and VDP are currently global measurements without making use of the information on spatial distribution, i.e., similar to other outcome measures of lung function such as spirometry (10). However, the automated parameters QDP and VDP were strongly correlated with visual scores so that regional information may be obtained from the color-coded visualizations by reader-based scoring, and analysis software for automated analysis at regional resolution is expected to become available in the near future (31).

In summary, this study demonstrates for the first time that MP-MRI enables simultaneous assessment of lung perfusion and ventilation over a broad range of CF lung disease severity employing 1.5 T MRI without the need for intravenous or inhaled contrast agents. Together with the possibility of fully automated data analysis, these properties support MP-MRI as novel functional non-invasive quantitative outcome measure for diagnostic monitoring and clinical trials in patients with CF and potentially other muco-obstructive lung diseases in conjunction with morphological MRI.

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 humans were approved by Ethikkommission der Charité Universitätsmedizin Berlin. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

FD: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. GB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JR: Data curation, Formal analysis, Investigation,

Writing – review & editing. MS: Supervision, Validation, Writing – review & editing, Conceptualization, Investigation, Methodology. HP: Data curation, Formal analysis, Investigation, Writing – review & editing. IS: Conceptualization, Data curation, Formal analysis, Investigation, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. OP: Data curation, Formal analysis, Investigation, Software, Writing – review & editing. OB: Investigation, Project administration, Software, Supervision, Writing – review & editing. MW: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – review & editing. MM: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Conflict of interest

MS reports research grants from the German Federal Ministry of Education and Research (BMBF), the German Research Foundation (DFG) and Vertex Pharmaceuticals paid to the institution; personal fees for participation in advisory boards or consulting from Vertex Pharmaceuticals. MW reports research grants from Vertex Pharmaceuticals paid to the institution; consulting fees from

Boehringer Ingelheim and Vertex Pharmaceuticals paid to the institution; lecture fees from Vertex Pharmaceuticals paid to the institution. MM reports research grants from the German Federal Ministry of Education and Research (BMBF), the German Research Foundation (DFG) and Vertex Pharmaceuticals paid to the institution; personal fees for participation in advisory boards or consulting from Arrowhead Pharmaceuticals, Boehringer Ingelheim, Enterprise Therapeutics, Kither Biotech, Vertex Pharmaceuticals; lecture fees from Vertex Pharmaceuticals; and travel reimbursement from Boehringer Ingelheim and Vertex Pharmaceuticals.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

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.

References

1. Woods JC, Wild JM, Wielpütz MO, Clancy JP, Hatabu H, Kauczor HU, et al. Current state of the art MRI for the longitudinal assessment of cystic fibrosis. *J Magn Reson Imaging*. (2020) 52:1306–20. doi: 10.1002/jmri.27030

2. Goralski JL, Stewart NJ, Woods JC. Novel imaging techniques for cystic fibrosis lung disease. *Pediatr Pulmonol*. (2021) 56:S40–54. doi: 10.1002/ppul.24931

3. Burgel PR, Bellis G, Olesen HV, Viviani L, Zolin A, Blasi F, et al. Future trends in cystic fibrosis demography in 34 European countries. *Eur Respir J*. (2015) 46:133–41. doi: 10.1183/09031936.00196314

4. Bell SC, Mall MA, Gutierrez H, Macek M, Madge S, Davies JC, et al. The future of cystic fibrosis care: a global perspective. *Lancet Respir Med*. (2020) 8:65–124. doi: 10.1016/S2213-2600(19)30337-6

5. Mall MA, Mayer-Hamblett N, Rowe SM. Cystic fibrosis: emergence of highly effective targeted therapeutics and potential clinical implications. *Am J Respir Crit Care Med*. (2020) 201:1193–208. doi: 10.1164/rccm.201910-1943SO

6. Pearce MS, Salotti JA, Little MP, McHugh K, Lee C, Kim KP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet*. (2012) 380:499–505. doi: 10.1016/S0140-6736(12)60815-0

7. Kuo W, Ciet P, Tiddens HAWM, Zhang W, Guillerman RP, van Straten M. Monitoring cystic fibrosis lung disease by computed tomography. Radiation risk in perspective. *Am J Respir Crit Care Med*. (2014) 189:1328–36. doi: 10.1164/rccm.201311-2099CI

8. Roach DJ, Crémillieux Y, Fleck RJ, Brody AS, Serai SD, Szczesniak RD, et al. Ultrashort echo-time magnetic resonance imaging is a sensitive method for the evaluation of early cystic fibrosis lung disease. *Ann Am Thorac Soc*. (2016) 13:1923–31. doi: 10.1513/AnnalsATS.201603-203OC

9. Grasemann H, Ciet P, Amin R, McDonald N, Klingel M, Tiddens HAWM, et al. Changes in magnetic resonance imaging scores and ventilation inhomogeneity in children with cystic fibrosis pulmonary exacerbations. *Eur Respir J*. (2017) 50:1700244. doi: 10.1183/13993003.00244-2017

10. Stahl M, Wielpütz MO, Graeber SY, Joachim C, Sommerburg O, Kauczor HU, et al. Comparison of lung clearance index and magnetic resonance imaging for assessment of lung disease in children with cystic fibrosis. *Am J Respir Crit Care Med*. (2017) 195:349–59. doi: 10.1164/rccm.201604-0893OC

11. Wielpütz MO, Puderbach M, Kopp-Schneider A, Stahl M, Fritzsching E, Sommerburg O, et al. Magnetic resonance imaging detects changes in structure and perfusion, and response to therapy in early cystic fibrosis lung disease. *Am J Respir Crit Care Med*. (2014) 189:956–65. doi: 10.1164/rccm.201309-1659OC

12. Altes TA, Johnson M, Fidler M, Botfield M, Tustison NJ, Leiva-Salinas C, et al. Use of hyperpolarized helium-3 MRI to assess response to ivacaftor treatment in patients with cystic fibrosis. *J Cyst Fibros*. (2017) 16:267–74. doi: 10.1016/j.jcf.2016.12.004

13. Graeber SY, Renz DM, Stahl M, Pallenberg ST, Sommerburg O, Naehrlich L, et al. Effects of elexacaftor/tezacaftor/ivacaftor therapy on lung clearance index and magnetic resonance imaging in patients with cystic fibrosis and one or two F508del alleles. *Am J Respir Crit Care Med*. (2022) 206:311–20. doi: 10.1164/rccm.202201-0219OC

14. Klimeš F, Voskrebenzev A, Gutberlet M, Speth M, Grimm R, Dohna M, et al. Effect of CFTR modulator therapy with elexacaftor/tezacaftor/ivacaftor on pulmonary ventilation derived by 3D phase-resolved functional lung MRI in cystic fibrosis patients. *Eur Radiol*. (2024) 34:80–9. doi: 10.1007/s00330-023-09912-6

15. Wucherpfennig L, Triphan SMF, Wege S, Kauczor HU, Heussel CP, Schmitt N, et al. Magnetic resonance imaging detects improvements of pulmonary and paranasal sinus abnormalities in response to elexacaftor/tezacaftor/ivacaftor therapy in adults with cystic fibrosis. *J Cyst Fibros*. (2022) 21:1053–60. doi: 10.1016/j.jcf.2022.03.011

16. Streibel C, Willers CC, Pusterla O, Bauman G, Stranzinger E, Brabandt B, et al. Effects of elexacaftor/tezacaftor/ivacaftor therapy in children with cystic fibrosis—a comprehensive assessment using lung clearance index, spirometry, and functional and structural lung MRI. *J Cyst Fibros*. (2023) 22:615–22. doi: 10.1016/j.jcf.2022.12.012

17. Eichinger M, Optazaite DE, Kopp-Schneider A, Hintze C, Biederer J, Niemann A, et al. Morphologic and functional scoring of cystic fibrosis lung disease using MRI. *Eur J Radiol*. (2012) 81:1321–9. doi: 10.1016/j.ejrad.2011.02.045

18. Wielpütz MO, von Stackelberg O, Stahl M, Jobst BJ, Eichinger M, Puderbach MU, et al. Multicentre standardisation of chest MRI as radiation-free outcome measure of lung disease in young children with cystic fibrosis. *J Cyst Fibros*. (2018) 17:518–27. doi: 10.1016/j.jcf.2018.05.003

19. Graeber SY, Vitzthum C, Pallenberg ST, Naehrlich L, Stahl M, Rohrbach A, et al. Effects of elexacaftor/tezacaftor/ivacaftor therapy on CFTR function in patients with cystic fibrosis and one or two *F508del* alleles. *Am J Respir Crit Care Med*. (2022) 205:540–9. doi: 10.1164/rccm.202110-2249OC

20. Stahl M, Roehmel J, Eichinger M, Doellinger F, Naehrlich L, Kopp MV, et al. Effects of lumacaftor/ivacaftor on cystic fibrosis disease progression in children 2 through 5 years of age homozygous for F508del-CFTR: a phase 2 placebo-controlled clinical trial. *Ann Am Thorac Soc*. (2023) 20:1144–55. doi: 10.1513/AnnalsATS.202208-684OC

21. Kanda T, Nakai Y, Oba H, Toyoda K, Kitajima K, Furui S. Gadolinium deposition in the brain. *Magn Reson Imaging*. (2016) 34:1346–50. doi: 10.1016/j.mri.2016.08.024

22. Gulani V, Calamante F, Shellock FG, Kanal E, Reeder SB. Gadolinium deposition in the brain: summary of evidence and recommendations. *Lancet Neurol*. (2017) 16:564–70. doi: 10.1016/S1474-4422(17)30158-8

23. Mathur M, Jones JR, Weinreb JC. Gadolinium deposition and nephrogenic systemic fibrosis: a radiologist's primer. *Radiographics*. (2020) 40:153–62. doi: 10.1148/rg.2020190110

24. Bauman G, Puderbach M, Heimann T, Kopp-Schneider A, Fritzsching E, Mall MA, et al. Validation of Fourier decomposition MRI with dynamic contrast-enhanced MRI using visual and automated scoring of pulmonary perfusion in young cystic fibrosis patients. *Eur J Radiol*. (2013) 82:2371–7. doi: 10.1016/j.ejrad.2013.08.018

25. Schönfeld C, Cebotari S, Voskrebenzev A, Gutberlet M, Hinrichs J, Renne J, et al. Performance of perfusion-weighted Fourier decomposition MRI for detection of chronic pulmonary emboli. *J Magn Reson Imaging*. (2015) 42:72–9. doi: 10.1002/jmri.24764

26. Nyilas S, Bauman G, Sommer G, Stranzinger E, Pusterla O, Frey U, et al. Novel magnetic resonance technique for functional imaging of cystic fibrosis lung disease. *Eur Respir J*. (2017) 50:1701464. doi: 10.1183/13993003.01464-2017

27. Bondesson D, Schneider MJ, Gaass T, Kühn B, Bauman G, Dietrich O, et al. Nonuniform Fourier-decomposition MRI for ventilation- and perfusion-weighted imaging of the lung. *Magn Reson Med*. (2019) 82:1312–21. doi: 10.1002/mrm.27803

28. Triphan SMF, Bauman G, Konietzke P, Konietzke M, Wielpütz MOInternational Workshop for Pulmonary Functional Imaging (IWPFI). Magnetic resonance imaging of lung perfusion. *J Magn Reson Imaging*. (2024) 59:784–96. doi: 10.1002/jmri.28912

29. Voskrebenzev A, Gutberlet M, Klimeš F, Kaireit TF, Schönfeld C, Rotärmel A, et al. Feasibility of quantitative regional ventilation and perfusion mapping with phase-resolved functional lung (PREFUL) MRI in healthy volunteers and COPD, CTEPH, and CF patients. *Magn Reson Med*. (2018) 79:2306–14. doi: 10.1002/mrm.26893

30. Bauman G, Puderbach M, Deimling M, Jellus V, Chefd'hotel C, Dinkel J, et al. Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of Fourier decomposition in proton MRI. *Magn Reson Med*. (2009) 62:656–64. doi: 10.1002/mrm.22031

31. Pusterla O, Heule R, Santini F, Weikert T, Willers C, Andermatt S, et al. MRI lung lobe segmentation in pediatric cystic fibrosis patients using a recurrent neural network

trained with publicly accessible CT datasets. *Magn Reson Med*. (2022) 88:391–405. doi: 10.1002/mrm.29184

32. Bauman G, Bieri O. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. *Magn Reson Med*. (2017) 77:336–42. doi: 10.1002/mrm.26096

33. Munidasa S, Zanette B, Couch M, Grimm R, Seethamraju R, Dumas MP, et al. Inter-and intravisit repeatability of free-breathing MRI in pediatric cystic fibrosis lung disease. *Magn Reson Med*. (2023) 89:2048–61. doi: 10.1002/mrm.29566

34. Behrendt L, Smith LJ, Voskrebenzev A, Klimeš F, Kaireit TF, Pöhler GH, et al. A dual center and dual vendor comparison study of automated perfusion-weighted phase-resolved functional lung magnetic resonance imaging with dynamic contrast-enhanced magnetic resonance imaging in patients with cystic fibrosis. *Pulm Circ*. (2022) 12:e12054. doi: 10.1002/pul2.12054

35. Bauman G, Scholz A, Rivoire J, Terekhov M, Friedrich J, De Oliveira A, et al. Lung ventilation-and perfusion-weighted Fourier decomposition magnetic resonance imaging: *in vivo* validation with hyperpolarized [3]He and dynamic contrast-enhanced MRI. *Magn Reson Med*. (2013) 69:229–37. doi: 10.1002/mrm.24236

36. Couch MJ, Munidasa S, Rayment JH, Voskrebenzev A, Seethamraju RT, Vogel-Claussen J, et al. Comparison of functional free-breathing pulmonary ^{1}H and hyperpolarized [129]Xe magnetic resonance imaging in pediatric cystic fibrosis. *Acad Radiol*. (2021) 28:e209–18. doi: 10.1016/j.acra.2020.05.008

37. Rayment JH, Couch MJ, McDonald N, Kanhere N, Manson D, Santyr G, et al. Hyperpolarised [129]Xe magnetic resonance imaging to monitor treatment response in children with cystic fibrosis. *Eur Respir J*. (2019) 53:1802188. doi: 10.1183/13993003.02188-2018

38. Stahl M, Steinke E, Graeber SY, Joachim C, Seitz C, Kauczor HU, et al. Magnetic resonance imaging detects progression of lung disease and impact of newborn screening in preschool children with cystic fibrosis. *Am J Respir Crit Care Med*. (2021) 204:943–53. doi: 10.1164/rccm.202102-0278OC

39. Bauman G, Pusterla O, Bieri O. Ultra-fast steady-state free precession pulse sequence for Fourier decomposition pulmonary MRI. *Magn Reson Med*. (2016) 75:1647–53. doi: 10.1002/mrm.25697

40. Sandkühler R, Jud C, Pezold S, Cattin PC. Adaptive graph diffusion regularisation for discontinuity preserving image registration In: S Klein, M Staring, S Durrleman and S Sommer, editors. *Biomedical image registration. WBIR 2018. Lecture notes in computer science*. Springer: Cham (2018). 24–34.

41. Mccormick M, Liu X, Jomier J, Marion C, Ibanez L. ITK: enabling reproducible research and open science. *Front Neuroinform*. (2014) 8:13. doi: 10.3389/fninf.2014.00013

42. Miller MR, Crapo R, Hankinson J, Brusasco V, Burgos F, Casaburi R, et al. General considerations for lung function testing. *Eur Respir J*. (2005) 26:153–61. doi: 10.1183/09031936.05.00034505

43. Graham BL, Steenbruggen I, Barjaktarevic IZ, Cooper BG, Hall GL, Hallstrand TS, et al. Standardization of spirometry 2019 update. An official American Thoracic Society and European Respiratory Society technical statement. *Am J Respir Crit Care Med*. (2019) 200:E70–88. doi: 10.1164/rccm.201908-1590ST

44. GLI Lung Function Calculator. Available at: https://gli-calculator.ersnet.org/index.html

45. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3–95-years age range: the global lung function 2012 equations. *Eur Respir J*. (2012) 40:1324–43. doi: 10.1183/09031936.00080312

46. Karlik SJ. Fundamentals of clinical research for radiologists: exploring and summarizing radiologic data. *Am J Roentgenol*. (2003) 180:47–54. doi: 10.2214/ajr.180.1.1800047

47. Rowntree D. *Statistics without tears: an introduction for non-mathematicians*. London, Great Britain: Penguin Books Ltd. (1981). 170 p.

48. Schiwek M, Triphan SMF, Biederer J, Weinheimer O, Eichinger M, Vogelmeier CF, et al. Quantification of pulmonary perfusion abnormalities using DCE-MRI in COPD: comparison with quantitative CT and pulmonary function. *Eur Radiol*. (2022) 32:1879–90. doi: 10.1007/s00330-021-08229-6

49. Nyilas S, Bauman G, Pusterla O, Ramsey K, Singer F, Stranzinger E, et al. Ventilation and perfusion assessed by functional MRI in children with CF: reproducibility in comparison to lung function. *J Cyst Fibros*. (2019) 18:543–50. doi: 10.1016/j.jcf.2018.10.003

50. Behrendt L, Voskrebenzev A, Klimeš F, Gutberlet M, Winther HB, Kaireit TF, et al. Validation of automated perfusion-weighted phase-resolved functional lung (PREFUL)-MRI in patients with pulmonary diseases. *J Magn Reson Imaging*. (2020) 52:103–14. doi: 10.1002/jmri.27027

51. Klimeš F, Voskrebenzev A, Gutberlet M, Kern AL, Behrendt L, Grimm R, et al. 3D phase-resolved functional lung ventilation MR imaging in healthy volunteers and patients with chronic pulmonary disease. *Magn Reson Med*. (2021) 85:912–25. doi: 10.1002/mrm.28482

52. Klimeš F, Voskrebenzev A, Gutberlet M, Obert AJ, Pöhler GH, Grimm R, et al. Repeatability of dynamic 3D phase-resolved functional lung (PREFUL) ventilation MR imaging in patients with chronic obstructive pulmonary disease and healthy volunteers. *J Magn Reson Imaging*. (2021) 54:618–29. doi: 10.1002/jmri.27543

53. Couch MJ, Morgado F, Kanhere N, Kowalik K, Rayment JH, Ratjen F, et al. Assessing the feasibility of hyperpolarized [129]Xe multiple-breath washout MRI in pediatric cystic fibrosis. *Magn Reson Med*. (2020) 84:304–11. doi: 10.1002/mrm.28099

54. Zha W, Nagle SK, Cadman RV, Schiebler ML, Fain SB. Three-dimensional isotropic functional imaging of cystic fibrosis using oxygen-enhanced MRI: comparison with hyperpolarized [3]He MRI. *Radiology*. (2019) 290:229–37. doi: 10.1148/radiol.2018181148

55. Chung SH, Huynh KM, Goralski JL, Chen Y, Yap PT, Ceppe AS, et al. Feasibility of free-breathing ^{19}F MRI image acquisition to characterize ventilation defects in CF and healthy volunteers at wash-in. *Magn Reson Med*. (2023) 90:79–89. doi: 10.1002/mrm.29630

56. Marshall H, Voskrebenzev A, Smith LJ, Biancardi AM, Kern AL, Collier GJ, et al. [129]Xe and free-breathing ^{1}H ventilation MRI in patients with cystic fibrosis: a dual-Center study. *J Magn Reson Imaging*. (2023) 57:1908–21. doi: 10.1002/jmri.28470

57. Bauman G, Pusterla O, Bieri O. Functional lung imaging with transient spoiled gradient echo. *Magn Reson Med*. (2019) 81:1915–23. doi: 10.1002/mrm.27535

58. Dohna M, Voskrebenzev A, Klimeš F, Kaireit TF, Glandorf J, Pallenberg ST, et al. PREFUL MRI for monitoring perfusion and ventilation changes after elexacaftor-tezacaftor-ivacaftor therapy for cystic fibrosis: a feasibility study. *Radiol Cardiothorac Imaging*. (2024) 6:e230104. doi: 10.1148/ryct.230104

A synthetic lung model (ASYLUM) for validation of functional lung imaging methods shows significant differences between signal-based and deformation-field-based ventilation measurements

Andreas Voskrebenzev[1,2]*, Marcel Gutberlet[1,2],
Filip Klimeš[1,2], Till F. Kaireit[1,2], Hoen-oh Shin[1,2],
Hans-Ulrich Kauczor[3,4], Tobias Welte[2,5], Frank Wacker[1,2] and
Jens Vogel-Claussen[1,2]

[1]Institute for Diagnostic and Interventional Radiology, Hannover Medical School, Hannover, Germany,
[2]Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Member of the
German Center for Lung Research, Hannover, Germany, [3]Department of Diagnostic and
Interventional Radiology, University Hospital of Heidelberg, Heidelberg, Germany, [4]Translational Lung
Research Center Heidelberg (TLRC), Member of the German Lung Research Center (DZL), Heidelberg,
Germany, [5]Clinic of Pneumology, Hannover Medical School, Hannover, Germany

Introduction: Validation of functional free-breathing MRI involves a comparison to more established or more direct measurements. This procedure is cost-intensive, as it requires access to patient cohorts, lengthy protocols, expenses for consumables, and binds working time. Therefore, the purpose of this study is to introduce a synthetic lung model (ASYLUM), which mimics dynamic MRI acquisition and includes predefined lung abnormalities for an alternative validation approach. The model is evaluated with different registration and quantification methods and compared with real data.

Methods: A combination of trigonometric functions, deformation fields, and signal combinations were used to create 20 synthetic image time series. Lung voxels were assigned either to normal or one of six abnormality classes. The images were registered with three registration algorithms. The registered images were further analyzed with three quantification methods: deformation-based or signal-based regional ventilation (JVent/RVent) analysis and perfusion amplitude (QA). The registration results were compared with predefined deformations. Quantification methods were evaluated regarding predefined amplitudes and with respect to sensitivity, specificity, and spatial overlap of defects. In addition, 36 patients with chronic obstructive pulmonary disease were included for verification of model interpretations using CT as the gold standard.

Results: One registration method showed considerably lower quality results (76% correlation vs. 92/97%, $p \leq 0.0001$). Most ventilation defects were correctly detected with RVent and QA (e.g., one registration variant with sensitivity $\geq 78\%$, specificity ≥ 88). Contrary to this, JVent showed very low sensitivity for lower lung quadrants (0–16%) and also very low specificity (1–29%) for upper lung quadrants. Similar patterns of defect detection differences between RVent and JVent were also observable in patient data: Firstly, RVent was more aligned with

CT than JVent for all quadrants ($p \leq 0.01$) except for one registration variant in the lower left region. Secondly, stronger differences in overlap were observed for the upper quadrants, suggesting a defect bias in the JVent measurements in the upper lung regions.

Conclusion: The feasibility of a validation framework for free-breathing functional lung imaging using synthetic time series was demonstrated. Evaluating different ventilation measurements, important differences were detected in synthetic and real data, with signal-based regional ventilation assessment being a more reliable method in the investigated setting.

1 Introduction

Functional proton lung MRI gained interest during the last few years, as it allows to assess lung function in free-breathing without ionizing radiation, inhalation of gases, or contrast agent application (1). These methods can be divided into two basic groups: signal-based and deformation-based. While both approaches require image registration to account for diaphragmatic and cardiac motion, the first derives ventilation and perfusion parameters from the signal values, and the latter uses the geometric information to derive ventilation based on the calculated expansion/shrinkage of voxels during registration.

Currently, the signal-based approach is more widespread, including Fourier decomposition (2), matrix pencil decomposition (3), self-gated non-contrast-enhanced functional lung (SENCEFUL) (4, 5), and phase-resolved functional lung (PREFUL) MRI (6, 7). Nevertheless, some studies showed promising results in 2D and 3D with the deformation-field-based approach (8–11). Since both approaches are surrogates, in addition to mandatory reproducibility measurements (12–15), it is also necessary to perform extensive validation with well-established reference methods such as inhaled gas MRI for ventilation or dynamic contrast-enhanced MRI for perfusion (16–20). This process requires access to patient cohorts, additional hardware (hyperpolarization, multinuclear RF coils), consumables, extended MRI protocols, manpower, and costly scanner time. Furthermore, it is not expected that inhaled gas MRI ventilation measurement performed in breath-hold will fully coincide with a free-breathing method purely due to physiologic reasons (21). Similarly, while correlations were observed in DCE and perfusion-weighted signal-based measurements, it is important to note that these methods rely on fundamentally different mechanisms. DCE measures signal changes due to T1 shortening, which is induced by the transport pathway of contrast media concentration. In contrast, signal-based methods measure the time-of-flight effect of unsaturated blood spins entering the slice. Additionally, patient compliance might limit the practicability in certain cases and might lead to increased variability (22).

This motivates us to approach the problem from a different perspective. Similar to conducting phantom measurements during sequence design, a synthetic lung model (ASYLUM) may be used as a complementary validation tool for post-processing algorithms. By applying known regional deformation and using model functions for signal time-series generation, different lung function states can be simulated and made directly available as a gold standard.

Therefore, the aim of this study is to describe a method to create ASYLUM and utilize this model by evaluating different registration and lung function quantification methods. Although it is possible to evaluate registrations by comparing structure alignment and other image similarity metrics (23), ASYLUM includes a known deformation for direct registration evaluation. The signal-based regional ventilation (RVent) (24) and deformation-field-based Jacobian determinant (JVent) (8) were selected to test whether the theoretical equivalence of both methods can be verified. Additionally, a simple implementation of a signal-based perfusion-weighted amplitude (QA) measurement was included to account for the perfusion abnormalities, which are also modeled in ASYLUM. Finally, all the described ventilation methods were evaluated in a patient cohort to see if the findings were also confirmed by real data using CT as a gold standard.

2 Theoretical considerations

2.1 Basic theory of signal-based pulmonary functional MRI

As described previously (25), the pulmonary time series of MR signal acquired with a fast spoiled gradient echo sequence during free-breathing can be described by four main components: proton density, time-of-flight (TOF) effect, movement, and noise. This is an incomplete model (e.g., T1, T2/T2* decay, diffusion, field inhomogeneities, direction of flow, artifacts), but previous validation studies (16–19, 26) suggest that these factors probably only play a minor role with proper imaging protocols and reconstructions (e.g., minimal TR in combination with low flip angles to achieve proton density (PD) weighting, asymmetric echo to reduce TE and hence T2* decay and further propel PD weighting, T-GRAPPA to exploit the dynamic nature of the acquisition, and apply surface coil intensity correction to avoid artificial regional signal variations). Therefore, for simplicity, these influences are omitted from the following description.

The variability due to movement is compensated by image registration, which will be discussed in the next section. During expiration, the decreasing lung volume leads to an increased proton

density and therefore increased MR signal, and vice versa during inspiration. Therefore, proton density is a function of respiration, and the inverse relationship of signal and volume is the foundation for ventilation measurements (27).

The continuous application of excitation pulses over a short period of time leads to incomplete relaxation, which finally converges to a steady state. Inflowing spins have not reached this state and have a higher initial signal in comparison with stationary spins. Thus, the pulsatile inflow of blood leads to further signal variation, also known as TOF, and is the foundation for perfusion measurement.

Due to the inherent low signal of the lung (28), averaging and filtering are required to achieve a sufficient signal-to-noise ratio (SNR). Ideally, after registration, averaging, and filtering, the variation of MR signal contains only information about ventilation and perfusion. Since these variations occur at different frequencies, they can be individually evaluated by Fourier analysis (29).

Using the inverse relationship, a relative volume change can be quantified as RVent (24), which is a generalized version of fractional and specific ventilation accounting for the registered volume:

$$RVent = \frac{v_{Insp} - v_{exp}}{v_{Reg}} = \frac{1/s_{Insp} - 1/s_{exp}}{1/s_{Reg}} = \frac{s_{Reg}\left(s_{exp} - s_{Insp}\right)}{s_{Insp}s_{exp}} \quad (1)$$

with inspiration (Insp), expiration (Exp), and registration (Reg) volumes (v) and signals (s). The registration volume is the final volume after registration, also denoted as a fixed image.

Quantification of perfusion is more complex as it involves the estimation of blood and exchange fractions (30, 31). Since the main concern of this study is the accurate detection of ventilation and perfusion defects, only a perfusion-weighted measurement was implemented in this study by estimating the amplitude of the TOF signal variation component.

The practical implementation of the described methods is described in a later section.

2.2 Basic theory of pulmonary deformation-based ventilation measurement

Although the registration process cannot be completely independent of signal values, which constitute the respective image, the ventilation result of deformation-based ventilation measurements is ultimately derived from geometric information. The deformation field includes the displacement vectors to map the voxels from the moving volume to the fixed volume.

The partial derivatives of this deformation field will reveal local expansion and deflation. For intuitive comprehension, the following extreme cases can be considered: If all voxels can be described by an arbitrary but constant displacement, no local expansion or inflation is present, and the derivatives will be zero. If instead only one voxel is displaced (or "copied") by one voxel, the partial derivative in this direction at this location will be 100% and corresponds rightly to an expansion factor of 2, since the voxel expanded its size from one voxel to two voxels.

Typically, the total displacement f is described as the sum of the identity and registration displacement operation. Then, the total

expansion/deflation calculation is mathematically described as the determinant of the Jacobi matrix. The Jacobi matrix includes all partial derivatives of the total displacement f, and the determinant can be interpreted as the area/volume parallelogram. Thus, the ventilation with the Jacobi method (JVent) can be calculated as follows:

$$JVent = \frac{v_{Insp} - v_{exp}}{v_{Reg}} = \det\left(J_f\right) - 1 \quad (2)$$

3 Materials and methods

3.1 Subjects

To validate the derived results from the model data, a subset of data from the COSYCONET study (32), assessing patients with chronic obstructive pulmonary disease (COPD), was carried out retrospectively [n total = 36, n female = 21, median age 63.5, age range 42–77, GOLD I (n = 10), II (n = 9), III (n = 12), IV (n = 5)].

The study inclusion criteria were as follows: COPD- or chronic bronchitis-diagnosed patients aged 40 or older years with availability for repeated study visits over at least 18 months.

The study exclusion criteria were as follows: major lung surgery, moderate or severe exacerbation within the last 4 weeks, presence of lung tumor, inability to walk or understand the intention of the project.

Baseline scans conducted between 16 February 2011 and 4 December 2013 at the Hannover site were included in this subset, comprising completed CT and MRI scans. COSYCONET was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice Guidelines. Approval was obtained from the ethics committees of the participating centers and from the concerned data security authority. All participants provided written informed consent.

3.2 Imaging procedure

3.2.1 MRI

Acquisition was performed on a 1.5 T system (Avanto, Siemens Healthineers, Erlangen, Germany) using a spoiled gradient echo sequence with the following settings provided as median with range (in cases of variability) in parenthesis: field-of-view (FOV) 500×500 mm^2, echo time (TE) 0.82 (0.65–0.82) ms, repetition time (TR) 3 ms, temporal resolution 288 (192–288) ms, slice thickness 15 mm, matrix 128×96 (128–64–128-128), parallel imaging acceleration factor with T-GRAPPA (33) 1 (1–2), receiver bandwidth (BW) 1,500 Hz/Pixel for a total of 200 (200–250) image frames for each slice. A total of four coronal slices were acquired in free-breathing for each patient, with the reference slice positioned at the tracheal bifurcation and two slices in the posterior and one in the anterior direction. The inter-slice distance was set to 3 mm (20%).

3.2.2 CT

Patients were scanned in supine position with a 64-slice scanner without intravenous contrast media with the following settings: tube current 120 kV, automatic tube modulation, table feed 39.375 mm/

gantry rotation, 0.625 mm slice thickness, and 0.7 mm reconstruction interval using a "standard" reconstruction kernel.

Scans were performed in full in- and expiration to capture the lung in total lung capacity and residual volume state.

3.3 ASYLUM—procedure outline

The creation of ASYLUM is a 3-fold procedure:

1. Create a base geometry, which is used as a starting point, and assign voxel classes to reflect different tissues present in the image (e.g., normal and abnormal lung voxels).
2. Change the geometry according to local expansion factors to reflect local expansion and deflation.
3. Fill the geometry with appropriate signals and noise, which reflect MR physics and are in concordance with the geometry change and voxel class defined in the first and second steps.

Steps 2 and 3 can be repeated with different expansion factors to create a time series of images. By changing the classes in step 1, variations of the model can be created. These three steps are described in detail in the following sections and are illustrated in Figure 1.

To reduce inaccuracies, especially during step 2, the calculations were performed on 4-fold upscaled data. Afterwards, the data were scaled back to original size. Before further processing, to mimic a typical PREFUL or Fourier decomposition acquisition, the data were once again interpolated to 256×256 matrix (factor 2).

3.4 ASYLUM—basic model and voxel classes

The initial 2D coronal lung morphology in expiration was constructed by three rectangles in a 128×128 matrix. This defines the three basic classes (1): the body (outer border 90×90) (2), the right and left lung (16×32 each), and (3) the remaining voxels as background (pure noise). Additionally, between both lungs, a vessel was defined as fourth class using a disk geometry with a radius of two voxels.

The lung class is further classified into eight subclasses:

1. Normally ventilated and perfused voxels (VV|QV) with expansion factor e
2. Ventilation and perfusion defect at constant (independent of respiration phase) signal corresponding to the halved value of normal lung parenchyma during inspiration phase (VD|QD$_{Low}$) with $e=0$ (e.g., emphysema)
3. Ventilation and perfusion defect with constant (independent of respiration phase) signal with a value corresponding to the value of normal lung parenchyma during inspiration phase (VD|QD$_{Insp}$) with $e=0$ (e.g., air trapping)
4. Ventilation and perfusion defect with constant (independent of respiration phase) signal with a value corresponding to the 5-fold value of normal lung parenchyma during expiration level (VD|QD$_{High}$) with $e=0$ (e.g., infiltrate)
5. Delayed ventilation and normally perfused voxels (VV$_{Delay}$|QV) (e.g., early disease manifestation) with expansion factor e
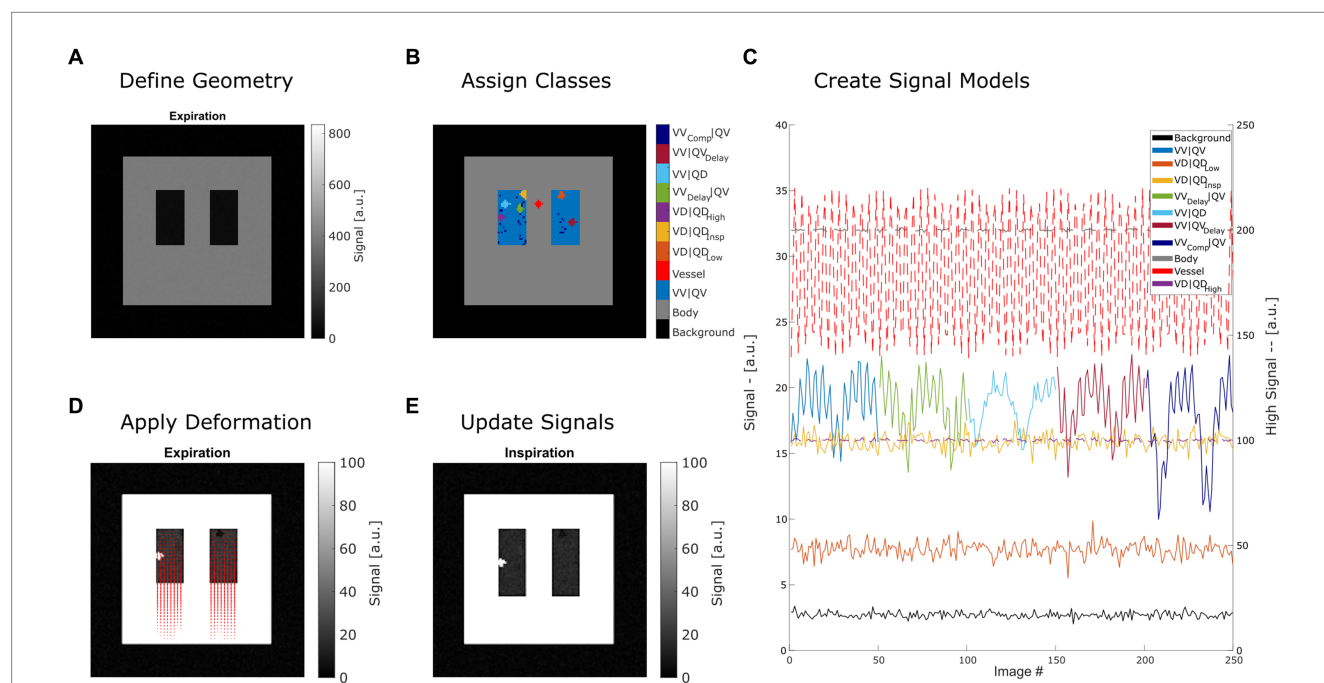
FIGURE 1
Summary of creating a synthetic lung model (ASYLUM) to mimic free-breathing MRI data. The steps consist of defining a simplified coronal lung anatomy **(A)**, random assigning of classes **(B)** for each lung voxel, application of deformation according to respiratory state and local expansion **(D)**, and thus creating an image time series **(E)** with respective time series **(C)**. Please note that for display purposes, some classes were displayed with truncated time series in **(C)**, and two axes were used (dashed line indicates that the high signal axis on the right was used). VV|QV: ventilated and perfused volume, VD|QD$_{Low}$: ventilation and perfusion defect below inspiration signal level, VD|QD$_{Insp}$: ventilation and perfusion defect at inspiration level, VD|QD$_{High}$: ventilation and perfusion defect at high signal level, VV$_{Delay}$|QV: ventilated and perfused volume with delayed ventilation, VV|QD: ventilated volume with perfusion defect, VV|QV$_{Delay}$: ventilated and perfused volume with delayed perfusion.

6. Ventilated perfusion defect (VV|QD) (e.g., V/Q mismatch) with expansion factor e

7. Voxel with normal ventilation but delayed perfusion (VV|QV$_{Delay}$) (e.g., early disease manifestation) with expansion factor e

8. Higher ventilated (compensatory) and normally perfused voxels (VV$_{Comp}$|QV) with expansion factor $e_{compensate}$.

The subclassification is performed using a randomized seed placement, which will result in different cluster locations for each class for repeated model generation. The size of each defect class was set to 60 voxels, equivalent to 6% of the expiratory lung volume (1024).

3.5 ASYLUM—warping

Since ASYLUM must reflect an acquisition in free-breathing, one main challenge is the modeling of respiratory movement. This can be achieved by inverting the registration problem, which tries to compensate movement. Therefore, by definition of adequate deformation, movement can be simulated. For simplicity, the respiratory movement was assumed to be one-dimensional (y-direction). This movement can be mathematically described by a deformation field $F_{x,y}$, which includes the amount of y-displacement for each voxel. To find this deformation field, an auxiliary matrix E containing all local expansion factors e was created according to the assigned lung classes (see previous section). This will initially result in unbalanced expansions, since some columns (y-direction) will include ventilation defects with no expansion. In addition, lung expansion must be compensated by the shrinking of the body class. Therefore, lung compensation voxels and body voxels were assigned with expansion factors to meet the following conditions:

$$C_1 = \frac{1}{32} \sum_{y=\text{Lung Start}}^{\text{Lung Start}+31} E_{x,y} = e \ \forall x \qquad (3)$$

$$C_2 = \sum_{y=0}^{127} E_{x,y} = 0 \ \forall x$$

The first condition, C_1, will ensure that the average lung expansion factor e is uniform (1.25) along all columns of the lung. The second condition, C_2, will establish adequate deflation of the body so that the overall expansion factor is one. Please note that 32 is the lung size in y-direction defined previously.

The deformation field in expiration space (forward deformation $F_{x,y}$) is obtained by applying the cumulative sum operation on the $E_{x,y}$ matrix in y-direction. Since warping is typically performed with inverse deformation fields (deformed or fixed space) to avoid problems of "holes" and both fields will be required anyway to simulate a "perfect registration," the inverse deformation field was calculated by performing a linear scattered interpolation on the irregular transformed grid (*i,j*) defined by the forward deformation:

$$\text{i} = \text{x} + 0 = x$$

$$j = y + E_{x,y}$$

Please note that there is no deformation in *x*-direction. Finally, the scattered interpolant S was evaluated at a regular grid in deformed space and reversed in direction to obtain the inverse deformation field $I_{x,y}$:

$$I_{x,y} = -S(x,y)$$

Thus, to obtain an image geometry $G(e)$ according to an expansion factor *e*, the expiration image is warped with $I_{x,y}$. A pseudo registration of this image would be performed by a consecutive image warping ($\bullet$) with the forward deformation $F_{x,y}$, which will in theory give the initial geometry of the expiration image. Therefore, all required deformation operations can be summarized as follows:

$$G(e) = G(1.0) \bullet I_{x,y}$$

$$G(1.0) = G(1.0) \bullet I_{x,y} \bullet F_{x,y} = G(e) \bullet F_{x,y}$$

3.6 ASYLUM—signal time-series generation

To fill the geometry with different signal time series $s(t)$, the Lujan et al. (34) formula, as modified by Bauman et al. (2) was used:

$$s(t) = s_0 - A_V \cos^{2n}\left(\frac{\pi t}{\tau_V} - \varphi_V\right) + A_Q \sin^{2m}\left(\frac{\pi t}{\tau_Q} - \varphi_Q\right),$$

with the ventilation (*V*) and perfusion (*Q*) signal amplitudes *A*, periods τ and phases φ and signal shape parameters *n* and *m*. The following parameters were set according to the provided example of Bauman et al.: $n = 3, m = 2, \tau_V = 5\,s, \tau_Q = 0.8$. The remaining parameters were adapted to fit the respective voxel class with arbitrarily set base signal levels $s_{\exp} = 20\,AU, s_{Body} = 200\,AU$:

1. VV|QV:
   $e = 0.25; s_0 = s_{\exp}; A_Q = 6; A_V = s_{\exp} \bullet (1 - 1/(1+e));$
   $\varphi_V = 0; \varphi_Q = 0$; the
   last two settings indicating no delay in ventilation and perfusion

2. VD|QD$_{Low}$: The same as first class, but with perfusion amplitude $A_Q = 0$, expansion factor $e = 0$ and $s_0 = s_{Insp}/2$

3. VD|QD$_{Insp}$: The same as second class, but with $s_0 = s_{\exp}/(1 + e_{Ventilated}) = s_{\exp}/1.25 = 16\,AU$

4. VD|QD$_{High}$: The same as second class, but with $s_0 = s_{\exp} \bullet 5 = 100\,AU$

5. $VV_{Delay}|QV$: The same as first class, but with $\varphi_V = \pi / 2$

6. $VV|QD$: The same as first class, but with $A_Q = 0$

7. $VV|QV_{Delay}$: The same as first class, but with $\varphi_Q = \pi / 2$

8. $VV_{Comp}|QV$: The same as first class, but with increased expansion factor $e_{Compensate}$ as defined in Equation 3.

The body and background class were assigned with a fixed signal as they contained no dynamic component. The vessel class was assigned with the following parameters: $s_0 = s_{Body} \cdot 0.9; A_Q = 120; A_V = 0; \varphi_V = 0; \varphi_Q = \pi / 2$.

Finally, an artificial sum-of-squares coil combination was performed to simulate MR signals with adequate noise distributions (35). For this purpose, the normally distributed noise of a pseudo-four-channel coil was added to the signal $s(t)$ with a standard deviation set in such a way, that VV voxels ended up with a SNR of 5.

To produce a time series of images, the expansion factor was modulated according to the ventilation phase during the specific time point. For this purpose, the inverse of the signal time series was used as a surrogate for volume $v(t,x)$ (27), from which the respiration factor $r(t,x)$ was derived to scale the expansion factor according to volume:

$$v(t,x) = 1 / \left(s_0 - A_V \cos^{2n} \left(\frac{\pi t}{\tau_V(x)} - \varphi_V(x) \right) \right)$$

$$e(t,x) = e(x) \cdot r(t,x) = e(x) \cdot \frac{v(t,x) - \min(v(t,x))}{\max(v(t,x)) - \min(v(t,x))}$$

Please note that the respiration factor $r(t,x)$ is used to scale $e(t,x)$ between 0 and $e(x)$ according to the respective volume during the time point t.

For the final creation of an image time series, deformation was performed according to $e(t,x)$, and the geometry was filled as described in this section.

The time was parametrized with 250 entries using an increment of 0.192 s (temporal resolution).

3.7 Registration and ventilation imaging

Non-rigid motion correction was performed with four methods with a one-step registration using the expiration state as a fixed image:

1. Reference registration (REF) by using the forward deformation field as described in the ASYLUM-warping section. The only error source in this case is the transformation between forward and inverse deformation fields.
2. Advanced normalization tools (ANTs) (36), which were successfully used in multiple previous studies (14, 37, 38). The b-spline model with cross-correlation as a similarity metric was used. Denoted as ANTs in the following.
3. The Forsberg registration package (39, 40) with polynomial expansion method and normalized cross-correlation and mean square error as metric, which was recently shown to deliver similar

quality to ANTs in the context of 3D PREFUL paired with a potential speed advantage (41). Denoted as F-REG in the following.
4. A diffeomorphic demons algorithm (42) with accumulated field smoothing is implemented as an official function "imregdemons" in MATLAB (R2020b). Denoted as M-REG in the following.

For the patient data, only algorithms 2–4 were feasible.

Prior to registration, the respiration factor $r(t)$ of the normal voxels was binned into 10 groups, each containing 10% of the data (25 images). Registration was performed toward the average expiration image.

The registered expiration and inspiration images (expiration and inspiration states according to VV|QV class) were averaged, and RVent was calculated according to Equation 1.

The deformation fields of the inspiration were averaged and the JVent was calculated according to Equation 2.

Due to low SNR or reduced functional signal amplitude, for the patient data, an image-guided edge-preserving filter was applied prior to RVent calculation (43) (the filtered result is denoted as RVent*). Otherwise, the processing was the same as for the ASYLUM data.

3.8 ASYLUM perfusion imaging

To assess the perfusion-related aspects of ASYLUM, perfusion-weighted analysis was performed as follows:

1. The registered image time series was high-pass filtered with a cutoff frequency of 0.9 Hz.
2. Phases with maximal and minimal signals were determined from the averaged signal in the lung ROI.
3. Perfusion-related amplitude (QA) was calculated as the difference between the maximal and minimal signal in lung parenchyma: $QA = s_{max} - s_{min}$.

3.9 Ventilation and perfusion defect (VD/QD) analysis

As published in recent studies (14, 44), the VD and QD were defined as regions with values below a threshold determined from the respective 90th-percentile value multiplied by 0.4. VD/QD percentage (VDP/QDP) was defined as the number of voxels below the threshold in relation to the total number of voxels in the lung ROI.

3.10 Parametric response mapping

CT data were analyzed with parametric response mapping (PRM) (45). For this end, registration of inspiration toward expiration was performed as follows:

1. Semiautomatic segmentation of lung lobes by applying a local, adaptive region growing algorithm in inspiration and expiration with a dedicated software (MeVisPULMO 3D, Fraunhofer MEVIS Bremen, Germany).
2. Downsampling the image resolution in x- and y-direction by a factor of 2.

3. Non-rigid registration with ANTs.
4. Labeling of lung voxels as normal, voxels with functional small airways disease (fSAD) or emphysema according to Galban et al. (45).

To allow a regional comparison with MRI, appropriate slice positions were identified. To find an adequate transformation between the coordinates of MR and CT measurements, a landmark (tracheal bifurcation) was manually identified on the anatomical 3D MRI scan (used as a localizer) and CT. Using transformed coordinates, the coronal slice locations were identified in the CT data and averaged to produce corresponding 15-mm slices, as in MR. Then CT was registered toward MRI expiration state using only mask information of CT and MRI with non-rigid registration by ANTs. The obtained transformation was subsequently applied to PRM map. The three-class PRM map was simplified to a binary VD map by combining fSAD and emphysema class into one VD class.

3.11 Image and statistical analysis

Raw registration performance of ANTs, F-REG, and M-REG on ASYLUM data was assessed with three quantitative metrics using REF as a reference:

1. Root mean squared relative error (RMSRE) of the inspiration state registered to the expiration state calculated inside the lung ROI of the expiration state:

$$RMSRE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{REG - REF}{REF}\right|^2},$$

With n being the number of samples in the lung ROI, REG the registered inspiration image with the respective registration method (46).

2. Pearson correlation of the y-profile of the respective registration to the y-profile of the reference
3. Registration time.

To obtain regional information for certain statistics, the whole lung ROI was divided into the following equally sized quadrants: upper right (UR), upper left (UL), lower right (LR), and lower left (LL).

Averaged whole lung ROI values of RVent, JVent, and QA were used to calculate median and interquartile range values for all registration methods of the repeated experiment and visualized as boxplots. For regional comparison, the mean and median of the relative difference to the defined functional value were determined quadrant-wise.

Quadrant-wise true positive (sensitivity) and true negative rate (specificity) were determined for the binary measurements VD and QD based on RVent, JVent, and QA. Sensitivity was also evaluated for the individual defect classes. For this purpose, the quadrant ROIs were modified by additional masking of the respective non-involved defect classes. This was necessary, as the VD and QD did not further differentiate between different kinds of defects, and therefore an evaluation of the whole ROI would lead to a bias in the measurement.

For the comparison of patient data, median and interquartile range values were determined for averaged lung ROI values. Additionally, VD and total overlap derived from RVent and JVent in relation to CT VD were compared quadrant-wise.

Statistical differences between registration and ventilation measurement methods were tested with Friedman's test as omnibus test. For *post-hoc* Wilcoxon signed rank test was performed at the 5% significance level.

4 Results

4.1 ASYLUM data

Dynamic data were successfully created as described in the Methods section (see Figure 2 for examples of time frames obtained with ASYLUM).

4.1.1 Registration performance

Significant but minor differences were found for the RMSRE metric for ANTs and the other registration algorithms (ANTs: 0.30, F-REG 0.27, and M-REG 0.27, $p \leq 0.05$). Correlation differences of the y-displacement profile were more pronounced, with F-REG showing the best result at 97% vs. 92% (ANTs) and 77% (M-REG), $p \leq 0.0001$. Similarly, the required registration time was significantly different between the algorithms: ranging from 150 min (ANTs) to 20 min (M-REG) and 7 min (F-REG). See Table 1, Part A, for a summary of all results.

Visual analysis (see example in Figure 3) confirmed the correlation results: M-REG showed clearly wrong displacement patterns and JVent maps, whereas ANTs and F-REG resulted in what seemed to be blurred versions of the REF displacement with an artificial increase of deformation from cranial to caudal as evident from the JVent maps.

4.1.2 Functional parameters

Although significantly different, REF and F-REG resulted in whole lung mean RVent and JVent values approximating the expected value of 0.25 within a margin of 0.02 or less for the values within IQR. Both REF and ANTs were nearly at the expected four arbitrary units (a.u.) mark for QA: 3.99 (REF) and 3.94 (ANTs). Contrary to this, F-REG and M-REG were off by a significantly higher margin: 3.26 (F-REG) and 3.60 (M-REG). See Table 1, Part B, for a summary of all results.

Visual inspection (see example in Figure 3) showed that increased RVent values in case of ANTs were mainly located at the lung boundary within the lung ROI. Similarly, M-REG resulted in increased JVent values at the boundary. In concordance with the previously described observations, the regional analysis (see Figure 4) shows a clear gradient in the JVent results for all registration results besides REF, manifesting as a negative (underestimation) relative difference for upper quadrants (e.g., −86% for ANTs UR) and positive (overestimation) for lower quadrants (e.g., 36% for F-REG LR). Parameters RVent and QA mainly slightly underestimated the functional parameters (e.g., −11% for F-REG UR and −2% for QA LR), with the exception of the ANTs mean values results, which showed high overestimation for lower quadrants (e.g., 116% for LR). This specific pattern disappeared for median values, while all other patterns prevailed. See Table 2 for the summary of all results.
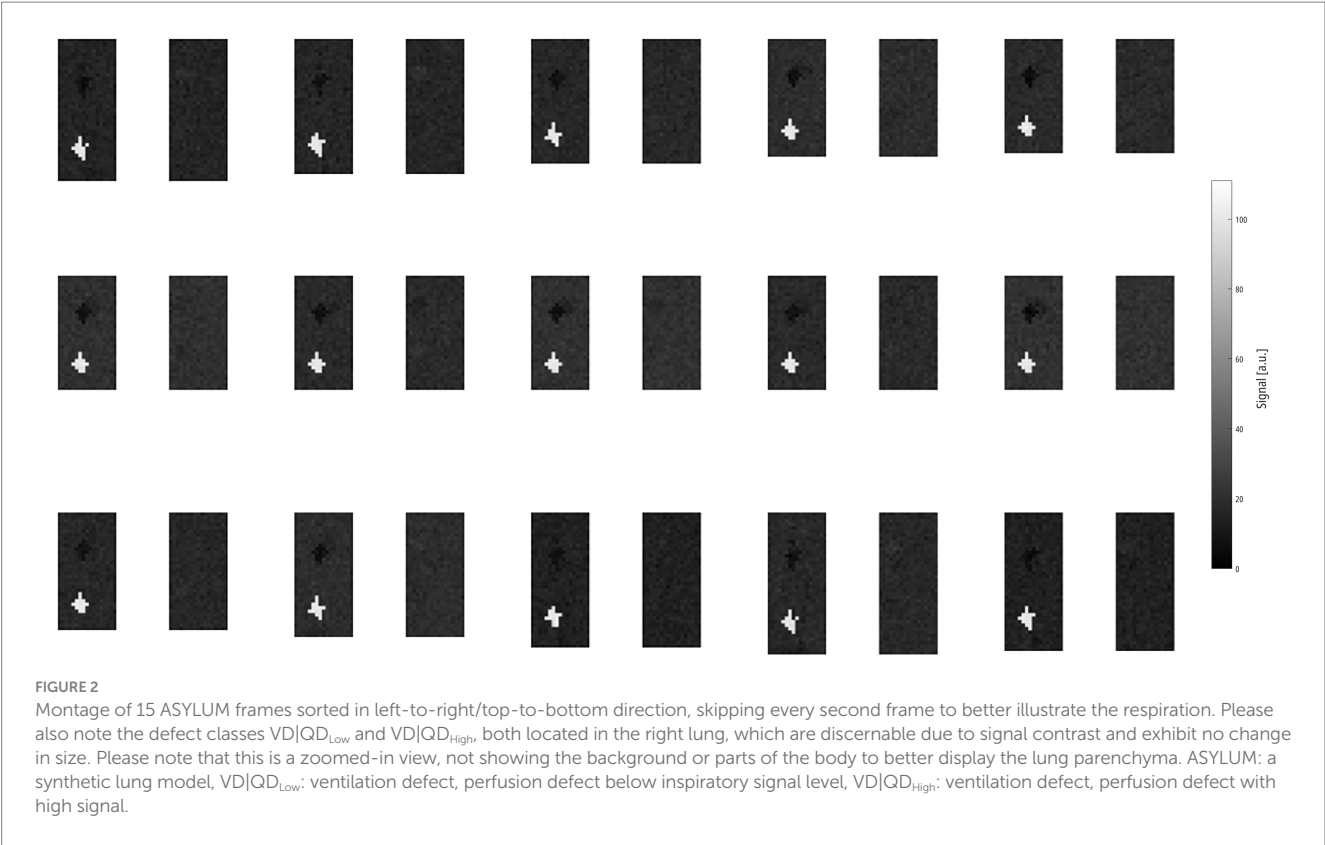
FIGURE 2
Montage of 15 ASYLUM frames sorted in left-to-right/top-to-bottom direction, skipping every second frame to better illustrate the respiration. Please also note the defect classes VD|QD$_{Low}$ and VD|QD$_{High}$, both located in the right lung, which are discernable due to signal contrast and exhibit no change in size. Please note that this is a zoomed-in view, not showing the background or parts of the body to better display the lung parenchyma. ASYLUM: a synthetic lung model, VD|QD$_{Low}$: ventilation defect, perfusion defect below inspiratory signal level, VD|QD$_{High}$: ventilation defect, perfusion defect with high signal.

TABLE 1 (A) Median and interquartile range values of registration performance metrics: root mean squared relative error (RMSRE) using the reference registration, Pearson correlation of the *y*-displacement profile, and registration time. (B) Median and interquartile range values of RVent, JVent, and QA parameters for all registrations and respective omnibus and *post-hoc* test results for difference.

(A) Registration	RMSRE	Correlation (%)	Registration time (min)
ANTs	0.30 (0.27–0.35)	91.91 (91.56–93.07)	149.84 (110.28–151.29)
F-REG	0.27 (0.26–0.28)	97.18 (96.84–97.37)	7.19 (5.71–9.48)
M-REG	0.27 (0.26–0.29)	76.64 (75.12–78.04)	19.64 (19.25–20.82)
Omnibus	1.06e−02 (*)	2.06e−09 (****)	5.33e−09 (****)
ANTs vs. F-REG	3.59e−03 (**)	8.86e−05 (****)	8.86e−05 (****)
ANTs vs. M-REG	1.11e−02 (*)	8.86e−05 (****)	1.03e−04 (***)
F-REG vs. M-REG	7.65e−01 (n.s.)	8.86e−05 (****)	8.86e−05 (****)

(B) Functional values	RVent	JVent	QA	RVent vs. JVent
REF	0.24 (0.24–0.24)	0.24 (0.24–0.24)	3.99 (3.99–4.01)	2.82e−03 (**)
ANTs	0.39 (0.35–0.55)	0.24 (0.23–0.24)	3.94 (3.92–3.96)	8.86e−05 (****)
F-REG	0.23 (0.23–0.23)	0.24 (0.24–0.24)	3.26 (3.24–3.30)	8.86e−05 (****)
M-REG	0.23 (0.23–0.23)	0.27 (0.26–0.27)	3.60 (3.54–3.64)	1.20e−04 (***)
Omnibus	2.21e−11 (****)	3.88e−12 (****)	5.88e−13 (****)	
REF vs. ANTs	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	
REF vs. F-REG	8.86e−05 (****)	4.05e−03 (**)	8.86e−05 (****)	
REF vs. M-REG	1.51e−03 (**)	8.86e−05 (****)	8.86e−05 (****)	
ANTs vs. F-REG	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	
ANTs vs. M-REG	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	
F-REG vs. M-REG	5.50e−01 (n.s.)	8.86e−05 (****)	8.86e−05 (****)	

Post-hoc tests are denoted with an asterisk indicating significance level: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p \leq 0.0001$ (****). Please note that the numbers displayed in this table are rounded to two decimal places. As a result, different values may appear identical despite minor differences in their actual values. RMSRE: root mean squared relative error, REF: reference (known) registration, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, JVent: Jacobian determinant ventilation, QA: perfusion amplitude.
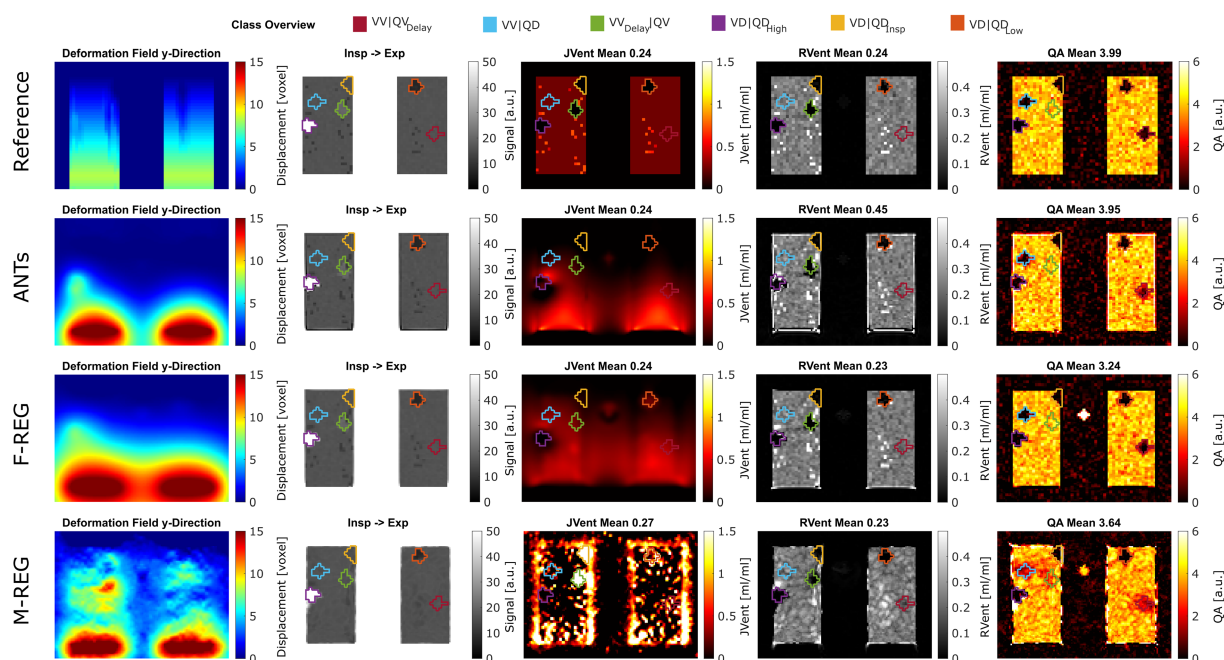
FIGURE 3
Exemplary performance of different registration and quantification methods for ASYLUM. For registration, reference (known deformation), advanced normalization tools (ANTs), Forsberg (F-REG), and Matlab (M-REG) algorithms were used (deformation fields in the *y*-direction are displayed in the first column). The registered image is displayed together with the color-coded class definitions in the second column. For ventilation, the Jacobian determinant (JVent, third column) and signal-based method (RVent, fourth column) were used. Additionally, a signal-based perfusion amplitude quantification (QA) is displayed in the last column. RMSRE: root mean squared relative error, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, JVent: Jacobian determinant ventilation, QA: perfusion amplitude, VV|QV: ventilated and perfused volume, VD|QD$_{Low}$: ventilation and perfusion defect below inspiration signal level, VD|QD$_{Insp}$: ventilation and perfusion defect at inspiration level, VD|QD$_{High}$: ventilation and perfusion defect at high signal level, VV$_{Delay}$|QV: ventilated and perfused volume with delayed ventilation, VV|QD: ventilated volume with perfusion defect, VV|QV$_{Delay}$: ventilated and perfused volume with delayed perfusion.

4.1.3 Sensitivity and specificity

Assessments of regional defect maps, as demonstrated in Figure 5, JVent showed a higher amount of VD in the upper lung regions, which led to a high true positive and low true negative rate in this case. Remarkably, JVent was able to identify the VD|QD$_{High}$ defect class, although it was located in the lower half region with higher sensitivity. All registration variants, except for REF, had a noticeable amount of VD/QD at the lung boundary. M-REG resulted in unusable JVent. In general, RVent and QA identified the defined VD and QD regions.

In concordance with the presented example, as summarized in Table 3 and Figure 6A, the sensitivity was very low for JVent (0–16%) for the non-REF registration variants in the lower quadrants in comparison to upper quadrants (29–99%, significant for ANTs, $p \leq 0.01$). Contrary to this, RVent and QA displayed overall high sensitivity (68–100%) across all quadrants and registration variants. Differences regarding quadrants were only found for QA with F-REG between UR and LR (100% vs. 97%, $p \leq 0.05$).

The specificity analysis (see Figure 6B) showed an inverted performance for JVent in comparison with previously described sensitivity results: 87% for LR/LL vs. 1 and 2% for UR/UL (ANTs). As before, this pattern was less pronounced but significant for F-REG and even more so for M-REG. The specificity for RVent and QA was very high (88–97%) across all registration variants and quadrants.

Sensitivity results as a function of individual defect classes (see Figure 6C) confirmed the better performance of JVent for VD|QD$_{High}$

in comparison to all other classes regarding median values and amount of dispersion. As expected, VV classes were nearly zero for RVent and QV classes were nearly zero for QA. Otherwise, the differences in class performance were less pronounced and less dispersed for RVent and QA in comparison to JVent.

4.2 Patient cohort

Except for ANTs (0.16 vs. 0.17, $p = 0.15$), JVent was significantly higher than the additionally filtered RVent* (0.22 vs. 0.13 (F-REG) and 0.08 vs. 0.04 (M-REG), $p \leq 0.01$). All registration variants resulted in significantly different values. See Table 4 for a summary of all results.

The quadrant analysis (see Table 5) showed significantly increased JVent-derived VD in comparison to RVent and CT for F-REG and especially ANTs: 54/60% UR/UL vs. 16/10% LR/LL (ANTs), 47/58% UR/UL vs. 27/25% LR/LL (F-REG). Contrary to this, RVent* showed a reversed defect ratio: 14/13% UR/UL vs. 38/42% LR/LL (ANTs) and 20/20% UR/UL vs. 44/46% LR/LL (F-REG). This defect distribution was also more similar to CT: 0.23/0.14 UR/UL vs. 42/37% LR/LL. M-REG demonstrated overall increased defect percentages and less pronounced differences between upper and lower lungs. These results were also reflected in the CT overlap coefficient (see Table 6), as UR and UL showed significantly higher coefficients for RVent* in comparison with JVent: 0.72/0.73 vs. 0.62/0.53 (ANTs), 0.68/0.70 vs. 0.57/0.51 (F-REG), and 0.46/0.43 vs. 0.37/0.32, $p \leq 0.001$.
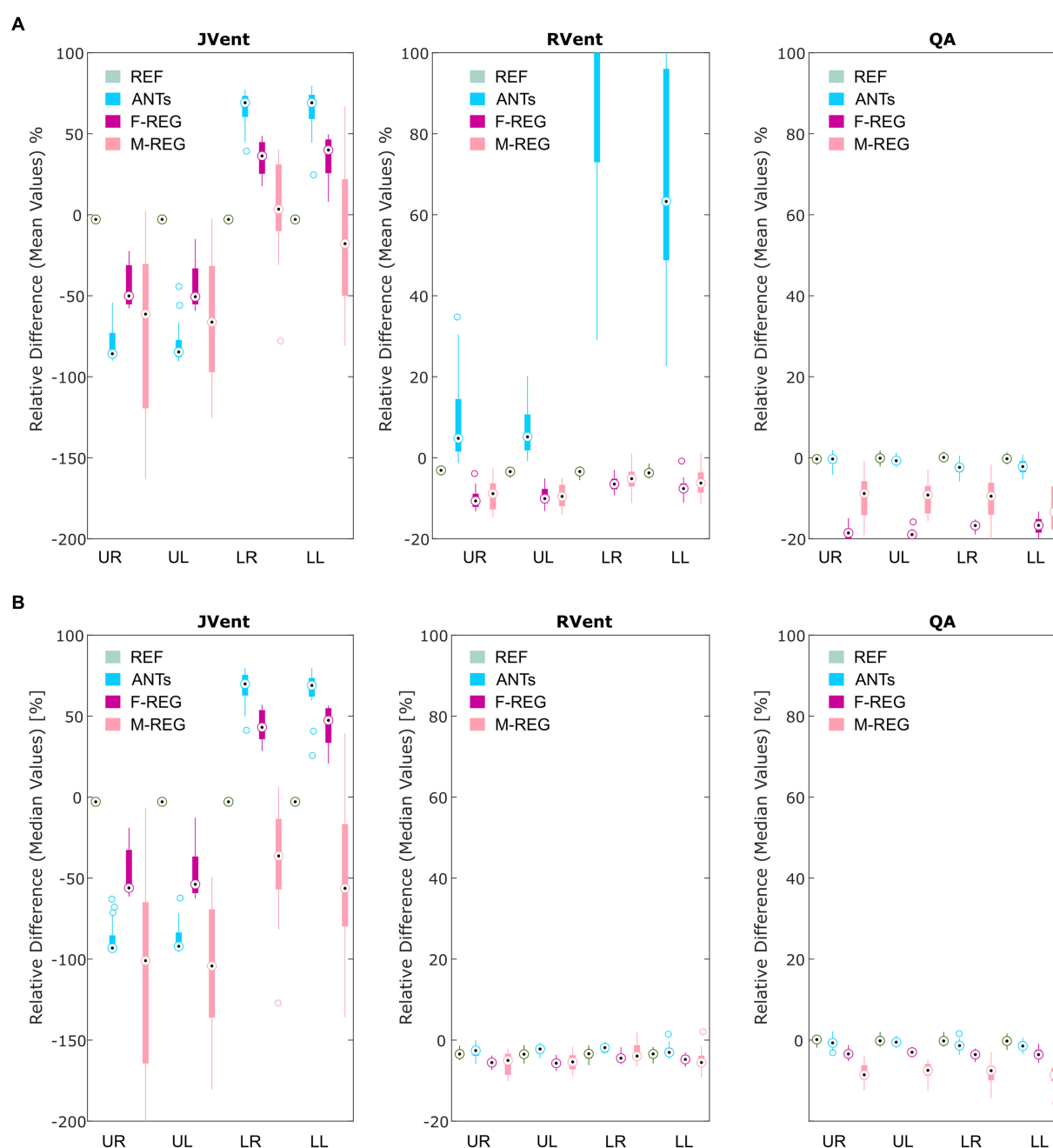
FIGURE 4
Boxplots of the mean **(A)** and median **(B)** relative differences in relation to the predefined functional values for all registration variants depending on quadrant. REF: reference (known) registration, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, JVent: Jacobian determinant ventilation, QA: perfusion amplitude, UR: upper right, UL: upper left, LR: lower right, LL: lower left.

Typical examples of RVent* and JVent maps with PRM and defect distributions are displayed in Figures 7–9. The first shows aligned RVent* and JVent defects with CT in UR/UL quadrants. The second demonstrates a mixed concordance of RVent* and JVent for aligned defects in UR/UL and major CT-aligned RVent defects in LR/LL with no corresponding JVent defects. The third demonstrates a case with CT-aligned RVent defects in LR, LL with no corresponding JVent defects. Contrary to this, JVent shows unaligned major defects in UR/UL. For all cases, the JVent map derived from M-REG was unusable.

5 Discussion

This study describes a feasible framework to create synthetic data mimicking a free-breathing lung MRI acquisition. ASYLUM was used to analyze differences between signal- and deformation-based lung ventilation measurements using different registration algorithms. Both the registration algorithms and functional measurement methods showed significantly different results. Overall, M-REG showed incorrect deformation fields, which resulted in unusable JVent, but more or less comparable measurements regarding RVent in

TABLE 2 Relative differences in relation to the predefined functional values for all registration variants depending on quadrant using mean (A) and median (B) values.

	(A) Relative difference (mean values)											
	JVent				RVent				QA			
Description	**REF**	**ANTs**	**F-REG**	**M-REG**	**REF**	**ANTs**	**F-REG**	**M-REG**	**REF**	**ANTs**	**F-REG**	**M-REG**
UR	−0.03 (−0.03 to −0.03)	−0.86 (−0.88 to −0.73)	−0.50 (−0.55 to −0.31)	−0.61 (−1.19 to −0.30)	−0.03 (−0.04 to −0.03)	0.05 (0.02 to 0.15)	−0.11 (−0.12 to −0.09)	−0.09 (−0.13 to −0.06)	−0.00 (−0.01 to 0.00)	−0.00 (−0.01 to 0.00)	−0.19 (−0.21 to −0.17)	−0.09 (−0.14 to −0.06)
UL	−0.03 (−0.03 to −0.03)	−0.85 (−0.88 to −0.77)	−0.51 (−0.55 to −0.33)	−0.66 (−0.97 to −0.32)	−0.03 (−0.04 to −0.03)	0.05 (0.02 to 0.11)	−0.10 (−0.11 to −0.08)	−0.10 (−0.12 to −0.07)	−0.00 (−0.01 to 0.00)	−0.01 (−0.02 to 0.00)	−0.19 (−0.20 to −0.19)	−0.09 (−0.14 to −0.07)
LR	−0.03 (−0.03 to −0.03)	0.69 (0.60 to 0.73)	0.36 (0.25 to 0.45)	0.03 (−0.10 to 0.31)	−0.03 (−0.04 to −0.03)	1.16 (0.73 to 2.03)	−0.07 (−0.08 to −0.05)	−0.05 (−0.07 to −0.03)	0.00 (−0.00 to 0.00)	−0.02 (−0.04 to −0.02)	−0.17 (−0.17 to −0.16)	−0.10 (−0.14 to −0.06)
LL	−0.03 (−0.03 to −0.03)	0.69 (0.59 to 0.74)	0.40 (0.26 to 0.46)	−0.18 (−0.50 to 0.22)	−0.04 (−0.04 to −0.03)	0.63 (0.49 to 0.96)	−0.08 (−0.09 to −0.06)	−0.06 (−0.09 to −0.04)	−0.00 (−0.01 to 0.00)	−0.02 (−0.04 to −0.01)	−0.17 (−0.19 to −0.15)	−0.13 (−0.18 to −0.07)
Omnibus quadrants	3.27e−09 (****)	1.84e−10 (****)	2.07e−10 (****)	1.49e−05 (****)	9.36e−01 (n.s.)	2.77e−10 (****)	4.85e−05 (****)	3.04e−04 (***)	6.02e−01 (n.s.)	6.02e−04 (***)	2.71e−04 (***)	6.82e−01 (n.s.)
UR vs. LR	1.01e−04 (***)	8.86e−05 (****)	8.86e−05 (****)	1.40e−04 (***)	− (−)	8.86e−05 (****)	5.93e−04 (***)	5.11e−03 (**)	− (−)	2.50e−03 (**)	1.11e−02 (*)	− (−)
UL vs. LL	3.81e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	1.02e−03 (**)	− (−)	8.86e−05 (****)	3.19e−03 (**)	5.73e−03 (**)	− (−)	7.19e−03 (**)	4.49e−04 (***)	− (−)
Description	**Omnibus REF**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**	**Omnibus F-REG**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**				
UR	1.95e−07 (****)	4.79e−02 (*)	8.86e−05 (****)	8.86e−05 (****)	2.06e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)				
UL	1.37e−07 (****)	3.19e−03 (**)	8.86e−05 (****)	8.86e−05 (****)	1.25e−08 (****)	8.86e−05 (****)	1.40e−04 (***)	8.86e−05 (****)				
LR	8.76e−08 (****)	1.11e−02 (*)	8.86e−05 (****)	8.86e−05 (****)	2.06e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)				
LL	5.06e−08 (****)	6.42e−03 (**)	8.86e−05 (****)	8.86e−05 (****)	2.06e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)				
Description	**Omnibus ANTs**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**	**Omnibus M-REG**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**				
UR	5.06e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	8.92e−04 (***)	5.03e−06 (****)	1.03e−04 (***)	1.40e−04 (***)	7.94e−01 (n.s.)				
UL	1.25e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	1.89e−04 (***)	5.29e−06 (****)	1.03e−04 (***)	1.03e−04 (***)	7.65e−01 (n.s.)				
LR	8.76e−08 (****)	1.94e−03 (**)	8.86e−05 (****)	8.86e−05 (****)	9.59e−04 (***)	4.79e−02 (*)	1.00e−02 (**)	2.19e−04 (***)				
LL	2.91e−07 (****)	2.96e−01 (n.s.)	8.86e−05 (****)	8.86e−05 (****)	2.12e−01 (n.s.)	− (−)	− (−)	− (−)				

	(B) Relative difference (median values)											
	JVent				RVent				QA			
Description	**REF**	**ANTs**	**F-REG**	**M-REG**	**REF**	**ANTs**	**F-REG**	**M-REG**	**REF**	**ANTs**	**F-REG**	**M-REG**
UR	−0.03 (−0.03 to −0.03)	−0.93 (−0.94 to −0.85)	−0.56 (−0.58 to −0.33)	−1.01 (−1.64 to −0.65)	−0.03 (−0.04 to −0.03)	−0.03 (−0.04 to −0.02)	−0.06 (−0.06 to −0.05)	−0.05 (−0.09 to −0.03)	0.00 (−0.01 to 0.01)	−0.01 (−0.01 to 0.00)	−0.03 (−0.04 to −0.03)	−0.09 (−0.10 to −0.06)

(Continued)

TABLE 2 (Continued)

(B) Relative difference (median values)											
	JVent				**RVent**				**QA**			
UL	−0.03 (−0.03 to −0.03)	−0.92 (−0.94 to −0.84)	−0.54 (−0.59 to −0.37)	−1.04 (−1.36 to −0.69)	−0.03 (−0.04 to −0.03)	−0.02 (−0.03 to −0.02)	−0.06 (−0.06 to −0.05)	−0.05 (−0.07 to −0.04)	−0.00 (−0.00 to 0.01)	−0.01 (−0.01 to 0.00)	−0.03 (−0.04 to −0.03)	−0.07 (−0.09 to −0.06)
LR	−0.03 (−0.03 to −0.03)	0.70 (0.63 to 0.75)	0.43 (0.36 to 0.54)	−0.36 (−0.57 to −0.13)	−0.03 (−0.05 to −0.03)	−0.02 (−0.03 to −0.01)	−0.04 (−0.05 to −0.03)	−0.04 (−0.05 to −0.01)	−0.00 (−0.01 to 0.01)	−0.01 (−0.02 to −0.01)	−0.04 (−0.04 to −0.03)	−0.08 (−0.10 to −0.06)
LL	−0.03 (−0.03 to −0.03)	0.69 (0.62 to 0.74)	0.47 (0.34 to 0.55)	−0.56 (−0.80 to −0.17)	−0.03 (−0.04 to −0.03)	−0.03 (−0.04 to −0.02)	−0.05 (−0.06 to −0.04)	−0.06 (−0.06 to −0.04)	−0.00 (−0.01 to 0.00)	−0.01 (−0.02 to −0.00)	−0.04 (−0.04 to −0.02)	−0.09 (−0.10 to −0.07)
Omnibus quadrants	1.00e+00 (n.s.)	2.01e−10 (****)	2.13e−10 (****)	1.62e−04 (***)	7.96e−01 (n.s.)	7.13e−02 (n.s.)	2.21e−03 (**)	1.92e−01 (n.s.)	5.16e−01 (n.s.)	1.48e−02 (*)	6.02e−01 (n.s.)	7.39e−01 (n.s.)
UR vs. LR	− (−)	8.86e−05 (****)	8.86e−05 (****)	2.93e−04 (***)	− (−)	− (−)	6.81e−04 (***)	− (−)	− (−)	3.33e−02 (*)	− (−)	− (−)
UL vs. LL	− (−)	8.86e−05 (****)	8.86e−05 (****)	3.59e−03 (**)	− (−)	− (−)	8.97e−03 (**)	− (−)	− (−)	2.51e−02 (*)	− (−)	− (−)
Description	**Omnibus REF**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**	**Omnibus F-REG**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**				
UR	1.95e−07 (****)	4.38e−02 (*)	8.86e−05 (****)	8.86e−05 (****)	5.33e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	1.63e−04 (***)				
UL	8.76e−08 (****)	2.51e−02 (*)	8.86e−05 (****)	8.86e−05 (****)	5.33e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	1.03e−04 (***)				
LR	1.37e−07 (****)	3.04e−02 (*)	8.86e−05 (****)	8.86e−05 (****)	2.50e−07 (****)	8.86e−05 (****)	8.86e−05 (****)	2.28e−02 (*)				
LL	8.76e−08 (****)	5.11e−03 (**)	8.86e−05 (****)	8.86e−05 (****)	5.06e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	1.51e−03 (**)				
Description	**Omnibus ANTs**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**	**Omnibus M-REG**	**JVent vs. Rvent**	**JVent vs. QA**	**RVent vs. QA**				
UR	2.64e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	7.80e−04 (***)	2.91e−07 (****)	8.86e−05 (****)	1.03e−04 (***)	5.11e−03 (**)				
UL	2.06e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	1.25e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	1.02e−03 (**)				
LR	1.95e−07 (****)	8.86e−05 (****)	8.86e−05 (****)	5.22e−02 (n.s.)	1.18e−06 (****)	1.89e−04 (***)	3.90e−04 (***)	2.19e−04 (***)				
LL	1.25e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	2.50e−03 (**)	3.90e−04 (***)	1.16e−03 (**)	1.71e−03 (**)	1.32e−03 (**)				

Respective omnibus and *post-hoc* test results for differences regarding quadrants (omnibus quadrants) and parameters (omnibus registration abbreviation) are included as well. $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p \leq 0.0001$ (****). Please note that the numbers displayed in this table are rounded to two decimal places. As a result, different values may appear identical despite minor differences in their actual values. REF: reference (known) registration, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, JVent: Jacobian determinant ventilation, QA: perfusion amplitude, UR: upper right, UL: upper left, LR: lower right, LL: lower left.

comparison to ANTs and F-REG. These registration variants also yielded deformation and JVent, similar to the known REF registration. Nevertheless, further analysis showed a strong regional JVent bias, which resulted in high defect classification in the upper lung regions with very low specificity. Contrary to this, the artificially created ventilation and perfusion defect regions were mostly correctly identified by the signal-based approach. Similar VD detection differences for JVent and RVent were also found in a patient cohort using CT as a gold standard.

Registration algorithms are an important part of functional lung MRI, have numerous parameters for tweaking, and are known to have a great impact on the final results. While registration algorithms are often tested initially with synthetic, but not necessarily realistic data (42, 47), in the realm of lung function, due to missing ground truth, the registration performance is mostly evaluated solely by testing the reproducibility of the final results and by image similarity metrics (e.g., segmented overlap of edges/features measured with Dice or structural similarity index measure) (38, 41). In this regard, the evaluation of a registration algorithm with a known lung deformation using ASYLUM is a novel approach in this specific field. Although M-REG performed as well as F-REG regarding RMSRE, the correlation to the known deformation revealed, that F-REG and ANTs
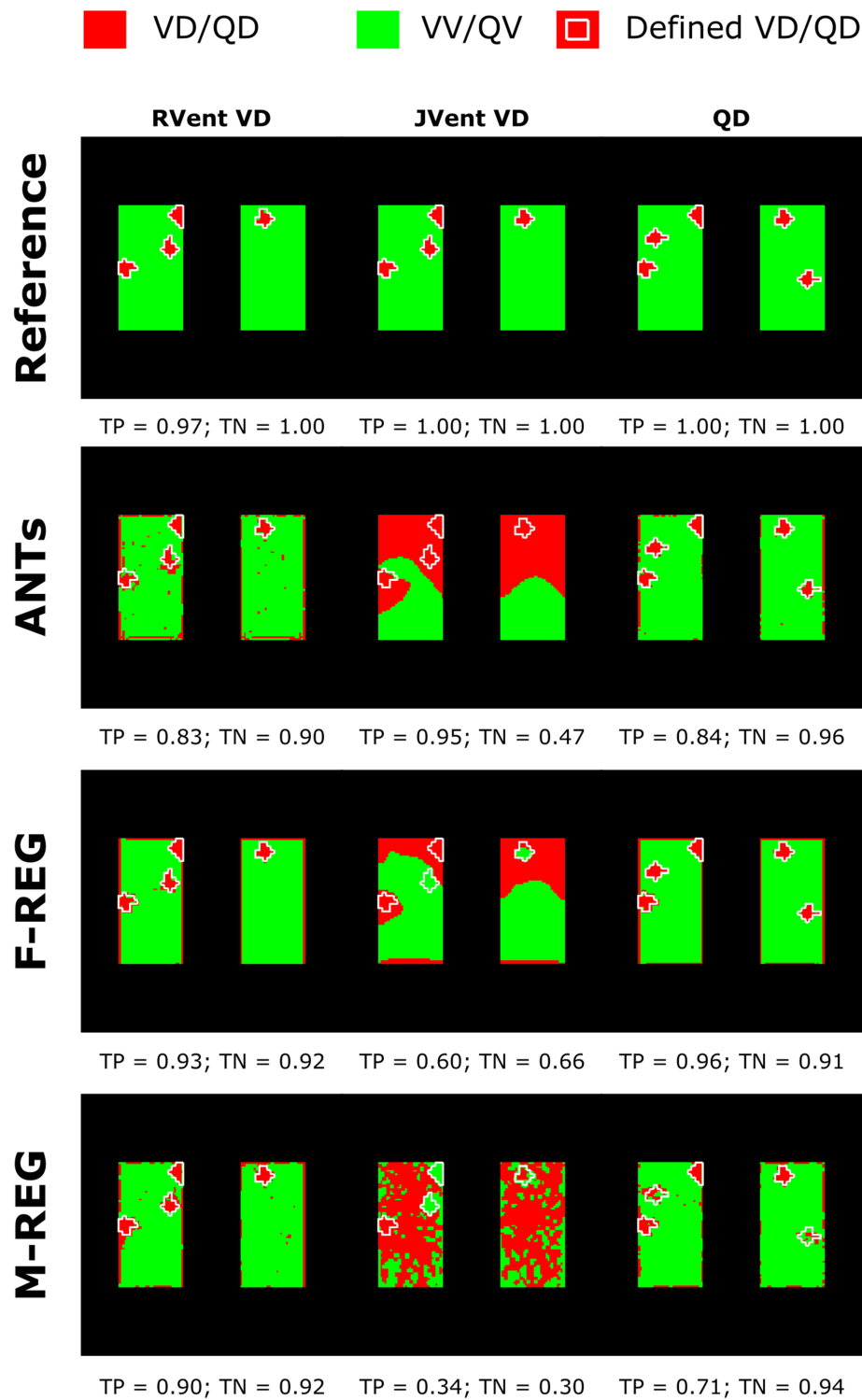
FIGURE 5
Regional concordance of the ventilation defects (VD) derived from RVent, JVent, and the perfusion defects (QD) derived from QA using registrations with Reference, ANTs, F-REG, and M-REG. As expected, the reference registration results in nearly perfect correspondence to the defined classes. VD derived from JVent results in overestimation, especially at the upper regions, leading to low specificity (true negative rate) scores. All VD/QD measurements show problems near the edges. REF: reference (known) registration, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, JVent: Jacobian determinant ventilation, QA: perfusion amplitude, TP: true positive rate, TN: true negative rate.

TABLE 3 True positive (A) and true negative (B) results based on VD calculated from JVent, RVent, and perfusion defect (QD) derived from QA.

(A) True positive rate											
	JVent				RVent				QA		
Description	REF	ANTs	F-REG	M-REG	REF	ANTs	F-REG	M-REG	REF	ANTs	F-REG	M-REG
UR	1.00 (0.50–1.00)	0.99 (0.38–1.00)	0.29 (0.00–0.86)	0.16 (0.00–0.21)	1.00 (0.48–1.00)	0.82 (0.00–0.95)	0.92 (0.33–1.00)	0.93 (0.43–0.97)	1.00 (1.00–1.00)	0.79 (0.67–0.85)	1.00 (0.98–1.00)	0.68 (0.42–0.92)
UL	1.00 (1.00–1.00)	0.99 (0.80–1.00)	0.50 (0.00–0.97)	0.30 (0.00–0.69)	1.00 (0.93–1.00)	0.84 (0.65–0.93)	0.92 (0.81–0.98)	0.94 (0.82–1.00)	1.00 (1.00–1.00)	0.78 (0.72–0.86)	0.97 (0.93–1.00)	0.73 (0.44–0.96)
LR	1.00 (0.50–1.00)	0.16 (0.00–0.62)	0.00 (0.00–0.48)	0.00 (0.00–0.51)	1.00 (0.46–1.00)	0.83 (0.30–0.97)	0.87 (0.25–0.97)	0.94 (0.28–0.99)	1.00 (1.00–1.00)	0.76 (0.72–0.84)	0.97 (0.94–1.00)	0.70 (0.51–1.00)
LL	1.00 (1.00–1.00)	0.06 (0.00–0.42)	0.00 (0.00–0.50)	0.16 (0.00–0.44)	1.00 (0.95–1.00)	0.84 (0.76–0.94)	0.90 (0.87–1.00)	0.87 (0.83–0.95)	1.00 (1.00–1.00)	0.78 (0.70–0.82)	0.98 (0.92–1.00)	0.74 (0.45–0.84)
Omnibus quadrants	7.53e−01 (n.s.)	5.21e−05 (****)	5.02e−01 (n.s.)	7.23e−01 (n.s.)	8.98e−01 (n.s.)	9.77e−01 (n.s.)	6.14e−01 (n.s.)	6.73e−01 (n.s.)	6.66e−01 (n.s.)	9.78e−01 (n.s.)	4.21e−02 (*)	7.51e−01 (n.s.)
UR vs. LR	− (−)	3.55e−03 (**)	− (−)	− (−)	− (−)	− (−)	− (−)	− (−)	− (−)	− (−)	3.98e−02 (*)	− (−)
UL vs. LL	− (−)	1.23e−03 (**)	− (−)	− (−)	− (−)	− (−)	− (−)	− (−)	− (−)	− (−)	6.58e−01 (n.s.)	− (−)

Description	Omnibus REF	JVent vs. RVent	JVent vs. QA	RVent vs. QA	Omnibus F-REG	JVent vs. Rvent	JVent vs. QA	RVent vs. QA
UR	8.23e−03 (**)	1.25e−01 (n.s.)	2.19e−01 (n.s.)	3.91e−02 (*)	4.61e−05 (****)	1.02e−02 (*)	2.00e−03 (**)	4.03e−03 (**)
UL	6.22e−02 (n.s.)	− (−)	− (−)	− (−)	9.36e−03 (**)	1.61e−02 (*)	1.30e−02 (*)	1.02e−01 (n.s.)
LR	6.63e−02 (n.s.)	− (−)	− (−)	− (−)	9.93e−04 (***)	3.23e−03 (**)	4.84e−04 (***)	1.33e−02 (*)
LL	2.24e−02 (*)	2.50e−01 (n.s.)	5.00e−01 (n.s.)	6.25e−02 (n.s.)	3.65e−06 (****)	3.29e−04 (***)	1.52e−04 (***)	1.51e−02 (*)

Description	Omnibus ANTs	JVent vs. RVent	JVent vs. QA	RVent vs. QA	Omnibus M-REG	JVent vs. Rvent	JVent vs. QA	RVent vs. QA
UR	8.68e−03 (**)	7.32e−04 (***)	3.90e−01 (n.s.)	6.47e−01 (n.s.)	2.15e−03 (**)	4.88e−04 (***)	1.48e−03 (**)	4.21e−01 (n.s.)
UL	2.01e−03 (**)	3.66e−04 (***)	1.41e−02 (*)	8.11e−01 (n.s.)	8.37e−04 (***)	7.10e−04 (***)	1.48e−02 (*)	8.40e−02 (n.s.)
LR	2.55e−03 (**)	8.54e−04 (***)	2.47e−03 (**)	8.11e−01 (n.s.)	2.67e−03 (**)	1.22e−03 (**)	1.17e−03 (**)	8.87e−01 (n.s.)
LL	2.06e−04 (***)	5.01e−04 (***)	3.41e−04 (***)	4.86e−02 (*)	1.92e−05 (****)	3.50e−04 (***)	1.96e−04 (***)	1.22e−01 (n.s.)

(B) True negative rate											
	JVent				RVent				QA		
Description	REF	ANTs	F-REG	M-REG	REF	ANTs	F-REG	M-REG	REF	ANTs	F-REG	M-REG
UR	1.00 (0.50–1.00)	0.01 (0.00–0.09)	0.29 (0.25–0.47)	0.25 (0.23–0.28)	1.00 (0.50–1.00)	0.92 (0.89–0.93)	0.91 (0.90–0.91)	0.92 (0.90–0.93)	1.00 (1.00–1.00)	0.97 (0.95–0.97)	0.91 (0.89–0.92)	0.95 (0.94–0.97)

(Continued)

TABLE 3 (Continued)

	(B) True negative rate											
	JVent				RVent				QA			
UL	1.00 (1.00–1.00)	0.02 (0.01–0.06)	0.29 (0.24–0.48)	0.25 (0.22–0.27)	1.00 (1.00–1.00)	0.91 (0.90–0.92)	0.91 (0.90–0.91)	0.92 (0.91–0.93)	1.00 (1.00–1.00)	0.95 (0.94–0.97)	0.91 (0.89–0.91)	0.95 (0.94–0.96)
LR	1.00 (0.50–1.00)	0.87 (0.85–0.90)	0.92 (0.86–0.94)	0.33 (0.31–0.36)	1.00 (1.00–1.00)	0.90 (0.88–0.91)	0.94 (0.93–0.94)	0.94 (0.91–0.95)	1.00 (1.00–1.00)	0.96 (0.95–0.97)	0.91 (0.90–0.91)	0.95 (0.93–0.97)
LL	1.00 (1.00–1.00)	0.87 (0.80–0.89)	0.91 (0.87–0.94)	0.31 (0.29–0.33)	1.00 (1.00–1.00)	0.88 (0.87–0.89)	0.93 (0.92–0.94)	0.93 (0.92–0.94)	1.00 (1.00–1.00)	0.95 (0.94–0.96)	0.91 (0.89–0.91)	0.94 (0.92–0.96)
Omnibus quadrants	7.53e−01 (n.s.)	1.45e−10 (****)	1.84e−10 (****)	6.67e−07 (****)	7.48e−01 (n.s.)	2.89e−05 (****)	1.23e−07 (****)	4.06e−02 (*)	6.66e−01 (n.s.)	4.40e−01 (n.s.)	6.50e−01 (n.s.)	9.32e−01 (n.s.)
UR vs. LR	− (−)	8.84e−05 (****)	8.86e−05 (****)	1.40e−04 (***)	− (−)	4.49e−04 (***)	1.63e−04 (***)	1.56e−01 (n.s.)	− (−)	− (−)	− (−)	− (−)
UL vs. LL	− (−)	8.86e−05 (****)	8.86e−05 (****)	2.19e−04 (***)	− (−)	2.93e−04 (***)	3.38e−04 (***)	4.38e−02 (*)	− (−)	− (−)	− (−)	− (−)
Description	**Omnibus REF**	**JVent vs. RVent**	**JVent vs. QA**	**RVent vs. QA**	**Omnibus F-REG**	**JVent vs. RVent**	**JVent vs. QA**	**RVent vs. QA**				
UR	6.95e−02 (n.s.)	− (−)	− (−)	− (−)	2.50e−07 (****)	8.86e−05 (****)	8.86e−05 (****)	7.94e−01 (n.s.)				
UL	1.00e+00 (n.s.)	− (−)	− (−)	− (−)	1.36e−07 (****)	8.86e−05 (****)	8.86e−05 (****)	3.34e−01 (n.s.)				
LR	1.74e−01 (n.s.)	− (−)	− (−)	− (−)	1.67e−04 (***)	2.90e−03 (**)	5.75e−01 (n.s.)	8.86e−05 (****)				
LL	1.35e−01 (n.s.)	− (−)	− (−)	− (−)	1.29e−03 (**)	8.96e−03 (**)	9.70e−01 (n.s.)	2.19e−04 (***)				
Description	**Omnibus ANTs**	**JVent vs. RVent**	**JVent vs. QA**	**RVent vs. QA**	**Omnibus M-REG**	**JVent vs. RVent**	**JVent vs. QA**	**RVent vs. QA**				
UR	2.06e−09 (****)	8.84e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	5.33e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	1.20e−04 (***)				
UL	2.06e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)	2.06e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	8.86e−05 (****)				
LR	9.66e−07 (****)	2.51e−02 (*)	1.03e−04 (***)	8.86e−05 (****)	5.33e−09 (****)	8.86e−05 (****)	8.86e−05 (****)	1.63e−04 (***)				
LL	1.37e−07 (****)	9.46e−03 (**)	8.86e−05 (****)	8.86e−05 (****)	5.06e−08 (****)	8.86e−05 (****)	8.86e−05 (****)	5.73e−03 (**)				

Respective omnibus and post-hoc test results for differences regarding quadrants (omnibus quadrants) and parameters (omnibus registration abbreviation) are included as well. $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p \leq 0.0001$ (****). Please note that the numbers displayed in this table are rounded to two decimal places. As a result, different values may appear identical despite minor differences in their actual values. REF: reference (known) registration, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, JVent: Jacobian determinant ventilation, QD: perfusion defect, UR: upper right, UL: upper left, LR: lower right, LL: lower left.

are recreating actually more accurate deformations. Thus, it was not surprising, that all JVent values of M-REG were unusable. Nevertheless, M-REG leads to similar RVent and QA results, which is an indication of the stability of signal-based calculations. Although this seems paradoxical at first, this finding can be explained by the fact that signal-based approaches require only correct registrations of signal groups and not necessarily correct, i.e., physiological sound movements of individual voxels.

Ensuring physiologic deformation vectors requires additional regularization, e.g., in the form of smoothing. Therefore, the degree of regularization might explain the blurred versions of the ANTs and F-REG registration results in comparison to REF. This blurring can explain the difficulty to distinguish defects from the surrounding JVent values. Both registrations also overestimated the movement in the lower lung and underestimated the movement in the upper lung regions, leading to a gradient not present in REF data. Then again, signal-based measurements were able to accurately identify defects, but blurring was also visible when inspecting the $VV_{Comp}|QV$ voxels, which were distributed as speckles and therefore are a good indicator for the accuracy of a registration. Registration artifacts at the edges can also be explained by smoothing and were most prominent for ANTs. As expected, the influence of these edge artifacts was prominent
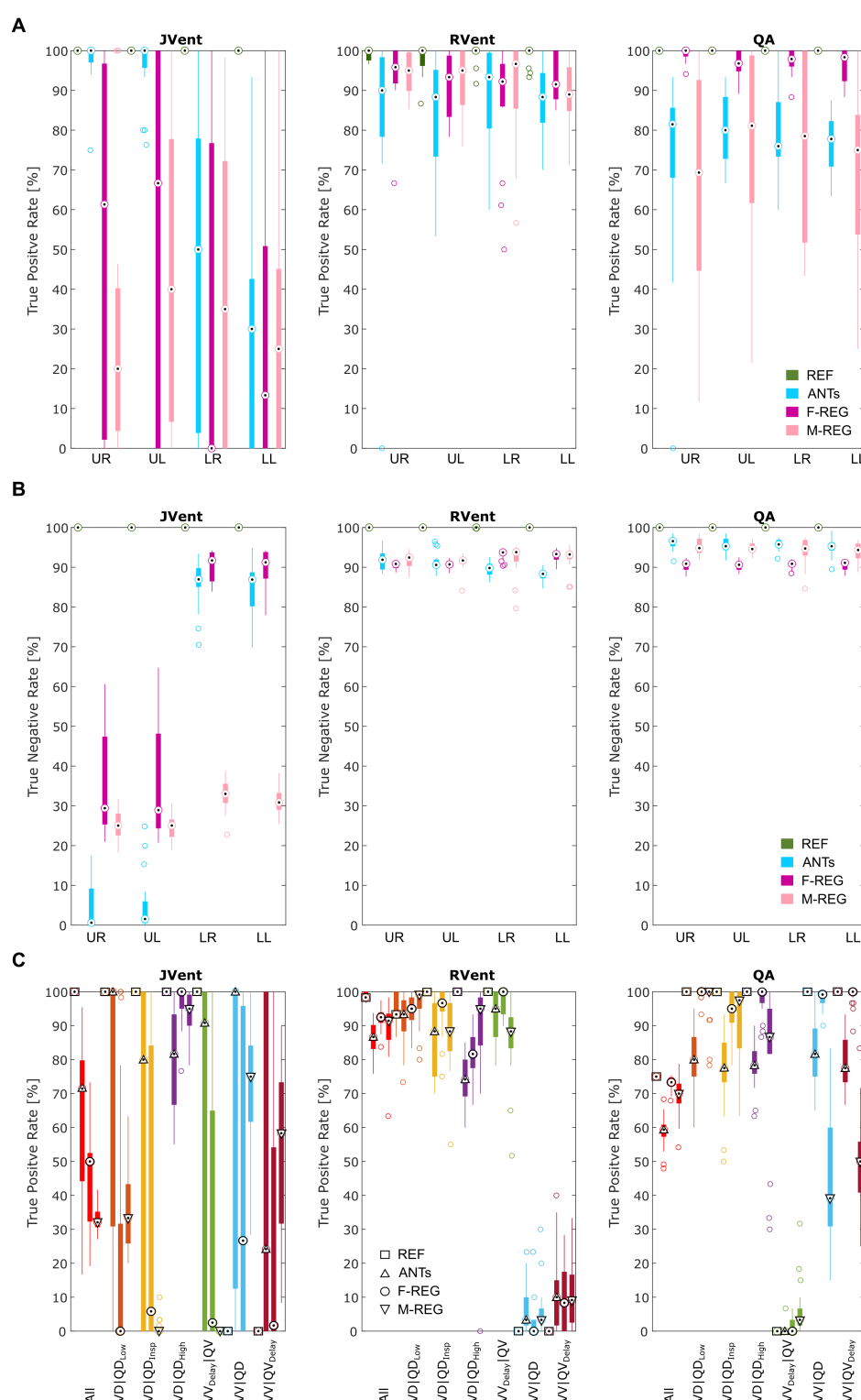
FIGURE 6

True positive **(A)** and true negative **(B)** rates for VD and QD based on thresholded JVent, RVent, and QA parameters according to the 90th percentile ×
0.4 with different registration methods. Please note the considerably lower rates of JVent in comparison to the signal-based RVent and QA. **(C)** Shows
the performance for the whole lung ROI depending on the defect class and registration algorithm (as indicated by different symbols at the median
position). REF: reference (known) registration, ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent:
regional ventilation, JVent: Jacobian determinant ventilation, QA: perfusion amplitude, All: all defect classes, $VD|QD_{Low}$: ventilation and perfusion defect
below inspiration signal level, $VD|QD_{Insp}$: ventilation and perfusion defect at inspiration level, $VD|QD_{High}$: ventilation and perfusion defect at high signal
level, $VV_{Delay}|QV$: ventilated and perfused volume with delayed ventilation, $VV|QD$: ventilated volume with perfusion defect, $VV|QV_{Delay}$: ventilated and
perfused volume with delayed perfusion, UR: upper right, UL: upper left, LR: lower right, LL: lower left.

TABLE 4 RVent and JVent median (interquartile) values of COPD patient cohort (*n* = 36).

Description	RVent*	JVent	RVent* vs. JVent
ANTs	0.17 (0.11–0.22)	0.16 (0.12–0.21)	1.47e−01 (n.s.)
F-REG	0.13 (0.08–0.18)	0.22 (0.16–0.32)	7.39e−27 (****)
M-REG	0.04 (0.02–0.06)	0.08 (0.02–0.15)	4.16e−09 (****)
Omnibus	1.13e−65 (****)	5.36e−44 (****)	—
ANTs vs. F-REG	7.39e−27 (****)	7.84e−27 (****)	—
ANTs vs. M-REG	5.37e−27 (****)	2.14e−13 (****)	—
F-REG vs. M-REG	8.00e−27 (****)	2.21e−25 (****)	—

Except for ANTs, all values were significantly different in regard to method (RVent vs. JVent) and registration algorithm. $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p \leq 0.0001$ (****). Please note that the numbers displayed in this table are rounded to two decimal places. As a result, different values may appear identical despite minor differences in their actual values. ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent*: filtered regional ventilation, JVent: Jacobian determinant ventilation.

TABLE 5 Overview of ventilation defects (VD) distribution for different quadrants, methods, and registration algorithms.

	ANTs						
	CT	JVent	RVent*	Omnibus ANTS	CT vs. JVent	CT vs. RVent*	JVent vs. RVent*
UR	0.23 (0.02–0.50)	0.54 (0.23–0.89)	0.14 (0.01–0.41)	4.47e−27 (****)	1.87e−12 (****)	9.53e−05 (****)	2.98e−22 (****)
UL	0.14 (0.02–0.44)	0.60 (0.20–0.96)	0.13 (0.01–0.35)	5.44e−28 (****)	3.36e−16 (****)	1.28e−02 (*)	1.72e−23 (****)
LR	0.42 (0.11–0.65)	0.16 (0.04–0.36)	0.38 (0.12–0.60)	9.86e−13 (****)	6.50e−12 (****)	2.85e−01 (n.s.)	1.52e−15 (****)
LL	0.37 (0.08–0.56)	0.10 (0.02–0.32)	0.42 (0.12–0.60)	1.15e−15 (****)	1.45e−11 (****)	2.58e−01 (n.s.)	2.37e−17 (****)
Omnibus quadrants	3.28e−13 (****)	3.86e−45 (****)	7.27e−22 (****)				
UR vs. LR	1.51e−05 (****)	6.92e−22 (****)	3.12e−11 (****)				
UL vs. LL	6.77e−08 (****)	1.98e−23 (****)	6.87e−14 (****)				

	F-REG						
	CT	JVent	RVent*	Omnibus F-REG	CT vs. JVent	CT vs. RVent*	JVent vs. RVent*
UR	0.23 (0.02–0.50)	0.47 (0.30–0.81)	0.20 (0.08–0.43)	1.04e−23 (****)	4.13e−14 (****)	5.88e−02 (n.s.)	1.66e−23 (****)
UL	0.14 (0.02–0.44)	0.58 (0.37–0.79)	0.20 (0.07–0.35)	6.99e−33 (****)	7.15e−18 (****)	6.89e−01 (n.s.)	2.65e−25 (****)
LR	0.42 (0.11–0.65)	0.27 (0.17–0.39)	0.44 (0.19–0.66)	5.66e−06 (****)	5.87e−05 (****)	2.64e−01 (n.s.)	1.14e−09 (****)
LL	0.37 (0.08–0.56)	0.25 (0.15–0.39)	0.46 (0.21–0.66)	2.23e−07 (****)	8.82e−03 (**)	3.78e−04 (***)	3.09e−12 (****)
Omnibus quadrants	3.28e−13 (****)	6.39e−29 (****)	2.49e−17 (****)				
UR vs. LR	1.51e−05 (****)	3.61e−17 (****)	1.22e−10 (****)				
UL vs. LL	6.77e−08 (****)	1.42e−19 (****)	3.30e−14 (****)				

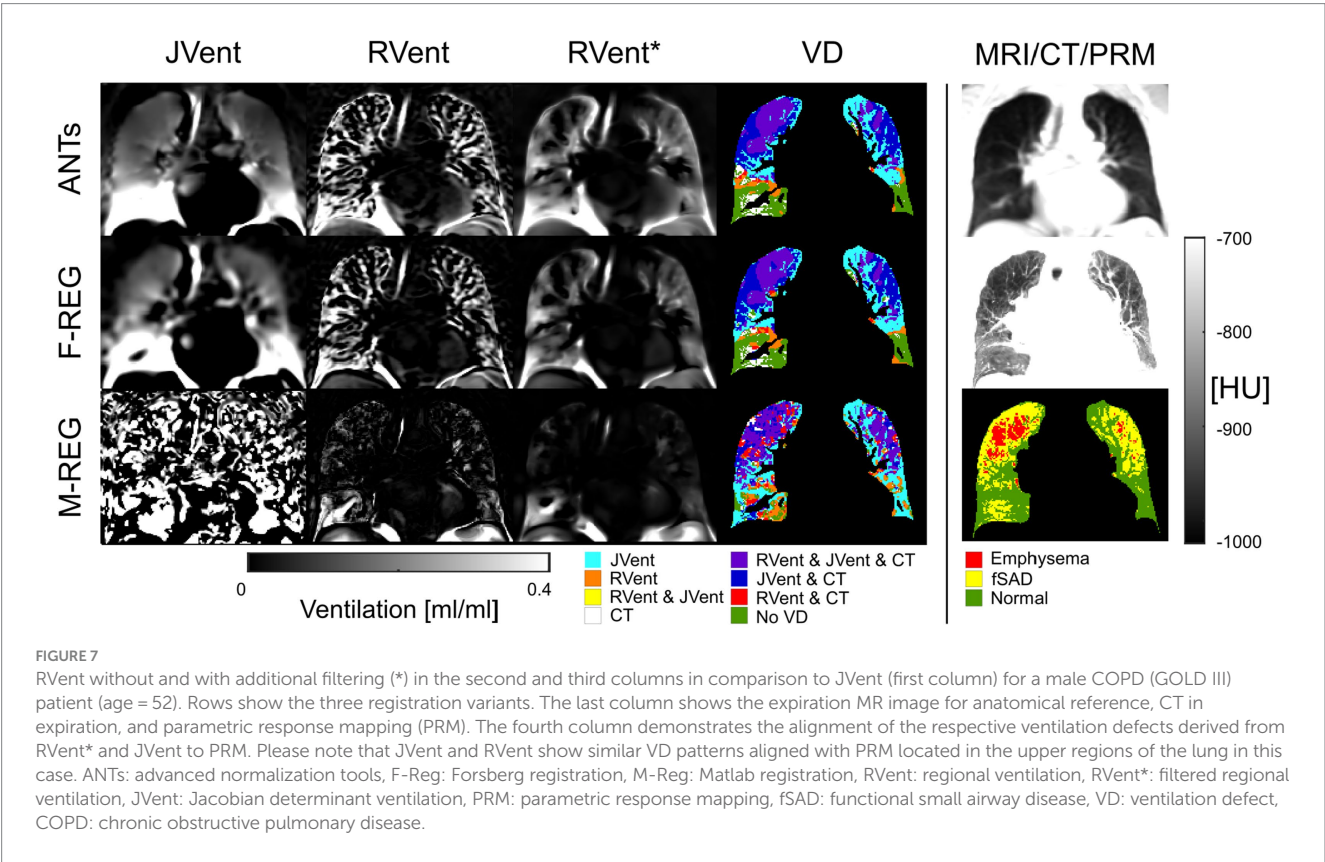|	M-REG						
	CT	JVent	RVent*	Omnibus M-REG	CT vs. JVent	CT vs. RVent*	JVent vs. RVent*
UR	0.23 (0.02–0.50)	0.75 (0.70–0.77)	0.65 (0.59–0.74)	1.16e−35 (****)	4.91e−24 (****)	6.57e−22 (****)	2.65e−19 (****)
UL	0.14 (0.02–0.44)	0.77 (0.72–0.80)	0.68 (0.60–0.75)	3.71e−38 (****)	1.01e−24 (****)	2.84e−23 (****)	6.14e−17 (****)
LR	0.42 (0.11–0.65)	0.70 (0.66–0.75)	0.71 (0.63–0.78)	2.50e−21 (****)	3.78e−19 (****)	5.68e−20 (****)	3.60e−01 (n.s.)
LL	0.37 (0.08–0.56)	0.73 (0.68–0.78)	0.73 (0.64–0.81)	6.11e−27 (****)	7.76e−21 (****)	1.04e−22 (****)	9.17e−01 (n.s.)
Omnibus quadrants	3.28e−13 (****)	2.99e−10 (****)	7.98e−03 (**)				
UR vs. LR	1.51e−05 (****)	6.40e−06 (****)	1.91e−02 (*)				
UL vs. LL	6.77e−08 (****)	5.66e−05 (****)	1.76e−03 (**)				

Significantly higher VD$_{JVent}$ were present in the upper lung regions in comparison to the lower regions for ANTs and for F-REG. RVent showed a reversed relationship for ANTs and F-REG similar to CT. Respective omnibus and post-hoc test results for differences regarding quadrants (omnibus quadrants) and parameters (omnibus registration abbreviation) are included as well. $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p \leq 0.0001$ (****). Please note that the numbers displayed in this table are rounded to two decimal places. As a result, different values may appear identical despite minor differences in their actual values. ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent*: filtered regional ventilation, JVent: Jacobian determinant ventilation, UR: upper right, UL: upper left, LR: lower right, LL: lower left, VD: ventilation defect.
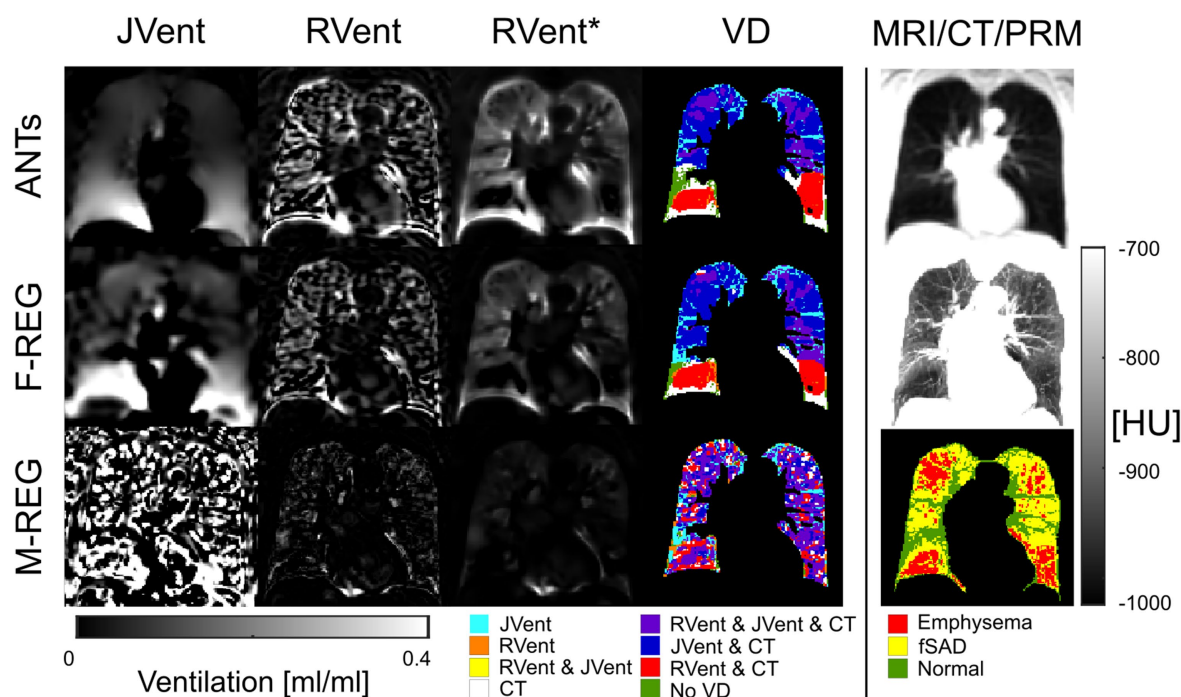
TABLE 6 Two-class overlap coefficient of ventilation defect (VD) and ventilated volume (VV) of RVent and JVent with CT as gold standard for different quadrants and registration algorithms.

Description	ANTs			F-REG			M-REG		
	RVent*	JVent	RVent* vs. JVent	RVent*	JVent	RVent* vs. JVent	RVent*	JVent	RVent* vs. JVent
UR	0.72 (0.59–0.89)	0.62 (0.44–0.77)	4.54e−07 (****)	0.68 (0.54–0.86)	0.57 (0.44–0.71)	3.21e−08 (****)	0.46 (0.39–0.54)	0.37 (0.29–0.53)	2.81e−10 (****)
UL	0.73 (0.54–0.90)	0.53 (0.34–0.77)	2.49e−10 (****)	0.70 (0.52–0.87)	0.51 (0.37–0.64)	3.36e−12 (****)	0.43 (0.32–0.53)	0.32 (0.23–0.47)	2.14e−12 (****)
LR	0.60 (0.52–0.76)	0.56 (0.42–0.76)	5.72e−04 (***)	0.59 (0.51–0.71)	0.53 (0.42–0.70)	6.04e−06 (****)	0.49 (0.41–0.60)	0.44 (0.35–0.55)	8.39e−09 (****)
LL	0.58 (0.48–0.74)	0.53 (0.41–0.83)	6.97e−02 (n.s.)	0.55 (0.46–0.68)	0.50 (0.41–0.76)	3.69e−03 (**)	0.47 (0.36–0.56)	0.42 (0.30–0.53)	4.15e−06 (****)
Omnibus quadrants	1.19e−11 (****)	1.23e−01 (n.s.)	—	1.77e−07 (****)	2.74e−02 (*)	—	4.22e−07 (****)	1.10e−16 (****)	—
UR vs. LR	9.43e−09 (****)	– (–)	—	2.26e−06 (****)	3.04e−01 (n.s.)	—	2.75e−02 (*)	4.93e−05 (****)	—
UL vs. LL	2.49e−06 (****)	– (–)	—	5.74e−06 (****)	3.44e−02 (*)	—	8.11e−03 (**)	4.75e−07 (****)	—

Especially for UR and UL, RVent showed significantly higher overlaps. $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p \leq 0.0001$ (****). Please note that the numbers displayed in this table are rounded to two decimal places. As a result, different values may appear identical despite minor differences in their actual values. ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent*: filtered regional ventilation, JVent: Jacobian determinant ventilation, UR: upper right, UL: upper left, LR: lower right, LL: lower left, VD: ventilation defect.

FIGURE 7
RVent without and with additional filtering (*) in the second and third columns in comparison to JVent (first column) for a male COPD (GOLD III) patient (age = 52). Rows show the three registration variants. The last column shows the expiration MR image for anatomical reference, CT in expiration, and parametric response mapping (PRM). The fourth column demonstrates the alignment of the respective ventilation defects derived from RVent* and JVent to PRM. Please note that JVent and RVent show similar VD patterns aligned with PRM located in the upper regions of the lung in this case. ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, RVent*: filtered regional ventilation, JVent: Jacobian determinant ventilation, PRM: parametric response mapping, fSAD: functional small airway disease, VD: ventilation defect, COPD: chronic obstructive pulmonary disease.

FIGURE 8
RVent without and with additional filtering (*) in the second and third columns in comparison to Jvent (first column) for a female COPD (Gold IV) patient (age = 70). Rows show the three registration variants. The last column shows the expiration MR image for anatomical reference, CT in expiration, and parametric response mapping (PRM). The fourth column demonstrates the alignment of the respective ventilation defects derived from RVent* and JVent to PRM. JVent and RVent showed PRM-aligned VD in the upper lung. RVent was also aligned in the lower parts. ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, RVent*: filtered regional ventilation, JVent: Jacobian determinant ventilation, PRM: parametric response mapping, fSAD: functional small airway disease, VD: ventilation defect, COPD: chronic obstructive pulmonary disease.

when dealing with mean values but disappeared in the median statistic. Defects $VD|QD_{High}$, which were clearly distinguishable from the surrounding parenchyma values, were also more easily identified with JVent. Such regions act as landmarks and probably lead to significant changes upon relocation in the minimization process during registration, resulting in better registration of such regions.

The defect patterns identified in ASYLUM were similarly observed in patients with COPD. Lower lung regions showed much lower defect percentage with JVent in comparison to upper lung regions. In addition, RVent showed defects in the lower lung region with no correspondence in JVent. CT showed higher correspondence of ventilation defects and RVent regarding the number of defects and their regional distribution. Therefore, the observed ventilation defect pattern of RVent and JVent is probably linked to the explanations outlined previously.

The observed strong gradient of decreasing JVent in the superior direction was not reported in studies with hyperpolarized gas MRI, which is considered a gold standard (26, 48) but can be observed in figures showing Jacobian determinant measurements from various studies (8, 10, 11). Similar to findings in our study, Castillo et al. found a good global correspondence of CT HU-derived ventilation in comparison to Jacobian-based methods, but a better correlation of CT HU when comparing on a regional level with SPECT/CT gold standard (49). The less pronounced differences might be explained by the fact that CT is an easier modality to achieve accurate registration as it offers more distinct landmarks due to its higher spatial resolution.

Partially different results and interpretations were reported by Tan et al. in a preprinted article involving six healthy volunteers. The authors found a better correspondence of Jacobian-derived ventilation to segmented lung volumes, concluding a less stable performance of signal-based measurements due to low SNR and registration errors (11). While a global better correspondence to lung volumes does not contradict our results, which mainly indicate problems of JVent on a regional level, in contrast, our results suggest a more stable performance of signal-based methods as discussed previously. In fact, registration errors would affect JVent more directly, by definition. Nevertheless, the authors raise a valid point regarding additional signal variations apart from proton density like T2*, which can affect signal-based methods and lead to errors. However, ultimately, both measurements are always intertwined to a certain degree as they depend on each other and result in identical or nearly identical results in theory or when using perfect registration, as demonstrated by ASYLUM. Although the slightly smaller relative differences (except for the edge artifacts) for RVent and QA favored ANTs over F-REG, F-REG was significantly faster. Overall, similar registration performance was observed for both algorithms, as reported by Klimeš et al. (41).

The idea of using digital models for validation is not new, but to the authors' knowledge, most models stem from radiotherapy and were never employed in the context of functional lung MRI. In general, models can be based on real data (50, 51), be built from scratch (52), or a mixture of both (53–56). Pure patient-based models
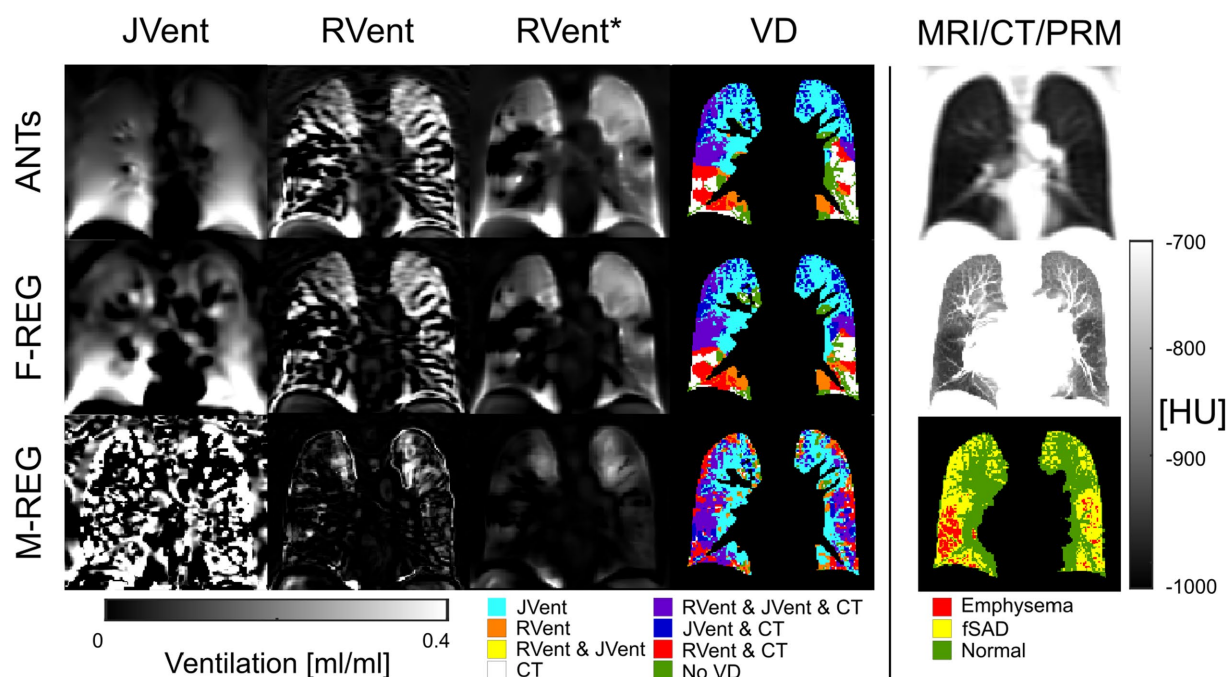
FIGURE 9
RVent without and with additional filtering (*) in the second and third columns in comparison to Jvent (first column) for a female COPD (GOLD III) patient (age = 63). Rows show the three registration variants. The last column shows the expiration image for anatomical reference, CT in expiration, and parametric response mapping (PRM). The fourth column demonstrates the alignment of the respective ventilation defects derived from RVent* and JVent to PRM. Please note that JVent and RVent show inversed VD patterns: RVent detects VD in the lower lung similar to PRM, while JVent shows mainly unmatched defects in the upper lung regions. ANTs: advanced normalization tools, F-Reg: Forsberg registration, M-Reg: Matlab registration, RVent: regional ventilation, RVent*: filtered regional ventilation, JVent: Jacobian determinant ventilation, PRM: parametric response mapping, fSAD: functional small airway disease, VD: ventilation defect, COPD: chronic obstructive pulmonary disease.

exhibit the most realistic data but require expert annotation or fiducial markers (57). Completely modeled variants allow for most control but might be too simplified. One of the more recent models, the 4D CT/MRI Breathing Anthropomorphic Thorax (CoMBAT) phantom (56) encompasses realistic movement and tissue parameters (T1, T2, and proton density) using real MR acquisitions and complex modeling of the acquisition and reconstruction side for organ motion quantification and management in image-guided radiotherapy. The CoMBAT model is much more complex than ASYLUM, but it later offers complete control over the regional movement of the different defect classes, which is crucial for this study and post-processing evaluations. Nevertheless, more elaborated models like CoMBAT might be helpful to create more realistic versions of ASYLUM by adapting certain aspects. In addition to the demonstrated application of examining the difference in registration performance with focus on signal and deformation-based ventilation measurements, ASYLUM can be used for completely different aspects, including (1) comparing similar post-processing methods in their performance (e.g., two signal-based approaches) (2), optimizing parameter settings of certain post-processing aspects like filter settings or registration, and (3) testing which minimal defect sizes can be detected with a certain method. Even slight variations of method implementation (e.g., different programming languages) can be validated by sharing ASYLUM results across different sites.

The limitations of this study include the fact that only a limited number of parameter variations were used for this initial study to remain within a reasonable scope. Results may vary depending on the chosen model parameters, such as expansion rate, SNR, defect severity spectrum (e.g., reduced ventilation instead of no ventilation), and defect size. A further limitation is the simplification of the model with regard to lung shape and the isolated one-dimensional movement of the lung. This might be important, as registration is guided by shapes and structures. Additional details in the lung (e.g., vessels) might lead to better registration results. Therefore, the presented results might underestimate the performance of registration, and further improvements of ASYLUM should address this point. On the other side, the 1D movement of ASYLUM should be easier to register in comparison to real motion, which also involves through-plane motion. In summary, both simplifications act as antagonistic factors regarding registration performance. Similarly, increased or decreased SNR will result in more or less accurate registration and parameter results. Based on the necessity to include additional filtering for the real MR data, the SNR was overestimated in the case of ASYLUM. MR sequence parameters, including TE, TR, and MR physics like relaxation, were not modeled with ASYLUM and therefore limit the model's capability to assess the acquisition aspect and might introduce additional discrepancies to real data. To test the whole capability of ASYLUM, especially in regard to its dynamic components/classes, a phase-sensitive analysis of the whole time series, as in PREFUL, is necessary but was omitted to maintain the concise scope of this study. In addition, registration parameters offer a lot of opportunities for tweaking and can substantially alter the results. For this study,

parameters, which delivered visually acceptable results for real data and were used in previous studies were employed. A systematic fine-tuning of parameters with ASYLUM was not performed and might lead to changes in the findings. In addition, models only approximate reality and require some sort of validation. Therefore, validation with real data can never be replaced completely.

Finally, the used gold standard in the patient cohort (CT PRM) is also dependent on registration algorithms and might contain a bias, and defects derived from the emphysema and fSAD classes do not necessarily correspond to ventilation defects. However, as discussed, the CT modality is probably more likely to achieve accurate registration as it offers more distinct landmarks due to its higher spatial resolution. Furthermore, a high correlation of ventilation defects and emphysema and fSAD was found previously (58).

6 Conclusion

The digital lung model framework ASYLUM was introduced for the validation of free-breathing functional lung MRI post-processing pipelines. As a first scenario, the influence of registration algorithms, two ventilation methods, and one perfusion quantification method was validated. The findings suggest that JVent, as derived from registration methods and parameters evaluated in this study, leads to a significant bias in the regional ventilation calculation and subsequent defect detection. Analysis of patient data and comparison with CT support these findings. Thus, without extensive registration testing and optimization, the use of JVent would result in unreliable defect classifications not suited for clinical/diagnostic decision-making. In contrast, signal-based regional ventilation assessment was a reliable method in the investigated setting.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://sourceforge.net/projects/asylum-repo/files/.

Ethics statement

The studies involving humans were approved by Ethik-Kommission der MHH—OE 9515—Carl-Neuberg-Str. 130625 Hannover. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. MG: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. FK: Conceptualization,

Methodology, Writing – original draft, Writing – review & editing. TK: Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. H-oS: Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. H-UK: Funding acquisition, Project administration, Resources, Writing – original draft, Writing – review & editing. TW: Funding acquisition, Project administration, Resources, Writing – original draft, Writing – review & editing. FW: Funding acquisition, Project administration, Resources, Writing – original draft, Writing – review & editing. JV-C: Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

FK, AV, and JV-C are shareholders of BioVisioneers GmbH, a company, which has interest in pulmonary magnetic resonance imaging methods.

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.

The reviewer HM declared past co-authorship with the authors AV, JV-C, FK, and TK to the handling editor.

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.

References

1. Sharma M, Wyszkiewicz PV, Desaigoudar V, Guo F, Capaldi DP, Parraga G. Quantification of pulmonary functional MRI: state-of-the-art and emerging image processing methods and measurements. *Phys Med Biol.* (2022) 67:22TR01. doi: 10.1088/1361-6560/ac9510

2. Bauman G, Puderbach M, Deimling M, Jellus V, Chefd'hotel C, Dinkel J, et al. Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of Fourier decomposition in proton MRI. *Magn Reson Med.* (2009) 62:656–64. doi: 10.1002/mrm.22031

3. Bauman G, Bieri O. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. *Magn Reson Med.* (2017) 77:336–42. doi: 10.1002/mrm.26096

4. Fischer A, Weick S, Ritter CO, Beer M, Wirth C, Hebestreit H, et al. SElf-gated non-contrast-enhanced FUnctional lung imaging (SENCEFUL) using a quasi-random fast low-angle shot (FLASH) sequence and proton MRI. *NMR Biomed.* (2014) 27:907–17. doi: 10.1002/nbm.3134

5. Mendes Pereira L, Wech T, Weng AM, Kestler C, Veldhoen S, Bley TA, et al. UTE-SENCEFUL: first results for 3D high-resolution lung ventilation imaging. *Magn Reson Med.* (2019) 81:2464–73. doi: 10.1002/mrm.27576

6. Voskrebenzev A, Gutberlet M, Klimeš F, Kaireit TF, Schönfeld C, Rotärmel A, et al. Feasibility of quantitative regional ventilation and perfusion mapping with phase-resolved functional lung (PREFUL) MRI in healthy volunteers and COPD, CTEPH, and CF patients. *Magn Reson Med.* (2018) 79:2306–14. doi: 10.1002/mrm.26893

7. Klimeš F, Voskrebenzev A, Gutberlet M, Kern AL, Behrendt L, Grimm R, et al. 3D phase-resolved functional lung ventilation MR imaging in healthy volunteers and patients with chronic pulmonary disease. *Magnetic resonance in medicine*, (2021) 85:912–925. doi: 10.1002/mrm.28482

8. Kjørstad Å, Corteville DMR, Henzler T, Schmid-Bindert G, Hodneland E, Zöllner FG, et al. Quantitative lung ventilation using Fourier decomposition MRI; comparison and initial study. *Magn Reson Mater Phys Biol Med.* (2014) 27:467–76. doi: 10.1007/s10334-014-0432-9

9. Chassagnon G, Martin C, Marini R, Vakalopolou M, Régent A, Mouthon L, et al. Use of elastic registration in pulmonary MRI for the assessment of pulmonary fibrosis in patients with systemic sclerosis. *Radiology.* (2019) 291:487–92. doi: 10.1148/radiol.2019182099

10. Boucneau T, Fernandez B, Larson P, Darrasse L, Maître X. 3D magnetic resonance spirometry. *Sci Rep.* (2020) 10:9649. doi: 10.1038/s41598-020-66202-7

11. Tan F, Zhu X, Chan M, Deveshwar N, Willmering M, Lustig M, et al. Pulmonary Ventilation Analysis Using 1H Ultra-Short Echo Time (UTE) Lung MRI: A Reproducibility Study. *bioRxiv.* (2023):563196. doi: 10.1101/2023.10.22.563196

12. Lederlin M, Bauman G, Eichinger M, Dinkel J, Brault M, Biederer J, et al. Functional MRI using Fourier decomposition of lung signal: reproducibility of ventilation- and perfusion-weighted imaging in healthy volunteers. *Eur J Radiol.* (2013) 82:1015–22. doi: 10.1016/j.ejrad.2012.12.003

13. Voskrebenzev A, Gutberlet M, Becker L, Wacker F, Vogel-Claussen J. Reproducibility of fractional ventilation derived by Fourier decomposition after adjusting for tidal volume with and without an MRI compatible spirometer. *Magn Reson Med.* (2016) 76:1542–50. doi: 10.1002/mrm.26047

14. Pöhler GH, Klimeš F, Behrendt L, Voskrebenzev A, Gonzalez CC, Wacker F, et al. Repeatability of phase-resolved functional lung (PREFUL)-MRI ventilation and perfusion parameters in healthy subjects and COPD patients. *J Magn Reson Imaging.* (2021) 53:915–27. doi: 10.1002/jmri.27385

15. Klimeš F, Voskrebenzev A, Gutberlet M, Obert AJ, Pöhler GH, Grimm R, et al. Repeatability of dynamic 3D phase-resolved functional lung (PREFUL) ventilation MR imaging in patients with chronic obstructive pulmonary disease and healthy volunteers. *J Magn Reson Imaging.* (2021) 54:618–29. doi: 10.1002/jmri.27543

16. Bauman G, Puderbach M, Heimann T, Kopp-Schneider A, Fritzsching E, Mall MA, et al. Validation of Fourier decomposition MRI with dynamic contrast-enhanced MRI using visual and automated scoring of pulmonary perfusion in young cystic fibrosis patients. *Eur J Radiol.* (2013) 82:2371–7. doi: 10.1016/j.ejrad.2013.08.018

17. Capaldi DPI, Sheikh K, Guo F, Svenningsen S, Etemad-Rezai R, Coxson HO, et al. Free-breathing pulmonary 1H and hyperpolarized 3He MRI: comparison in COPD and bronchiectasis. *Acad Radiol.* (2015) 22:320–9. doi: 10.1016/j.acra.2014.10.003

18. Behrendt L, Voskrebenzev A, Klimeš F, Gutberlet M, Winther HB, Kaireit TF, et al. Validation of automated perfusion-weighted phase-resolved functional lung (PREFUL)-MRI in patients with pulmonary diseases. *J Magn Reson Imaging.* (2020) 52:103–14. doi: 10.1002/jmri.27027

19. Kaireit TF, Kern A, Voskrebenzev A, Pöhler GH, Klimes F, Behrendt L, et al. Flow volume loop and regional ventilation assessment using phase-resolved functional lung (PREFUL) MRI: comparison with 129 xenon ventilation MRI and lung function testing. *J Magn Reson Imaging JMRI.* (2021) 53:1092–105. doi: 10.1002/jmri.27452

20. Marshall H, Voskrebenzev A, Smith LJ, Biancardi AM, Kern AL, Collier GJ, et al. 129Xe and free-breathing 1H ventilation MRI in patients with cystic fibrosis: a dual-center study. *J Magn Reson Imaging.* (2023) 57:1908–21. doi: 10.1002/jmri.28470

21. Vogel-Claussen J. Functional lung MRI: deep learning turns proton into helium ventilation maps—the Battle is on! *Radiology.* (2021) 298:439–40. doi: 10.1148/radiol.2020204069

22. Hughes PJC, Smith L, Chan HF, Tahir BA, Norquay G, Collier GJ, et al. Assessment of the influence of lung inflation state on the quantitative parameters derived from hyperpolarized gas lung ventilation MRI in healthy volunteers. *J Appl Physiol.* (2019) 126:183–92. doi: 10.1152/japplphysiol.00464.2018

23. Murphy K, van Ginneken B, Reinhardt JM, Kabus S, Ding K, Deng X, et al. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. *IEEE Trans Med Imaging.* (2011) 30:1901–20. doi: 10.1109/TMI.2011.2158349

24. Klimeš F, Voskrebenzev A, Gutberlet M, Kern A, Behrendt L, Kaireit TF, et al. Free-breathing quantification of regional ventilation derived by phase-resolved functional lung (PREFUL) MRI. *NMR Biomed.* (2019) 32:e4088. doi: 10.1002/nbm.4088

25. Voskrebenzev A, Kaireit TF, Klimeš F, Pöhler GH, Behrendt L, Biller H, et al. PREFUL MRI depicts dual bronchodilator changes in COPD: a retrospective analysis of a randomized controlled trial. *Radiol Cardiothorac Imaging.* (2022) 4:e210147. doi: 10.1148/ryct.210147

26. Marshall H, Voskrebenzev A, Biancardi A, Smith L, Tahir B, Wild J, et al. Comparison of 1H MRI and 3He MRI ventilation images in patients with cystic fibrosis and patients with lung cancer. *Eur Respir J.* (2018) 52. doi: 10.1183/13993003.congress-2018.PA387

27. Zapke M, Topf HG, Zenker M, Kuth R, Deimling M, Kreisler P, et al. Magnetic resonance lung function – a breakthrough for lung imaging and functional assessment? A phantom study and clinical trial. *Respir Res.* (2006) 7. doi: 10.1186/1465-9921-7-106

28. Wild JM, Marshall H, Bock M, Schad LR, Jakob PM, Puderbach M, et al. MRI of the lung (1/3): methods. *Insights Imaging.* (2012) 3:345–53. doi: 10.1007/s13244-012-0176-x

29. Deimling M, Jellus V, Geiger B, Chefd'hotel C. Time resolved lung ventilation imaging by Fourier decomposition. *Proc Intl Soc Mag Reson Med.* (2008) 16:2639.

30. Kjørstad Å, Corteville DMR, Fischer A, Henzler T, Schmid-Bindert G, Zöllner FG, et al. Quantitative lung perfusion evaluation using fourier decomposition perfusion MRI. *Magn Reson Med.* (2014) 72:558–62. doi: 10.1002/mrm.24930

31. Glandorf J, Klimeš F, Behrendt L, Voskrebenzev A, Kaireit TF, Gutberlet M, et al. Perfusion quantification using voxel-wise proton density and median signal decay in PREFUL MRI. *Magn Reson Med.* (2021) 86:1482–93. doi: 10.1002/mrm.28787

32. Karch A, Vogelmeier C, Welte T, Bals R, Kauczor HU, Biederer J, et al. The German COPD cohort COSYCONET: aims, methods and descriptive analysis of the study population at baseline. *Respir Med.* (2016) 114:27–37. doi: 10.1016/j.rmed.2016.03.008

33. Breuer FA, Kellman P, Griswold MA, Jakob PM. Dynamic autocalibrated parallel imaging using temporal GRAPPA (TGRAPPA). *Magn Reson Med.* (2005) 53:981–5. doi: 10.1002/mrm.20430

34. Lujan AE, Larsen EW, Balter JM, Ten Haken RK. A method for incorporating organ motion due to breathing into 3D dose calculations. *Med Phys.* (1999) 26:715–20. doi: 10.1118/1.598577

35. Dietrich O, Raya JG, Reeder SB, Ingrisch M, Reiser MF, Schoenberg SO. Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. *Magn Reson Imaging.* (2008) 26:754–62. doi: 10.1016/j.mri.2008.02.001

36. 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

37. Voskrebenzev A, Gutberlet M, Kaireit TF, Wacker F, Vogel-Claussen J. Low-pass imaging of dynamic acquisitions (LIDA) with a group-oriented registration (GOREG) for proton MR imaging of lung ventilation. *Magn Reson Med.* (2017) 78:1496–505. doi: 10.1002/mrm.26526

38. Ljimani A, Hojdis M, Stabinska J, Valentin B, Frenken M, Appel E, et al. Analysis of different image-registration algorithms for Fourier decomposition MRI in functional lung imaging. *Acta Radiol.* (2021) 62:875–81. doi: 10.1177/0284185120944902

39. Forsberg D, Andersson M, Knutsson H. Extending image registration using polynomial expansion to diffeomorphic deformations. In: SSBA Symposium on Image Analysis. (2012):4.

40. Forsberg D. Fordanic/image-registration. Published online 2022. Accessed April 12, 2022. Available at: https://github.com/fordanic/image-registration

41. Klimeš F, Voskrebenzev A, Gutberlet M, Grimm R, Wacker F, Vogel-Claussen J. Evaluation of image registration algorithms for 3D phase-resolved functional lung ventilation magnetic resonance imaging in healthy volunteers and chronic obstructive pulmonary disease patients. *NMR Biomed.* (2023) 36:e4860. doi: 10.1002/nbm.4860

42. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage.* (2009) 45:S61–72. doi: 10.1016/j.neuroimage.2008.10.040

43. He K, Sun J, Tang X. Guided image filtering. *IEEE Trans Pattern Anal Mach Intell.* (2013) 35:1397–409. doi: 10.1109/TPAMI.2012.213

44. Vogel-Claussen J, Kaireit TF, Voskrebenzev A, Klimeš F, Glandorf J, Behrendt L, et al. Phase-resolved functional lung (PREFUL) MRI–derived ventilation and perfusion parameters predict future lung transplant loss. *Radiology.* (2023) 307:e221958. doi: 10.1148/radiol.221958

45. Galbán CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, et al. Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. *Nat Med.* (2012) 18:1711–5. doi: 10.1038/nm.2971

46. Göçken M, Özçalıcı M, Boru A, Dosdoğru AT. Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Syst Appl.* (2016) 44:320–31. doi: 10.1016/j.eswa.2015.09.029

47. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal.* (1998) 2:243–60. doi: 10.1016/S1361-8415(98)80022-4

48. Smith LJ, Horsley A, Bray J, Hughes PJC, Biancardi A, Norquay G, et al. The assessment of short- and long-term changes in lung function in cystic fibrosis using 129Xe MRI. *Eur Respir J.* (2020) 56:2000441. doi: 10.1183/13993003.00441-2020

49. Castillo R, Castillo E, Martinez J, Guerrero T. Ventilation from four-dimensional computed tomography: density versus Jacobian methods. *Phys Med Biol.* (2010) 55:4661–85. doi: 10.1088/0031-9155/55/16/004

50. Aubert-Broche B, Griffin M, Pike GB, Evans AC, Collins DL. Twenty new digital brain phantoms for creation of validation image data bases. *IEEE Trans Med Imaging.* (2006) 25:1410–6. doi: 10.1109/TMI.2006.883453

51. Vandemeulebroucke J, Sarrut D, Clarysse P. The POPI-model, a point-validated pixel-based breathing thorax model. *In Proceeding of the XVth ICCR Conference, Toronto, Canada.* (2007).

52. Pretorius PH, Xia W, King MA, Tsui BMW, Pan T, Villegas BJ. Determination of left and right ventricular volume and ejection fraction using a mathematical cardiac torso phantom for gated blood pool SPECT. *J Nucl Med.* (1996) 37:97P.

53. Segars WP, Lalush DS, Tsui BMW. A realistic spline-based dynamic heart phantom. In: 1998 IEEE Nuclear Science Symposium Conference Record. 1998 IEEE Nuclear Science Symposium and Medical Imaging Conference (Cat. No.98CH36255). Vol 2.; (1998):1175–1178

54. Sharif B, Bresler Y. Adaptive real-time cardiac MRI using paradise: validation by the physiologically improved NCAT phantom. In: 2007 4th IEEE international symposium on Biomedical imaging: From Nano to macro. (2007):1020–1023.

55. Nie K, Chuang C, Kirby N, Braunstein S, Pouliot J. Site-specific deformable imaging registration algorithm selection using patient-based simulated deformations. *Med Phys.* (2013) 40:041911. doi: 10.1118/1.4793723

56. Paganelli C, Summers P, Gianoli C, Bellomi M, Baroni G, Riboldi M. A tool for validating MRI-guided strategies: a digital breathing CT/MRI phantom of the abdominal site. *Med Biol Eng Comput.* (2017) 55:2001–14. doi: 10.1007/s11517-017-1646-6

57. Madan H, Pernuš F, Likar B, Špiclin Ž. A framework for automatic creation of gold-standard rigid 3D–2D registration datasets. *Int J Comput Assist Radiol Surg.* (2017) 12:263–75. doi: 10.1007/s11548-016-1482-4

58. Voskrebenzev A, Kaireit TF, Gutberlet M, Klimeš F, Behrendt L, Hoen-Oh S, et al. Detection of emphysema using phase-resolved functional lung imaging parametric response mapping (PREFUL PRM). *In: Proc Intl Soc Mag Reson Med.* (2023) 31:4669.

Glossary

A	Signal amplitude		REF	Reference (known) registration
ANTs	Advanced normalization tools registration package		Reg	Fixed respiration state
ASYLUM	A synthetic lung model		RMSRE	Root mean squared relative error
BW	Bandwidth		ROI	Region of interest
C	Condition for ventilation compensation		RVent	Signal-based regional ventilation measurement
CoMBAT	4D CT/MRI breathing anthropomorphic thorax		s	Signal
CT	Computed tomography		S	Scattered interpolant
e	Expansion factor		s(t)	signal time-series
E	Local expansion matrix		SENCEFUL	Self-gated noncontrast-enhanced functional lung
Exp	Expiration		SNR	Signal-to-noise ratio
FOV	Field-of-view		SPECT	Single-photon emission computed tomography
F-REG	Forsberg registration package		TE	Echo time
$F_{x,y}$	Forward deformation field		TOF	Time-of-flight effect
G	Image geometry		TR	Repetition time
HU	Hounsfield units		UL	Upper-left quadrant
i,j	Irregular grid		UR	Upper-right quadrant
Insp	Inspiration		v	Volume
$I_{x,y}$	Inverse deformation field		V	Ventilation
j	Phase		v(t,x)	volume surrogate voxel time-series
JVent	Deformation-based regional ventilation measurement		VD	Ventilation defect
LL	Lower-left quadrant		$VD|QD_{High}$	Ventilation and perfusion defect at high signal
LR	Lower-right quadrant		$VD|QD_{Insp}$	Ventilation and perfusion defect at inspiration signal
MBW	Multiple breath wash out		$VD|QD_{Low}$	Ventilation and perfusion defect at signal below inspiration level
M-REG	A diffeomorphic demons registration algorithm implemented in MATLAB		VV	Ventilated volume
MRI	Magnetic Resonance Imaging		VV|QD	Ventilated perfusion defect
PREFUL	Phase-resolved functional lung		VV|QV	Normally ventilated and perfused voxels
PRM	Parametric Response Mapping		$VV|QV_{Delay}$	Voxel with normal ventilation but delayed perfusion
Q	Perfusion		$VV_{Comp}|QV$	Compensatory ventilated and normally perfused voxels
QA	Signal-based perfusion-weighted amplitude measurement		$VV_{Delay}|QV$	Delayed ventilation and normally perfused voxels
QD	Perfusion defect		x,y	Regular grid
r(t,x)	Respiration factor derived from v(t,x)		•	Warping operator

Automated lung segmentation on chest MRI in children with cystic fibrosis

Friedemann G. Ringwald[1,2], Lena Wucherpfennig[2,3,4],
Niclas Hagen[1,2], Jonas Mücke[1], Sebastian Kaletta[1],
Monika Eichinger[2,3,4], Mirjam Stahl[2,5,6,7], Simon M. F. Triphan[2,3,4],
Patricia Leutz-Schmidt[2,3,4], Sonja Gestewitz[2,3,4],
Simon Y. Graeber[2,5,6,7], Hans-Ulrich Kauczor[2,3,4],
Abdulsattar Alrajab[3], Jens-Peter Schenk[3], Olaf Sommerburg[2,8,9],
Marcus A. Mall[2,5,6,7], Petra Knaup[1,2], Mark O. Wielpütz[2,3,4] and
Urs Eisenmann[1,2]*

[1]Institute of Medical Informatics, Heidelberg University, Heidelberg, Germany, [2]Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), Heidelberg, Germany, [3]Department of Diagnostic and Interventional Radiology, University Hospital Heidelberg, Heidelberg, Germany, [4]Department of Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany, [5]Department of Pediatric Respiratory Medicine, Immunology and Critical Care Medicine, Charité-Universitätsmedizin Berlin, Berlin, Germany, [6]German Center for Lung Research (DZL), Associated Partner Site, Berlin, Germany, [7]Berlin Institute of Health (BIH) at Charité-Universitätsmedizin Berlin, Berlin, Germany, [8]Division of Pediatric Pulmonology & Allergy and Cystic Fibrosis Center, Department of Pediatrics, University Hospital Heidelberg, Heidelberg, Germany, [9]Department of Translational Pulmonology, University Hospital Heidelberg, Heidelberg, Germany

Introduction: Segmentation of lung structures in medical imaging is crucial for the application of automated post-processing steps on lung diseases like cystic fibrosis (CF). Recently, machine learning methods, particularly neural networks, have demonstrated remarkable improvements, often outperforming conventional segmentation methods. Nonetheless, challenges still remain when attempting to segment various imaging modalities and diseases, especially when the visual characteristics of pathologic findings significantly deviate from healthy tissue.

Methods: Our study focuses on imaging of pediatric CF patients [mean age, standard deviation (7.50 ± 4.6)], utilizing deep learning-based methods for automated lung segmentation from chest magnetic resonance imaging (MRI). A total of 165 standardized annual surveillance MRI scans from 84 patients with CF were segmented using the nnU-Net framework. Patient cases represented a range of disease severities and ages. The nnU-Net was trained and evaluated on three MRI sequences (BLADE, VIBE, and HASTE), which are highly relevant for the evaluation of CF induced lung changes. We utilized 40 cases for training per sequence, and tested with 15 cases per sequence, using the Sørensen-Dice-Score, Pearson's correlation coefficient (r), a segmentation questionnaire, and slice-based analysis.

Results: The results demonstrated a high level of segmentation performance across all sequences, with only minor differences observed in the mean Dice coefficient: BLADE (0.96 ± 0.05), VIBE (0.96 ± 0.04), and HASTE (0.95 ± 0.05). Additionally, the segmentation quality was consistent across different disease severities, patient ages, and sizes. Manual evaluation identified specific challenges, such as incomplete segmentations near the diaphragm and dorsal

regions. Validation on a separate, external dataset of nine toddlers (2–24 months) demonstrated generalizability of the trained model achieving a Dice coefficient of 0.85 ± 0.03.

Discussion and conclusion: Overall, our study demonstrates the feasibility and effectiveness of using nnU-Net for automated segmentation of lung halves in pediatric CF patients, showing promising directions for advanced image analysis techniques to assist in clinical decision-making and monitoring of CF lung disease progression. Despite these achievements, further improvements are needed to address specific segmentation challenges and enhance generalizability.

1 Introduction

Cystic fibrosis (CF) is an inherited multi-organ disease, which largely effects the lungs. Repeated bacterial infections and inflammation can result in lung damage, causing most of the morbidity and mortality seen in CF (1, 2). The early detection and monitoring of CF-related lung disease is a prerequisite for optimized care and improved long-term outcomes (3–7).

Recently, chest magnetic resonance imaging (MRI), a radiation-free modality, has shown great promise in assessing structural and functional CF lung abnormalities. Studies have shown chest MRI can detect changes as early as in infancy, and is capable of monitoring disease progression and therapeutic response throughout adulthood (7–14).

To semi-quantitatively assess the severity of lung abnormalities in CF patients, a morpho-functional chest MRI scoring system, also referred to as the Eichinger Score, was developed in 2012 (15). This scoring system includes items for morphological lung abnormalities, as well as perfusion abnormalities (8–11, 15, 16). To automate this scoring process, a critical step is automating the lung segmentation process.

In medical imaging, segmentation refers to identifying an organ or specific tissue of interest by extracting the boundaries and the inner region. This process allows for downstream analysis and extracting important quantitative information within that region. Precise segmentation may support accurate decisions on diagnosis, treatment plans, disease monitoring, and guiding of interventions (17). In the last decade, automated segmentation methods improved in performance and precision, resulting in the possibility of fully automated segmentation in different medical disciplines and imaging modalities (18). Machine learning methods, particularly neural networks, have demonstrated remarkable performance, often outperforming conventional methods, especially when analyzing large datasets (19–21). The nnU-Net, an advanced deep learning framework tailored for medical applications, stands out in its performance (22). It permits the training of networks to perform semantic segmentation with high accuracy and performance, eliminating the need for numerous configuration steps due to its self-configuring training parameters and layer settings. However, difficulties arise when attempting to adapt the nnU-Net to a variety of imaging modalities and diseases. This is particularly challenging when the visual characteristics of pathologic findings deviate significantly from healthy tissue, indicating a change in tissue composition within the same organ (23).

Automated lung segmentation in MR images, especially in the CF population, also have inherent challenges. In MRI, difficulties arise due to the limited spatial resolution and the low contrast between the lungs and the adjacent tissue. In CF patients, breathing artifacts, most notably in young children; cardiac pulsation artifacts; chest growth in children, lung abnormalities displacing air contents, and the deformation associated with disease progression, all contribute to the complexity of the segmentation task (8, 24).

Despite these challenges, many studies are beginning to show promising results incorporating neural networks to automate MRI lung segmentation, even in different underlying pathologies, replacing conventional segmentation approaches (25, 26). Zha et al. applied convolutional neural networks (CNNs) on 3D radial ultra-short echo-times (UTE) oxygen-enhanced MRI in a dataset of 45 subjects (age 10+ with CF, asthma, or healthy) and achieved Dice coefficients of 0.97 and 0.96 for the right and left lung, respectively (27). Furthermore, researchers tested other MRI sequences, such as fast UTE with stack-of-spirals trajectory and matrix pencil decomposition MRI, in CF patients (age 5+) yielding Dice coefficients of 0.96 for children and 0.89 for adults (28, 29).

Notably, Astley et al. tested 2D and 3D nnU-Nets for lung segmentation of patients with varying pulmonary pathologies. In their patient cohort (median age 34 yrs.), analysis of a dataset comprising 809 spoiled-gradient-recalled and UTE MRI scans, even across different vendors, demonstrated a remarkable performance, reaching a median Dice coefficient of 0.96 internally and 0.97 on an external test set (30). To enhance the accuracy of automated lung segmentation, by inclusion of artificially generated images with consolidations, Cristoso et al. reached a Dice coefficient of 0.94 on a cohort of healthy volunteers and patients (31). In a 2023 study of neonates, either healthy or suffering from bronchopulmonary dysplasia, the authors employed CNNs for lung segmentation on quiet-breathing MRI and achieved a Dice coefficient of 0.908 on an internal test set and 0.88 on an independent test set (32). Most recently, a new approach for lung segmentation on healthy adults using thresholding and clustering on an enhanced deep-inspiration-breath-hold reached a Dice coefficient of 0.94 (33, 34). A high benchmark for lung lobe segmentation using pseudo-MRI images derived from CT and three concatenated CNNs achieved a Dice coefficient of 0.95 on a dataset of 100 CF patients over the age of 4.7 years old (35).

To the best of our knowledge, we are the first to demonstrate pediatric lung half segmentations for patients across the entire pediatric age range with different stages of cystic fibrosis using chest MRI on the commonly used sequences BLADE, VIBE, and HASTE. We selected a total of 165 MRI examinations from 84 patients in our internal monocentric CF database. This database contains 1,312 highly standardized annual surveillance MRIs, acquired over more than a decade from 266 patients. Segmentations were created manually by three observers.

2 Materials and methods

2.1 Study population

This ongoing prospective longitudinal observational study (clinicaltrials.gov identifiers NCT00760071, NCT02270476) was approved by the institutional ethics committee and informed written consent was obtained from the parents or legal guardians of all patients. The CF diagnosis was confirmed by increased sweat chloride (Cl-) concentrations ($\geq$60 mmoL/L) and cystic fibrosis transmembrane conductance regulator (CFTR) mutation analysis. In pancreatic-sufficient patients with borderline sweat test results (sweat Cl- 30–60 mmoL/L), the diagnosis was further supported by assessing CFTR function in rectal biopsies, as previously described (36). We included 165 cases in the study. Some patients were included in our previous reports on morpho-functional MRI (8, 37–39).

2.2 Magnetic resonance imaging

We performed standardized chest MRI after the initial CF diagnosis or after referral to our center as early as at the age of 3 months. We repeated exams annually using two 1.5 T scanner models from the same manufacturer (Magnetom Symphony and Magnetom Avanto, Siemens Healthcare, Erlangen, Germany). We kept the scanning protocol constant during the study period, apart from minor updates to new software versions as previously described (8–16, 24). We acquired T1-weighted sequences before and after intravenous application of contrast material and T2-weighted sequences before contrast. Children aged 5 years and younger were routinely sedated with oral or rectal chloral hydrate (100 mg/kg body weight, maximum dose of 2 g).

2.3 Staging CF lung disease

One observer (MOW) with more than 15 years of experience in chest MRI, who also evaluated all previous studies, assessed all MRI examinations using the established chest MRI scoring system (8–11, 13–15, 40). The MRI scoring system assigns a numerical disease severity score to each lobe (e.g., 0 = no presence, 1 = <50% of a lobe affected, and 2 = $\geq$50% of a lobe affected) for each of the morphological score items bronchiectasis/wall thickening, mucus plugging, sacculation/abscess, consolidation, and special finding/pleural lesion, as well as for perfusion abnormalities. The sum of morphological findings becomes the MRI morphology score, perfusion abnormalities create the MRI perfusion score, and the sum of both results in the MRI global score, ranging from 0 to 72.

TABLE 1 MRI sequence details.

	BLADE	VIBE	HASTE
Slice thickness (mm)	4	4	6
Pixel spacing in plane (mm) (min-max)	0.9375*0.9375 - 1.25*1.25	0.78125*0.78125 - 0.879*0.879	0.839*0.839 - 1.875*1.875
Matrix (min-max)	320*320 - 384*384	512*512	512*512

Since image acquisition protocols changed slightly over the years, pixel spacing and matrix size have different min/max values.

2.4 Image sequence selection

Three MRI sequences in coronal orientation were used (Table 1):

1. Balanced Steady State Free Precession Line Acquisition with Undersampling (BLADE): This is a T2-weighted turbo spin echo-based 2D sequence designed to reduce motion artifacts in MRI. It is particularly useful for imaging areas of the body that are prone to movement, like the lungs, or for imaging patients who have difficulty remaining still (41). Its acquisition can be split among multiple breath-holds (i.e., slices are not necessarily at the same depth of inspiration) or triggered using a navigator signal.
2. Volumetric Interpolated Breath-Hold Examination (VIBE): This is a T1-weighted 3D gradient echo sequence acquired after injection of a contrast agent. It was acquired in a single breath-hold and allows for high spatial resolution (42).
3. Half-Fourier Acquisition Single-Shot Turbo Spin-Echo (HASTE): This is a T2-weighted turbo spin echo 2D sequence that acquires each slice from a single echo train, minimizing motion effects at the cost of noticeable blurring in the phase encoding direction (43).

2.5 Dataset composition

From our database with 1,312 CF examinations from 266 patients, we selected 55 examinations for each MRI sequence (BLADE, VIBE, and HASTE), resulting in an overall 165 examinations from 84 patients (Figure 1). All cases were chosen to ensure an even distribution of age and gender, and to include varying levels of disease severity based on the global MRI score. To achieve this, the overall distribution of age, gender and disease severity was visualized and cases were then selected manually. From this overall dataset with 165 cases, 45 cases (15 for each sequence) were selected in a stratified manner, to represent the underlying distribution of age, gender, and global MRI score for the creation of the internal test set. This internal test set was not used for training, and solely utilized to test the final performance of the networks. In the internal test set, the median age was 9 years (± 4.92) (range 2 months–17 years) (Table 2, internal test set) with 46.7% male cases. The remaining cases were used for training the neural networks in the so-called training set. The training set had a median age of 9 years (± 4.78) (range 2 months–17 years) and 49.1% male cases (Table 3;
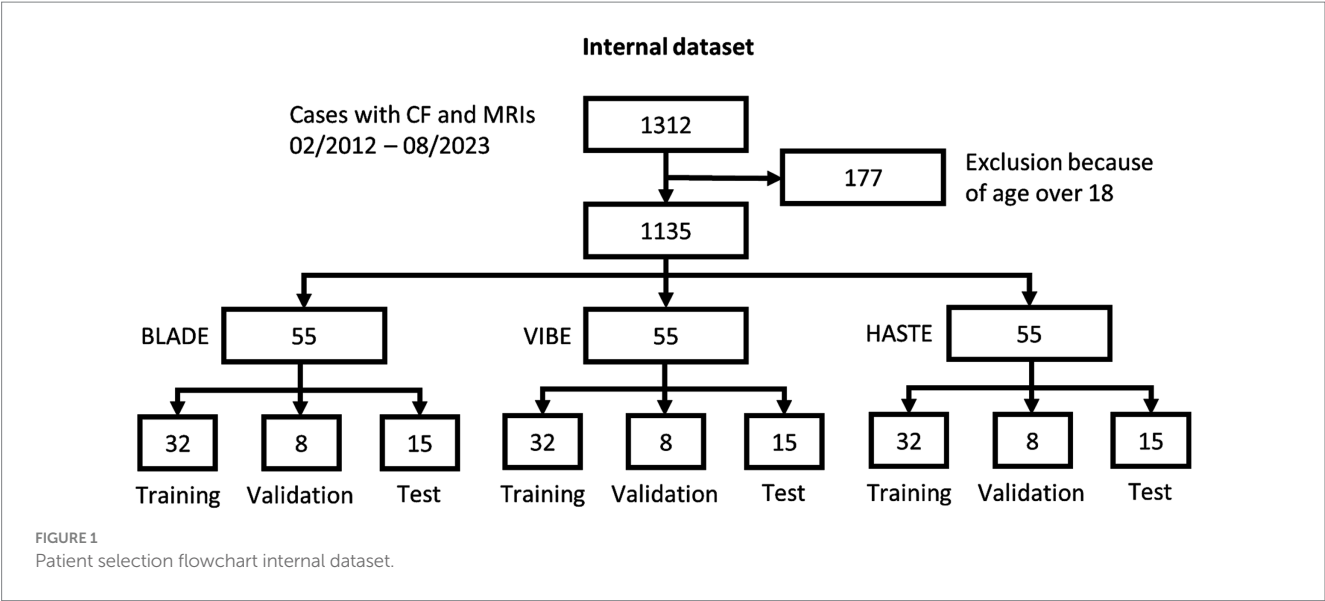
FIGURE 1
Patient selection flowchart internal dataset.

TABLE 2 Overview of the internal and external test sets.

	Internal test set	External test set
Cases, *n*	45	9
Age (years) median [range] (std)	9 [0.13–17.1] ± 4.92	0.79 [0.16–2.0] ± 0.63
Sex (m/f)	21 / 24	4/5
Height (cm) median [range] (std)	133.45 [93.0–174.8] ± 24.43	75.16 [62.0–83.0] ± 7.73
Weight (kg) median [range] (std)	28.03 [12.4–55.9] ± 14.65	9.2 [6.0–10.6] ± 1.73
CF*TR* genotype, *n* (%)	38 (84)	4 (44)
F508del/F508del	18 (40)	2 (22)
F508del/other	15 (33)	3 (33)
Other/other	4 (9)	
Pancreatic insufficiency, *n* (%)	31 (67)	9 (100)
Spirometry, *n* (%)	29 (65)	0 (0)
ppFEV1 median [range] (std)	97.7 [62.2–123.7] ± 15.8	-
Multiple breath washout, *n* (%)	26 (58)	6 (67)
LCI N₂ median [range] (std)	7.74 [4.74–10.57] ± 1.4	7.8 [6.58–9.05] ± 1.2
Global MRI Score median [range] (std)	12 [0–43] ± 9.8	10.22 [7–16] ± 2.72

TABLE 3 Patient characteristics of the internal dataset (training set).

	Internal training set
Cases, *n*	120
Age (years) median [range] (std)	9 [0.16–17.0] ± 4.78
Sex (m/f)	59/61
Height (cm) median [range] (std)	135 [52.7–175.3] ± 24.80
Weight (kg) median [range] (std)	28.2 [12.8–70.9] ± 16.05
CF*TR* genotype, *n* (%)	108 (90)
F508del/F508del	44 (37)
F508del/other	48 (40)
Other/other	12 (10)
Pancreatic insufficiency, *n* (%)	90 (75)
Spirometry, *n* (%)	46 (38)
ppFEV1 median [range] (std)	91.58 [42.8–108.5] ± 16.3
Multiple breath washout, *n* (%)	45 (37.5)
LCI N₂ median [range] (std)	8.28 [3.91–15.3] ± 2.46
Global MRI score median [range] (std)	12 [0–39] ± 7.83

notably younger, as HASTE is a contrast agent-free alternative to VIBE. The comparison between the internal training and test data revealed no statistically significant changes in either age ($p = 0.06$) or global MRI score ($p = 0.97$). The available data from all three sequences resulted in 6,010 2D slices. The training set comprised a total of 4,290 slices, with an average of 34 slices for BLADE, 49 slices for VIBE and 24 slices for HASTE per MRI. In the internal test set, a total of 1,720 slices were used, with an average of 36 slices for BLADE, 52 slices for VIBE, and 27 slices for HASTE.

Additionally, we collected an external dataset from two different centers (Center A: one case, Center B: eight cases) comprising nine HASTE acquisitions from nine cases (Table 2) (44). Compared to the internal dataset, the age distribution of the external test set was dominated by very young patients (Table 2). This was reflected in the statistically significant difference in age between the internal and
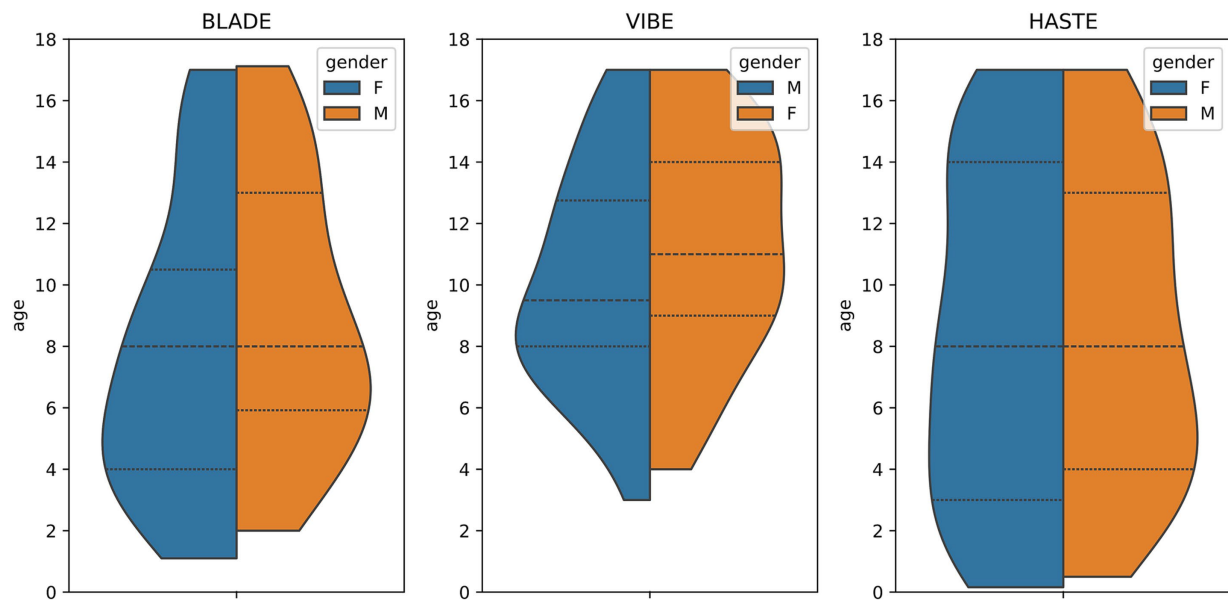
Figures 1, 2). Selecting the cases in such a way may allow good segmentations over all age and disease classes. We included cases only if relevant image data were available. No cases were excluded due to artifacts or poor quality. With regard to similarity, no notable differences were observed in the global MRI score across the three utilized sequences ($p = 0.78$). A notable difference in age was observed between the three sequences ($p = 0.006$). Patients undergoing imaging using HASTE were

FIGURE 2
Violin plot of age and gender distribution for the three MRI sequences.

external datasets ($p < 0.0005$). Regarding disease severity, however, the external patients were similar to those of the internal dataset. A comparison of the global MRI score exhibited no notable differences ($p = 0.86$). The external test set had an overall number of 174 2D slices with an average of 19 slices per MRI.

2.6 Segmentation ground truth generation

We manually segmented all MRIs using three independent observers with 1 (JM), 2.5 (FGR), and 5 years (LW) experience in lung MRI segmentation, respectively. They created reference segmentations of the lung halves using the open-source software Medical Imaging Interaction Toolkit (MITK, version 2021.10) in combination with a Wacom Cintiq 16 tablet and pen. In the event of disagreement among observers, agreement was reached by individual comparison, collective discussion and consensus among the observers.

2.7 Segmentation questionnaire

In cooperation with the experienced radiologists, we designed a qualitative questionnaire for fine-grained evaluation of the segmentations (Supplementary Figure 1). The questionnaire evaluated the overall segmentation quality on an 11-point Likert-scale (45), ranging from 0 (worst quality) to 10 (best quality). Furthermore, the questionnaire included a detailed evaluation of the lung segmentation and specific information regarding the segmentation performance in specific anatomical regions (ventrally, dorsally, mediastinum, periphery, apex, and diaphragm). Information on incomplete segmentation or over-segmentation in specific areas could be provided. Lastly, the observer was given the opportunity to provide an open text response to the segmentation.

2.8 Slice based qualitative analysis

To gain further insight into the segmentation quality, all lungs from the internal test set were subjected to a detailed examination by a radiologist to identify any instances of incorrect segmentation. For each lung, the number of slices requiring correction was annotated. In conjunction with the data on the overall slices, this provides an indication of the quantity of usable slices. The number of slices requiring correction is reported as a mean percentage, with standard deviation and maximum.

2.9 nnU-Net implementation

The latest implementation of the 2D nnU-Net (Version 2) was utilized in its default configuration. It is a self-configuring framework, which automatically adapts its architecture, pre-processing, and training pipeline to a given dataset. The nnU-Net framework employs a U-Net-based architecture comprising an encoder-decoder structure. On the encoder path, the spatial dimensions of the input image are successively reduced through convolutional layers and max-pooling, thereby capturing increasingly abstract feature representations. On the decoder path, upsampling is applied to restore the spatial dimensions, concatenating feature maps from the corresponding encoder layers. This allows for high-level semantic information and precise localization. For further details to the nnU-Net, please refer to (22). Three individual nnU-Net configurations were trained, one for each sequence using the following steps: based on the 55 study cases per sequence, the data were partitioned into 58% as training set, 14% as validation set, and 27% as test set. This resulted in 32 cases being used for training, eight cases for validation and 15 cases for testing per sequence. The training and validation sets were utilized for the initial and fine-tuning training of the neural

network, while the test set was withheld for final evaluation. To ensure an average performance, we validated the models utilizing 5-fold cross-validation with different training and validation set partitions as per default nnU-Net configuration. All calculations were performed on two Tesla V100S PCIe 32GB with 1,000 epochs and an average run time of 33 h per fold. The batch size was subject to variation during the training phase, with values of 14, 32, and 33, respectively, being applied to BLADE, VIBE, and HASTE. Stochastic gradient descent was employed for optimization purposes, with a weight decay of $3e^{-5}$ and an initial learning rate of 0.01. Z-score normalization was utilized as the normalization method. For inference and producing the final predictions the nnU-Net uses an ensemble of all five folds, reporting one final result for the test set. For external validation a separate dataset was utilized. This external dataset was chosen to simulate real-world scenarios and challenges, ensuring a comprehensive examination of the model's performance across diverse imaging conditions.

2.10 Statistical analyses

A one-way ANOVA test was used to determine if there were statistically significant differences in global MRI score and ages among the different datasets. Results were considered significant at $p < 0.05$.

Furthermore, the Sørensen-Dice-Score (DSC), calculated from the spatial overlap between the ground truth segmentation (GT) and predicted segmentation (PS), was utilized to evaluate the entire MRI sequence (46). The DSC ranges from 0 to 1, evaluating the quality of the segmentation indicated by the overlap and is defined as follows:

$$DSC = \frac{2 \left| PS \cap GT \right|}{\left| PS \right| + \left| GT \right|}$$

First, the Dice coefficient was calculated between each manual segmentation and predicted mask, and subsequently, the mean value was obtained for the entire stack of slices. This process was conducted for both the right and left lungs, as well as for the combination of both lung halves.

Data were analyzed with Python (Version 3.9) using the package SciPy (Version 1.11.4) (47). The Pearson correlation coefficient, indicating strength of linear relationship, was calculated for the DSC vs. age and DSC vs. the global MRI score (48). In general, the Pearson correlation coefficient measures the linear correlation of two sets of data and is defined as:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Since the Sørensen-Dice-Score does not provide any indications regarding the location of incorrect segmentations or crucial errors, we deployed an additional questionnaire, which was filled out once for each internal case. To assess the generalizability and robustness of lung segmentation, we conducted an evaluation using an external dataset distinct from the training and validation sets. Due to data availability, only the HASTE model was tested.

3 Results

3.1 Internal and external test set demographics

A total of 45 cases were utilized for the internal test set, while the external test set consisted of nine cases from two distinct centers. The cases from the external dataset are notably younger, with a median age of 0.79 years, whereas the internal test set had a median age of 9 years (Table 2). With regard to the global MRI score, the external dataset exhibited a slightly lower median of 10.22, as compared to the internal test set, which had a median global MRI score of 12.

3.2 BLADE, HASTE, and VIBE are equally well suited for nnU-Net training

Using VIBE and BLADE, the nnU-Net achieved a mean DSC of 0.96 (Table 4; Figures 3, 4). HASTE demonstrated comparable performance with a mean DSC of 0.95. For the BLADE sequence, the right lung exhibited slightly superior segmentation, whereas both lungs demonstrated equivalent performance in the VIBE sequence. On the HASTE sequence, the left lung reached a higher DSC compared to the right lung with a DSC of 0.96 and 0.93, respectively (Table 4).

3.3 Questionnaire confirms segmentation quality

Our analysis of the questionnaire for the 45 internal test cases showed similar results to the overall high Dice coefficients. The segmentations derived from all three sequences were evaluated with a median score of nine out of 10 points (9/10) on the Likert Scale, with a standard deviation of 2.02, 1.48, and 1.18 for BLADE, VIBE, and HASTE, respectively (Figure 5). Additional information can be found in Supplementary Figures 2, 3. In addition to the quality of the segmentation, the observer provided information about inconsistencies or errors in the segmentations. Three general trends were identified (Supplementary Figure 3):

1. missing ventral segmentations;
2. missing segmentations near the costodiaphragmatic recess; and
3. incorrect segmentation of the lower mediastinum.

Further, in some cases, segmentations were incomplete in the lung periphery, leaving a small space unaccounted for close to the edge of the lung (Figure 6).

TABLE 4 Sørensen-Dice-Score (DSC) results for three sequences, showing mean (stdv).

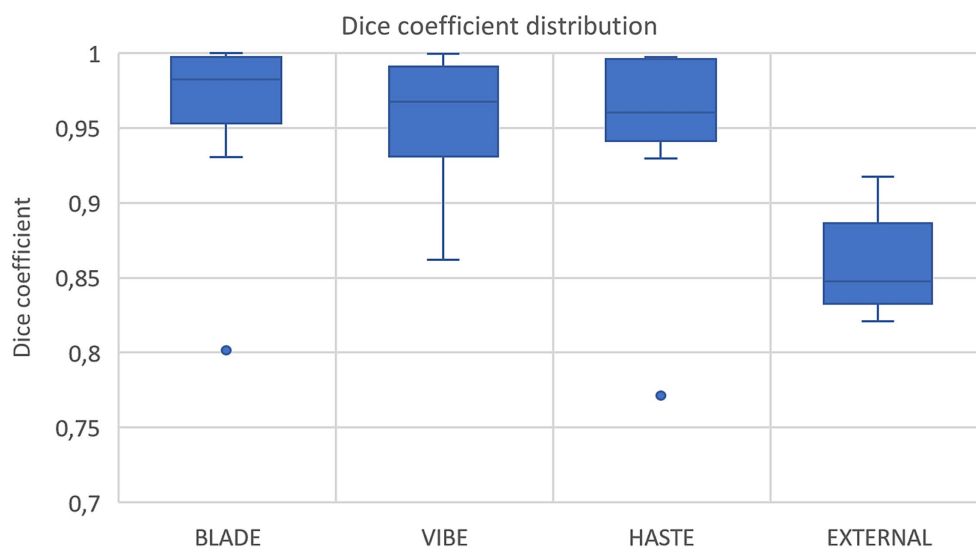| DSC	BLADE	VIBE	HASTE
Whole lung	0.96 ± 0.05	0.96 ± 0.04	0.95 ± 0.05
Left lung	0.95 ± 0.09	0.96 ± 0.04	0.96 ± 0.03
Right lung	0.97 ± 0.03	0.96 ± 0.04	0.93 ± 0.10

FIGURE 3
Box plots with dice coefficient of internal test sets from BLADE, VIBE, and HASTE and the external test set.

3.4 nnU-Net performance is independent of age and disease severity

Anatomy such as form and size of the chest change with age, and lung disease severity alters anatomy and signal of the lungs. Thus, we correlated DSC with the patient age and DSC with disease severity. Age as well as the MRI global score did not show an association with nnU-Net performance $r=0.09$ and $r=-0.12$, respectively (Supplementary Figures 4, 5).

3.5 nnU-Net shows acceptable performance on external validation data

Nine cases with corresponding HASTE MRI from two external centers, with ages ranging from 3 months to 2 years were segmented using the network trained on the HASTE imaging data. The average DSC across all validated cases was 0.85 (± 0.03) with a range from 0.82 to 0.92 for both lung halves combined (Figure 3, external data). Regarding lung halves, the left lung was segmented better, with a DSC ranging from 0.79 to 0.92, compared to the right lung with a DSC of 0.70 to 0.92. This indicates acceptable, but not perfect performance.

3.6 Slice based analysis highlights segmentation quality

A visual inspection was conducted on all data from the internal test set to ascertain the quantity of slices that would require manual correction. A total of 1,720 slices from the 45 internal test cases were subjected to quality control. The results are consistent with the responses provided in question 1 of the questionnaire. The CF case with the lowest score assigned by the radiologist (2/10) exhibited the highest number of slices requiring correction. Specifically, 79% of slices in the right lung and 29% of slices in the left lung were of

insufficient quality. Overall, the mean percentage of slices in the right lung and left lung that required correction was 10.60% (±16.46) and 8.75% (±9.39), respectively (Supplementary Figure 6).

4 Discussion

Segmentation can play a vital role as a pre-processing step before applying machine learning-based image analysis methods. In our work, lung half segmentation of pediatric MRIs of CF patients using three different sequences, BLADE, VIBE, and HASTE were created utilizing the nnU-Net neural network. A dataset comprising 165 cases, with 55 cases for each of the three sequences, was employed for the training, validation, and testing of the nnU-Net. For each sequence, the nnU-Net was trained individually using a training set of 40 cases and a testing set of 15 cases. For evaluation, the Sørensen-Dice-Score was used in combination with a tailored questionnaire and a slice-based analysis to provide a more detailed insight into the quality of the segmentations.

Overall, the segmentation performance achieved a mean Dice of 0.95 or higher for all sequences and lung halves except for the right lung on the HASTE sequence, which reached a mean Dice of 0.93. With the patient's age ranging from just a few months to 17 years, the segmentation performance was correlated with age. Generally, it was visible that the segmentation quality stayed constant across all pediatric age classes, further supported by Pearson correlation coefficient $r=0.09$. Due to the different disease status of the patients, the global MRI score was correlated with the Dice coefficient. Patients with both lower and higher global MRI score were segmented equally well, which is supported by the Pearson correlation coefficient of $r=-0.12$. This demonstrates the excellent performance of the nnU-Net for lung lobe segmentation in pediatric chest MRIs within our cohort. The high mean DSC indicates robust segmentation performance, independent of the underlying pathological changes induced by CF in the pediatric stage. An improvement in segmentation
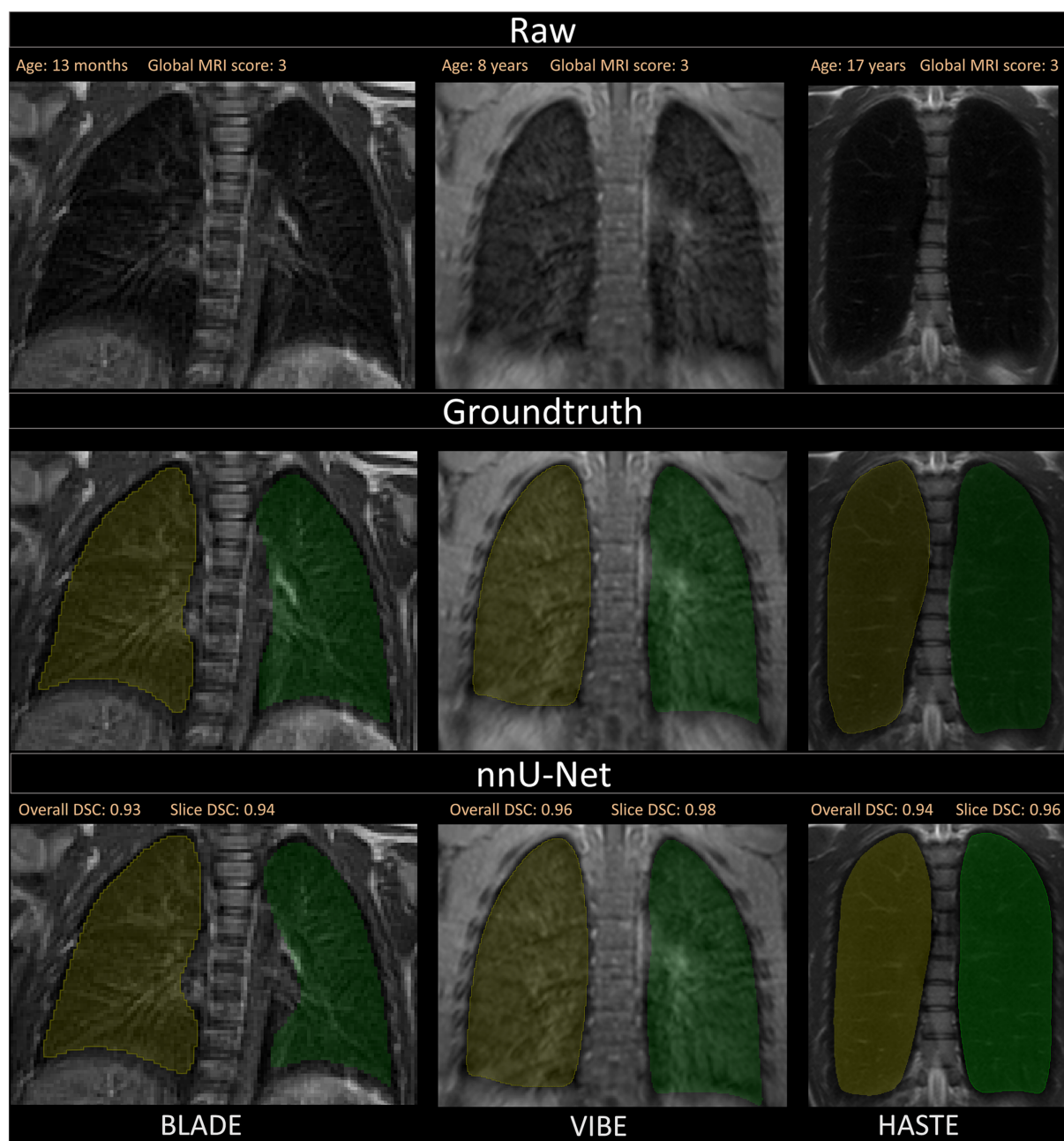
FIGURE 4
Visualization of the three different sequences with ground truth segmentations and good segmentations produced by the nnU-Net. Overall DSC corresponds to the dice coefficient of the entire lung and Slice DSC to the dice coefficient of the visualized slice. The segmentation of the right lung is indicated in yellow and the left lung segmentation in green. Each column corresponds to one MRI sequence. In the top row, the raw images are shown. The second row contains the manually annotated lung halves (ground truth). In the third row, the segmentation calculated by the corresponding nnU-Net is depicted. The three shown patients are of ascending age from left to right, thus the different lung sizes. All three patients have a global MRI score of 3. The results of the questionnaire indicated that the lungs were rated with a score of 10/10 for BLADE, 9/10 for VIBE, and 9/10 for HASTE. The different contrasts and gray levels are due to the different sequences. Both the ground truth and the segmentation of the lung halves appear to be very similar. Although the right and left lung differ in size and shape, segmentation performance seems to be almost equal. In general, the high Sørensen-Dice-Score and corresponding high segmentation performance are evident.

performance might be expected with an increased amount of training data (49). However, in segmentation tasks that require precise ground truth annotations, which are extremely time-intensive to generate, necessary trade-offs must be made.

For a qualitative analysis, we provided a questionnaire to the observers for the purpose of evaluating the segmentations manually in addition to the Sørensen-Dice-Score. Consistent segmentation errors in the ventral and dorsal areas of the lung, as well as around the costodiaphragmatic recess were detected. These errors can be caused by the thickness of the image slices, which directly affects the appearance of the tissue. When the slice thickness increases, tissue other than lung tissue becomes included, which may lead to the partial

FIGURE 5
Segmentation quality by sequence with the questionnaire evaluating the segmentations from 0 (bad) to 10 (good) of the internal test set.

volume effect (PVE) (50). PVE occurs in volumetric imaging, when more than one tissue type is present in a voxel. In cases where the lung parenchyma ends in the middle of the slice, the voxel will have a different shade of gray compared to voxels completely inside or outside the lung. Depending on the patient and amount of non-lung area on the entire slice, this gray level complicates manual and automated segmentation (Figure 7). Furthermore, even the slightest movement in the ventral and dorsal areas may introduce additional artifacts or blurring. While the observers annotated certain areas as lung tissue, the neural network failed to do so. To overcome this challenge, more annotated data could improve the dorsal and ventral segmentation performance. In addition to the questionnaire, the observer conducted a slice-based analysis annotating which slices required manual correction. The analysis revealed that the left lung necessitated more corrections than the right lung.

To achieve a more general evaluation of the trained neural networks, an external dataset from two other centers was segmented and evaluated with the Sørensen-Dice-Score. This was a useful test to explore whether generalization had been achieved, allowing the processing of data from a different source than the training data. Generally, overfitting on the training data is common, leading to very good performances on the training and test set from the same distribution but poor performance on external data. Despite the fact that the Sørensen-Dice-Score for the external dataset did not exceed 0.92, with a range of 0.7–0.92, it suggests that the trained neural network has overall generalizability, given the differences between the two test sets regarding age and number of slices. A comparison of the number of MRI slices in the internal and external test sets reveals notable differences: the internal dataset averages 38 slices per MRI, while the external dataset averages only 19 slices. Since the Dice coefficient is more sensitive to segmentation errors when the overall segmented area is smaller, this difference in slice count and corresponding segmentation area must be taken into account when interpreting the results (51).

Factors such as different MRI scanner specifications, protocols, slice thickness, and resolution affect image quality and therefore segmentation performance. The small number of external MRIs ($n = 9$) limited the general interpretability. Efforts to improve segmentation performance on the external dataset could include retraining the nnU-Net configuration with external MRI data to reduce segmentation errors. The overall segmentation performances on the three MRI sequences were comparable with existing work in literature. Lung CT scans have been segmented fully automated, reaching high accuracies for lung lobe segmentations with a mean Dice coefficient of up to 0.97 (52). In chest MRI, lung segmentation can be achieved using traditional approaches such as thresholding, but neural network-based segmentation approaches have recently been shown to outperform traditional methods (25). Astley et al. even showed the nnU-Net can be trained to perform well across several sequences, diseases, and vendors reaching a median Dice coefficient of 0.96 on the internal and 0.97 on an independent test set (30). Moreover, their results demonstrated that the 3D-Unet exhibited superior performance compared to the 2D version, which, in turn, outperformed the conventional segmentation approach, spatial fuzzy C-means. In contrast to their work, our study focused on pediatric MRIs of the entire pediatric range of patients with varying degrees of CF disease severity. Efforts toward improving MRI-based lung segmentation include artificially created images to increase robustness in case of severe pathologies (31). For hyperpolarized 129Xe MRI, segmentation performances with a Dice score of 0.929 and above were demonstrated using multiple different methods, highlighting the superiority of the nnU-Net over conventional segmentation methods (53, 54). Neonatal lung segmentations showed a Sørensen-Dice-Score of 0.908 and 0.880 on an independent test set with segmentations automatically by a combination of U-Nets (32).

The resulting segmentations for both the internal and external test sets of the underlying study exhibited variability, yet never attained a Dice coefficient of 1.0. This raises a pivotal question about the criteria for determining whether a segmentation is suitable for subsequent processing or downstream analyses. While a Dice coefficient of 1.0 represents perfect segmentation, striving to improve the coefficient from an already high mean value such as 0.95, may demand a
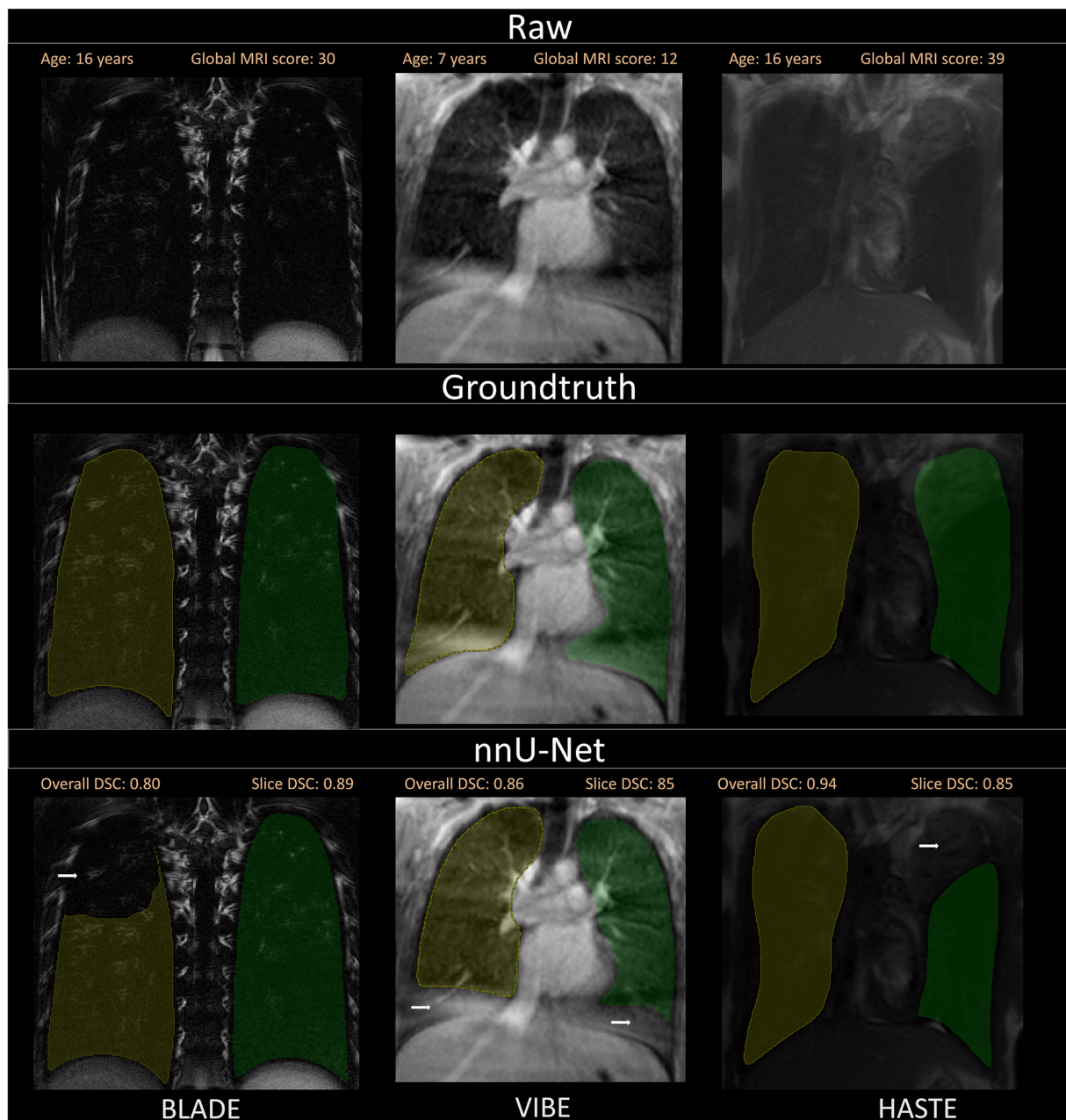
FIGURE 6
A selection of segmentations with a slightly lower Sørensen-Dice-Score, as well as visual discrepancies between ground truth and nnU-Net segmentation is shown. Overall DSC corresponds to the dice coefficient of the entire lung and Slice DSC to the dice coefficient of the visualized slice. Segmentation errors are indicated with white arrows in the second row. Three common segmentation mistakes are shown: Incomplete segmentations for BLADE, wrong segmentations due to breathing motion or other artifacts for VIBE and pathological changes influencing segmentation performance on the patient captured with the HASTE protocol. Based on the results of the questionnaire, the lungs were rated with a score of 2/10 for BLADE, 6/10 for VIBE, and 7/10 for HASTE.

disproportionate amount of time, effort and computational resources. In practice, the pursuit of marginal improvements—bringing the Dice score closer to 1.0—often results in diminishing returns. Such refinements may have minimal impact on the overall effectiveness or accuracy of downstream tasks, particularly when the current segmentation quality is already deemed suitable for clinical decision-making or research purposes. Therefore, it is critical to assess whether the additional time and computational effort invested in further optimizing segmentation is justified, or whether the existing

performance is sufficient for the intended applications. When observers segment lungs, they hardly ever reach complete agreement. Segmentation tasks are always dependent on the reader, their experience in the domain, and the tools used. Literature has shown that a Dice coefficient below 0.9 is not uncommon as reader agreement (55). Given that the overarching objective is to automate the Eichinger score, it can be argued that segmentation errors that do not significantly impact the majority of a lung half might be considered acceptable. However, in other research questions, this threshold may
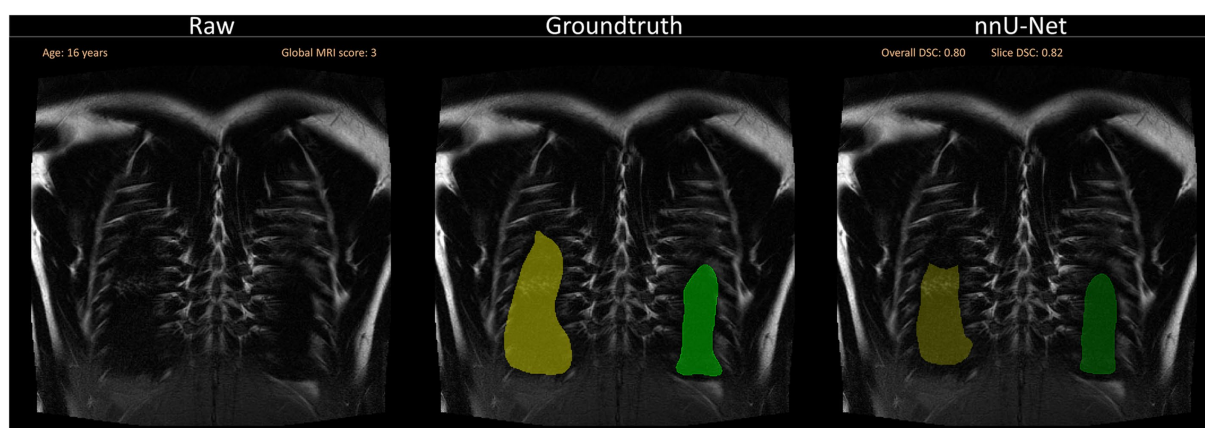
FIGURE 7
Segmentation errors dorsally due to partial volume effects and general difficulty to segment close to the ribs. The visualization shows a slice close to the back of the patient without any segmentations (left), with the annotated ground truth (middle) and the suggested segmentation by the neural network (right). Overall DSC corresponds to the dice coefficient of the entire lung and Slice DSC to the dice coefficient of the visualized slice.

have to be set differently, for example, in the context of tumor resection or radiotherapy planning, where segmentations require a higher degree of accuracy (56, 57).

In summary, the results obtained in this study are comparable to those reported in similar studies. To the best of our knowledge, we are the first to demonstrate successful pediatric lung half segmentations for patients with different stages of cystic fibrosis on the MRI sequences BLADE, VIBE, and HASTE.

Our study has some limitations that require discussion. The sample size of 55 cases with corresponding MRIs for each sequence, especially in the age group of patients under 1 year, is relatively small. Compared to the external dataset, all internal cases have larger lungs due to the higher mean age, which could influence the segmentation performance. Moreover, the majority of included patients had a global MRI score of 20 or less. Therefore, it is unclear whether our results are transferable to cohorts of older patients or patients with more advanced lung disease. Future work may focus on this aspect as well as an extension to other sequences.

Recent advancements in this field of research are driving the development of various methods for automated segmentation (19). In the future, it may be valuable to explore these approaches on this dataset and consider expanding the current model to include the remainder of the patients and cases. Especially, since the overall goal of automating the Eichinger score, works toward automated lung lobe segmentations should be explored. Pusterla et al. showed recently that automated lung segmentation with a combination of neural networks is possible with high accuracy (35). Earlier studies demonstrated that segmentation of perfusion maps with a 3D U-Net is an effective approach. However, the evaluation of lung lobes on MRI is challenging due to the difficulty in discerning lobe fissures, if they are visible at all. A lung atlas-based approach, which is independent of age and disease status, may prove advantageous, particularly in light of the findings reported by Tutison et al. regarding the segmentation of the lung (58). With the automated lung segmentation in place, further complex deep learning-based analysis techniques can be applied to assist radiologists in monitoring treatment response, therapy progression, and overall lung health of CF patients, potentially saving time. These results reinforce the efforts toward automated analysis of chest MRIs of patients with cystic fibrosis.

In conclusion, the performance of the nnU-Net in segmenting the lung halves of MRIs from pediatric CF patients demonstrated good agreement with manual segmentations. The segmentation performance of pediatric CF patients does not appear to be significantly influenced by age or disease status.

Data availability statement

The data analyzed in this study are subject to the following licenses/restrictions: the data that support the findings of this work are available from the corresponding author upon reasonable request. Requests to access these datasets should be directed to urs.eisenmann@med.uni-heidelberg.de.

Ethics statement

The studies involving humans were approved by clinicaltrials.gov identifiers NCT00760071 and NCT02270476. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

FR: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. LW: Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. NH: Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology,

Funding

Acknowledgments

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 reviewer OP declared a past co-authorship with the authors MW and MM to the handling editor.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1401473/full#supplementary-material

References

1. Gibson RL, Burns JL, Ramsey BW. Pathophysiology and Management of Pulmonary Infections in cystic fibrosis. *Am J Respir Crit Care Med.* (2003) 168:918–51. doi: 10.1164/rccm.200304-505SO

2. Welsh MJ, Ramsey BW, Accurso F, Cutting GR. Cystic fibrosis In: DL Valle, S Antonarakis, A Ballabio, AL Beaudet and GA Mitchell, editors. The Online Metabolic and Molecular Bases of Inherited Disease. New York, NY: McGraw-Hill Education (2019)

3. Bell SC, Mall MA, Gutierrez H, Macek M, Madge S, Davies JC, et al. The future of cystic fibrosis care: a global perspective. *Lancet Respir Med.* (2020) 8:65–124. doi: 10.1016/S2213-2600(19)30337-6

4. Grasemann H, Ratjen F. Early lung disease in cystic fibrosis. *Lancet Respir Med.* (2013) 1:148–57. doi: 10.1016/S2213-2600(13)70026-2

5. Stick S, Tiddens H, Aurora P, Gustafsson P, Ranganathan S, Robinson P, et al. Early intervention studies in infants and preschool children with cystic fibrosis: are we ready? *Eur Respir J.* (2013) 42:527–38. doi: 10.1183/09031936.00108212

6. Sly PD, Wainwright CE. Diagnosis and early life risk factors for bronchiectasis in cystic fibrosis: a review. *Expert Rev Respir Med.* (2016) 10:1003–10. doi: 10.1080/17476348.2016.1204915

7. Wielpütz MO, Eichinger M, Biederer J, Wege S, Stahl M, Sommerburg O, et al. Bildgebung der Lunge bei Mukoviszidose und klinische Interpretation. *RöFo.* (2016) 188:834–45. doi: 10.1055/s-0042-104936

8. Wielpütz MO, Puderbach M, Kopp-Schneider A, Stahl M, Fritzsching E, Sommerburg O, et al. Magnetic resonance imaging detects changes in structure and perfusion, and response to therapy in early cystic fibrosis lung disease. *Am J Respir Crit Care Med.* (2014) 189:956–65. doi: 10.1164/rccm.201309-1659OC

9. Stahl M, Wielpütz MO, Graeber SY, Joachim C, Sommerburg O, Kauczor H-U, et al. Comparison of lung clearance index and magnetic resonance imaging for assessment of lung disease in children with cystic fibrosis. *Am J Respir Crit Care Med.* (2017) 195:349–59. doi: 10.1164/rccm.201604-0893OC

10. Wielpütz MO, von Stackelberg O, Stahl M, Jobst BJ, Eichinger M, Puderbach MU, et al. Multicentre standardisation of chest MRI as radiation-free outcome measure of lung disease in young children with cystic fibrosis. *J Cyst Fibros.* (2018) 17:518–27. doi: 10.1016/j.jcf.2018.05.003

11. Wielpütz MO, Eichinger M, Wege S, Eberhardt R, Mall MA, Kauczor H-U, et al. Midterm reproducibility of chest magnetic resonance imaging in adults with clinically stable cystic fibrosis and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* (2019) 200:103–7. doi: 10.1164/rccm.201812-2356LE

12. Woods JC, Wild JM, Wielpütz MO, Clancy JP, Hatabu H, Kauczor H-U, et al. Current state of the art MRI for the longitudinal assessment of cystic fibrosis. *J Magn Reson Imaging.* (2020) 52:1306–20. doi: 10.1002/jmri.27030

13. Stahl M, Steinke E, Graeber SY, Joachim C, Seitz C, Kauczor H-U, et al. Magnetic resonance imaging detects progression of lung disease and impact of newborn screening

in preschool children with cystic fibrosis. *Am J Respir Crit Care Med*. (2021) 204:943–53. doi: 10.1164/rccm.202102-0278OC

14. Wucherpfennig L, Triphan SMF, Wege S, Kauczor H-U, Heussel CP, Sommerburg O, et al. Elexacaftor/Tezacaftor/Ivacaftor improves bronchial artery dilatation detected by magnetic resonance imaging in patients with cystic fibrosis. *Ann Am Thorac Soc*. (2023) 20:1595–604. doi: 10.1513/AnnalsATS.202302-168OC

15. Eichinger M, Optazaite D-E, Kopp-Schneider A, Hintze C, Biederer J, Niemann A, et al. Morphologic and functional scoring of cystic fibrosis lung disease using MRI. *Eur J Radiol*. (2012) 81:1321–9. doi: 10.1016/j.ejrad.2011.02.045

16. Puderbach M, Eichinger M, Haeselbarth J, Ley S, Kopp-Schneider A, Tuengerthal S, et al. Assessment of morphological MRI for pulmonary changes in cystic fibrosis (CF) patients. *Investig Radiol*. (2007) 42:715–24. doi: 10.1097/RLI.0b013e318074fd81

17. Asgari Taghanaki S, Abhishek K, Cohen JP, Cohen-Adad J, Hamarneh G. Deep semantic segmentation of natural and medical images: a review. *Artif Intell Rev*. (2021) 54:137–78. doi: 10.1007/s10462-020-09854-1

18. Wang R, Lei T, Cui R, Zhang B, Meng H, Nandi AK. Medical image segmentation using deep learning: a survey. *IET Image Process*. (2022) 16:1243–67. doi: 10.1049/ipr2.12419

19. Liu X, Song L, Liu S, Zhang Y. A review of deep-learning-based medical image segmentation methods. *Sustain For*. (2021) 13:1224. doi: 10.3390/su13031224

20. Kumar Y, Brar TPS, Kaur C, Singh C. A comprehensive study of deep learning methods for kidney tumor, cyst, and stone diagnostics using CT images. *Arch Computat Methods Eng*. (2024) 31:4163–4188. doi: 10.1007/s11831-024-10112-8

21. Qureshi I, Yan J, Abbas Q, Shaheed K, Riaz AB, Wahid A, et al. Medical image segmentation using deep semantic-based methods: a review of techniques, applications and emerging trends. *Inform Fusion*. (2023) 90:316–52. doi: 10.1016/j.inffus.2022.09.031

22. Isensee F, Petersen J, Klein A, Zimmerer D, Jaeger PF, Kohl S, et al. (2018). nn U-Net: Self-adapting framework for U-net-based medical image segmentation. Available online at: https://arxiv.org/pdf/1809.10486.pdf

23. Siddique N, Paheding S, Elkin CP, Devabhaktuni V. U-net and its variants for medical image segmentation: a review of theory and applications. *IEEE Access*. (2021) 9:82031–57. doi: 10.1109/ACCESS.2021.3086020

24. Thukral BB. Problems and preferences in pediatric imaging. *Indian J Radiol Imag*. (2015) 25:359–64. doi: 10.4103/0971-3026.169466

25. Kohlmann P, Strehlow J, Jobst B, Krass S, Kuhnigk J-M, Anjorin A, et al. Automatic lung segmentation method for MRI-based lung perfusion studies of patients with chronic obstructive pulmonary disease. *Int J Comput Assist Radiol Surg*. (2015) 10:403–17. doi: 10.1007/s11548-014-1090-0

26. Heimann T, Eichinger M, Bauman G, Bischoff A, Puderbach M, Meinzer H-P (2012). "Automated scoring of regional lung perfusion in children from contrast enhanced 3D MRI" in *Medical Imaging 2012: Computer-Aided Diagnosis*. SPIE. 83150U (SPIE Proceedings).

27. Zha W, Fain SB, Schiebler ML, Evans MD, Nagle SK, Liu F. Deep convolutional neural networks with multiplane consensus labeling for lung function quantification using UTE proton MRI. *J Magn Reson Imaging*. (2019) 50:1169–81. doi: 10.1002/jmri.26734

28. Weng AM, Heidenreich JF, Metz C, Veldhoen S, Bley TA, Wech T. Deep learning-based segmentation of the lung in MR-images acquired by a stack-of-spirals trajectory at ultra-short echo-times. *BMC Med Imaging*. (2021) 21:79. doi: 10.1186/s12880-021-00608-1

29. Willers C, Bauman G, Andermatt S, Santini F, Sandkühler R, Ramsey KA, et al. The impact of segmentation on whole-lung functional MRI quantification: repeatability and reproducibility from multiple human observers and an artificial neural network. *Magn Reson Med*. (2021) 85:1079–92. doi: 10.1002/mrm.28476

30. Astley JR, Biancardi AM, Hughes PJC, Marshall H, Collier GJ, Chan H-F, et al. Implementable deep learning for multi-sequence proton MRI lung segmentation: a multi-center, multi-vendor, and multi-disease study. *J Magn Reson Imaging*. (2023) 58:1030–44. doi: 10.1002/jmri.28643

31. Crisosto C, Voskrebenzev A, Gutberlet M, Klimeš F, Kaireit TF, Pöhler G, et al. Artificially-generated consolidations and balanced augmentation increase performance of U-net for lung parenchyma segmentation on MR images. *PLoS One*. (2023) 18:e0285378. doi: 10.1371/journal.pone.0285378

32. Mairhörmann B, Castelblanco A, Häfner F, Koliogiannis V, Haist L, Winter D, et al. Automated MRI lung segmentation and 3D morphologic features for quantification of neonatal lung disease. *Radiol Artif Intell*. (2023) 5:e220239. doi: 10.1148/ryai.220239

33. Missimer JH, Emert F, Lomax AJ, Weber DC. Automatic lung segmentation of magnetic resonance images: a new approach applied to healthy volunteers undergoing enhanced deep-inspiration-breath-hold for motion-mitigated 4D proton therapy of lung tumors. *Phys Imag Radiat Oncol*. (2024) 29:100531. doi: 10.1016/j.phro.2024.100531

34. Taran TV, Pavlova OS, Gulyaev MV, Dmitriev DS, Pistrak AG, Ryabikov KN, et al. Automated image registration and perfusion sorting algorithms for PREFUL MRI. *Appl Magn Reson*. (2024) 55:741–52. doi: 10.1007/s00723-024-01684-6

35. Pusterla O, Heule R, Santini F, Weikert T, Willers C, Andermatt S, et al. MRI lung lobe segmentation in pediatric cystic fibrosis patients using a recurrent neural network trained with publicly accessible CT datasets. *Magn Reson Med*. (2022) 88:391–405. doi: 10.1002/mrm.29184

36. Hirtz S, Gonska T, Seydewitz HH, Thomas J, Greiner P, Kuehr J, et al. CFTR cl-channel function in native human colon correlates with the genotype and phenotype in cystic fibrosis. *Gastroenterology*. (2004) 127:1085–95. doi: 10.1053/j.gastro.2004.07.006

37. Sommerburg O, Wielpütz MO, Trame J-P, Wuennemann F, Opdazaite E, Stahl M, et al. Magnetic resonance imaging detects chronic rhinosinusitis in infants and preschool children with cystic fibrosis. *Ann Am Thorac Soc*. (2020) 17:714–23. doi: 10.1513/AnnalsATS.201910-777OC

38. Wucherpfennig L, Wuennemann F, Eichinger M, Schmitt N, Seitz A, Baumann I, et al. Longitudinal magnetic resonance imaging detects onset and progression of chronic rhinosinusitis from infancy to school age in cystic fibrosis. *Ann Am Thorac Soc*. (2023) 20:687–97. doi: 10.1513/AnnalsATS.202209-763OC

39. Wucherpfennig L, Wuennemann F, Eichinger M, Seitz A, Baumann I, Stahl M, et al. Long-term effects of lumacaftor/ivacaftor on paranasal sinus abnormalities in children with cystic fibrosis detected with magnetic resonance imaging. *Front Pharmacol*. (2023) 14:1161891. doi: 10.3389/fphar.2023.1161891

40. Triphan SMF, Stahl M, Jobst BJ, Sommerburg O, Kauczor H-U, Schenk J-P, et al. Echo time-dependence of observed lung T1 in patients with cystic fibrosis and correlation with clinical metrics. *J Magn Reson Imaging*. (2020) 52:1645–54. doi: 10.1002/jmri.27271

41. Ciet P, Serra G, Bertolo S, Spronk S, Ros M, Fraioli F, et al. Assessment of CF lung disease using motion corrected PROPELLER MRI: a comparison with CT. *Eur Radiol*. (2016) 26:780–7. doi: 10.1007/s00330-015-3850-9

42. Dobritz M, Radkow T, Nittka M, Bautz W, Fellner FA. VIBE mit paralleler Akquisitionstechnik—eine neue Möglichkeit der dynamischen kontrastverstärkten MRT der Leber. *RöFo*. (2002) 174:738–41. doi: 10.1055/s-2002-32223

43. Semelka RC, Kelekis NL, Thomasson D, Brown MA, Laub GA. HASTE MR imaging: description of technique and preliminary results in the abdomen. *J Magn Reson Imaging*. (1996) 6:698–9. doi: 10.1002/jmri.1880060420

44. Stahl M, Wielpütz MO, Ricklefs I, Dopfer C, Barth S, Schlegtendal A, et al. Preventive inhalation of hypertonic saline in infants with cystic fibrosis (PRESIS). A randomized, double-blind, controlled study. *Am J Respir Crit Care Med*. (2019) 199:1238–48. doi: 10.1164/rccm.201807-1203OC

45. Robinson J. Likert scale In: AC Michalos, editor. Encyclopedia of Quality of Life and Well-Being Research. Dordrecht: Springer (2014). 3620–1.

46. Dice LR. Measures of the amount of ecologic association between species. *Ecology*. (1945) 26:297–302. doi: 10.2307/1932409

47. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. Sci Py 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. (2020) 17:261–72. doi: 10.1038/s41592-019-0686-2

48. Kirch W ed. Pearson's correlation coefficient. In: Encyclopedia of public health. Dordrecht: Springer (2008).

49. Yang J, Zhang Z, Gong Y, Ma S, Guo X, Yang Y, et al. (2022). Do deep neural networks always perform better when eating more data? Available online at: http://arxiv.org/pdf/2205.15187.pdf

50. González Ballester MA, Zisserman AP, Brady M. Estimation of the partial volume effect in MRI. *Med Image Anal*. (2002) 6:389–405. doi: 10.1016/S1361-8415(02)00061-0

51. Maier-Hein L, Reinke A, Godau P, Tizabi MD, Buettner F, Christodoulou E, et al. Metrics reloaded: recommendations for image analysis validation. *Nat Methods*. (2024) 21:195–212. doi: 10.1038/s41592-023-02151-z

52. Weinheimer O, Konietzke P, Heussel CP, Kauczor H-U, Robinson TE, Galban CJ, et al. (2019). "Improving pulmonary lobe segmentation on expiratory CTs by using aligned inspiratory CTs" in *Medical Imaging 2019: Computer-Aided Diagnosis*, San Diego, California, United States. Bellingham, Washington, USA: SPIE. p. 126 (Progress in biomedical optics and imaging; vol. 20, no. 48). February 17–20, 2019.

53. Leewiwatwong S, Lu J, Dummer I, Yarnall K, Mummy D, Wang Z, et al. Combining neural networks and image synthesis to enable automatic thoracic cavity segmentation of hyperpolarized 129Xe MRI without proton scans. *Magn Reson Imaging*. (2023) 103:145–55. doi: 10.1016/j.mri.2023.07.001

54. Astley JR, Biancardi AM, Hughes PJC, Marshall H, Smith LJ, Collier GJ, et al. Large-scale investigation of deep learning approaches for ventilated lung segmentation using multi-nuclear hyperpolarized gas MRI. *Sci Rep*. (2022) 12:10566. doi: 10.1038/s41598-022-14672-2

55. Jin L, Ma Z, Li H, Gao F, Gao P, Yang N, et al. Interobserver agreement in automatic segmentation annotation of prostate magnetic resonance imaging. *Bioengineering*. (2023) 10. doi: 10.3390/bioengineering10121340 Available at: https://www.mdpi.com/2306-5354/10/12/1340

56. Liu Z, Tong L, Chen L, Jiang Z, Zhou F, Zhang Q, et al. Deep learning based brain tumor segmentation: a survey. *Complex Intell Syst*. (2023) 9:1001–26. doi: 10.1007/s40747-022-00815-5

57. Gandotra S, Kumar Y, Modi N, Choi J, Shafi J, Ijaz MF. Comprehensive analysis of artificial intelligence techniques for gynaecological cancer: symptoms identification, prognosis and prediction. *Artif Intell Rev*. (2024) 57. doi: 10.1007/s10462-024-10872-6

58. Tustison NJ, Qing K, Wang C, Altes TA, Mugler JP. Atlas-based estimation of lung and lobar anatomy in proton MRI. *Magn Reson Med*. (2016) 76:315–20. doi: 10.1002/mrm.25824

Frontiers in Medicine

Translating medical research and innovation into improved patient care

A multidisciplinary journal which advances our medical knowledge. It supports the translation of scientific advances into new therapies and diagnostic tools that will improve patient care.

Discover the latest Research Topics

See more →

frontiers | Research Topics