



OPEN ACCESS

EDITED AND REVIEWED BY
Claudio Sette,
Catholic University of the Sacred Heart,
Rome, Italy

*CORRESPONDENCE

Mónica Hebe Vazquez-Levin
✉ mhvazl@gmail.com;
✉ mhvazquez@ibyme.conicet.gov.ar

SPECIALTY SECTION

This article was submitted to
Cancer Genetics,
a section of the journal
Frontiers in Oncology

RECEIVED 07 February 2023

ACCEPTED 20 February 2023

PUBLISHED 31 March 2023

CITATION

Vazquez-Levin MH, Reventos J and Zaki G
(2023) Editorial: Artificial intelligence:
A step forward in biomarker discovery
and integration towards improved
cancer diagnosis and treatment.
Front. Oncol. 13:1161118.
doi: 10.3389/fonc.2023.1161118

COPYRIGHT

© 2023 Vazquez-Levin, Reventos and Zaki.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Editorial: Artificial intelligence: A step forward in biomarker discovery and integration towards improved cancer diagnosis and treatment

Mónica Hebe Vazquez-Levin^{1*}, Jaume Reventos²
and George Zaki³

¹Instituto de Biología y Medicina Experimental (IBYME), Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina (CONICET) Fundación IBYME (FIBYME), Buenos Aires, Argentina,

²Institut d'Investigació Biomedica de Bellvitge (IDIBELL) and Universitat Internacional de Catalunya, Barcelona, Spain, ³Frederick National Laboratory for Cancer Research, National Cancer Institute at Frederick (NIH), Frederick, MD, United States

KEYWORDS

cancer, artificial intelligence, machine learning, digital twin, precision medicine

Editorial on the Research Topic

[Artificial intelligence: A step forward in biomarker discovery and integration towards improved cancer diagnosis and treatment](#)

In cancer, a biomarker refers to a substance or process indicative of the presence of cancer in the body. However, the idea of “one-molecule (or process) marker” indicated by its presence, and the existence of an undergoing transforming cancer process is currently a utopia. During the past decade, there has been a fundamental shift in cancer research and clinical decision-making, moving from qualitative data to quantitative digital data. A large wealth of cancer biomarkers and images has come from research laboratories and clinical institutions worldwide. Moreover, the major bulk of information has arisen from genomics, proteomics, metabolomics, and other omics, but also from oncology clinics, imaging, epidemiology and more. Artificial Intelligence (AI) is a unique technology that is able to combine all the above, and particularly suited to establish novel therapies and predictive models of drug response (1, 2). The combination of several biomarkers, by means of Machine Learning (ML) algorithms, would reach unprecedented conclusions in diagnosis, prediction and general decision making of novel anticancer therapies (3–5). In addition, the multimodal temporal data collected from patients with cancers can feed to initialize and track a Digital Twin to experiment with multiple possible treatments *in silico*.

This Research Topic has gathered 10 selected contributions in the area of ML tools, Deep Learning and Cancer Digital Twin technologies in the field of Precision Oncology, and contains one review, one minireview and eight original contributions.

The review paper by Asada *et al.* emphasizes the relevance of Precision Oncology and the integration of whole genome sequencing analysis, epigenome analyses and the use of ML, and opens a discussion about future perspectives in the field.

Networks of cellular systems and arrays of biological models described by ordinary and partial differential equations were developed in the last decades towards a better understanding of biological systems. [Mertins](#) minireview describes the use of ML algorithms to analyze computational dynamic ordinary differential equation models in combination with omics data, towards the discovery of novel biomarkers and novel molecular targets.

Tumor cell heterogeneity has been for many years a distortion factor in the interpretation of cellular and molecular findings in oncology. Tumor degree of purity is playing an important role in optimizing the correlation of the research findings with therapeutic anticancer strategies. By using Random Forest ML, [Yang et al.](#) were able to assess tumor purity in children CNS tumors, which will imply genomic, biological and clinical implications.

More than 60% of cervical cancers are caused by Human Papilloma Virus (HPV) 16 genotype, classified into lineages A, B, C, and D. In their contribution, [Asensio-Puig et al.](#) report the development of a Random Forest-based new model to assess HPV16 lineage. Authors highlight that their model is 40 times faster than current assessment done with Maximum Likelihood Tree, which requires a manual annotation and cannot assess poorly sequenced samples.

The work by [Chang et al.](#) proposes a novel ML predictive model utilizing a three-Dimensional Convolutional Neural Network (3D-CNN) to predict the presence of lymph node metastasis and the postoperative positive margin status based on preoperative CT scans. Their report provides a proof of concept for the preoperatively use of radiomics and 3D-CNN deep learning framework to improve the prediction of positive resection margins as well as the presence of lymph node metastatic disease.

In the report by [Spiller et al.](#), the utility and feasibility of imaging, computer vision and ML to determine patient-derived organoids vital status is reported. By acquiring bright field images at different time points without relying upon vital dyes, authors track the dynamic response of individual organoids to various drugs. In addition, authors report a web-based data visualization tool, called the Organoizer, available for public use.

[Abousamra et al.](#) present a Deep Learning workflow that generates Tumor Infiltrating Lymphocytes (TIL) maps to study their abundance and spatial distribution in 23 cancer types. Authors trained three state-of-the-art CNN architectures (namely VGG16, Inception-V4, ResNet-34) with training data from The Cancer Genome Atlas, combining manual annotations from pathologists and computer-generated labels from a first-generation TIL model. It also incorporates automated thresholding to convert model predictions into binary classifications to generate TIL maps.

With the aim to identify putative biomarkers for lung cancer and to elucidate the pathogenesis of this disease, [Bahado-Singh et al.](#) combined AI and DNA methylation analysis of circulating cell-free tumor DNA. The study analyzes six AI platforms, including Support Vector ML and Deep Learning, to measure cytosine (CpG) methylation changes across the genome in lung cancer. Training sets and validation sets are generated and 10-fold cross validation performed. To elucidate lung cancer pathogenesis, gene enrichment analysis using g:profiler and GREAT enrichment is done.

Triple-negative breast cancer (TNBC) always requires neoadjuvant chemotherapy (NACT) for a pathological complete response and improved long-term survival. [Irajizad et al.](#) previously identified a polyamine biomarker suitable to assess which patient will respond to NACT. In their contribution, [Irajizad et al.](#) identified TNBC patients who will be insensitive to NACT, by using ML methods.

Finally, the contribution by [Batch et al.](#) aimed to improve the detection of metastatic disease over time from structured radiology reports with the ultimate goal of building and updating a Digital Twin to model long-term prognosis. By exposing prediction models to historical information using Natural Language Processing (NLP), the authors were able to extract and encode relevant features from medical text reports, and use these features to develop, train, and validate models. Over 700 thousand radiology reports were used for model development to predict the presence of metastatic disease. The model uses features from consecutive structured patient text radiology reports. Three models were developed to classify the type of metastatic disease: a simple CNN, a CNN augmented with an attention layer, and Recurrent Neural Network labels. To develop the models, a subset of the reports was curated for ground-truth. Results from the three models were compared (accuracy, precision, recall, and F1-score) to a single-report model previously developed to analyze one report instead of multiple past reports. Results suggest that NLP models can extract cancer progression patterns from multiple consecutive reports and predict the presence of metastatic disease in multiple organs with higher performance when compared with a single-report-based prediction.

In summary, contributions to this special edition highlight how AI will accelerate the advancement of Personalized Medicine and cancer care, by improving patient diagnosis, treatment, and prognosis.

Author contributions

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

Conflict of interest

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

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. Kehl KL, Xu W, Gusev A, Bakouny Z, Choueiri TK, Riaz IB, et al. Artificial intelligence-aided clinical annotation of a large multi-cancer genomic dataset. *Nat Commun* (2021) 12(1):7304. doi: 10.1038/s41467-021-27358-6
2. You Y, Lai X, Pan Y, Zheng H, Vera J, Liu S, et al. Artificial intelligence in cancer target identification and drug discovery. *Sig Transduct Target Ther* (2022) 7(1):156. doi: 10.1038/s41392-022-00994-0
3. Koh DM, Papanikolaou N, Bick U, Illing R, Kahn CE Jr, Kalpathi-Cramer J, et al. Artificial intelligence and machine learning in cancer imaging. *Commun Med (Lond)* (2022) 2:133. doi: 10.1038/s43856-022-00199-0
4. Kong J, Ha D, Lee J, Kim I, Park M, Im SH, et al. Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nat Commun* (2022) 13(1):3703. doi: 10.1038/s41467-022-31535-6
5. Nguyen L, Van Hoeck A, Cuppen E. Machine learning-based tissue of origin classification for cancer of unknown primary diagnostics using genome-wide mutation features. *Nat Commun* (2022) 13(1):4013. doi: 10.1038/s41467-022-31666-w