

EDITORIAL OPEN



Artificial intelligence for the next generation of precision oncology

npj Precision Oncology (2021)5:79; <https://doi.org/10.1038/s41698-021-00216-w>

The concept of precision oncology involves the prescription of therapies that target the molecular driver alterations of an individual patient's tumor. This treatment paradigm has been accelerated in recent years through increasing integration of molecular profiling approaches into mainstream clinical oncology, and the approval of a variety of molecularly targeted agents that have improved the clinical outcomes of patients across multiple cancer types. Targeted therapies such as trastuzumab or vemurafenib have become the standards of care for patients with HER2-expressing breast cancers or BRAF-mutant melanomas, respectively, and immune checkpoint inhibitors recently received tissue-agnostic approval for treating patients whose tumors exhibit microsatellite instability. With the advent of personalized cancer medicine, the portfolio of anti-tumor agents and companion diagnostic assays has experienced an unprecedented expansion, and a plethora of clinical trials are exploring the most effective treatments for the best-matching patients. In this context, data-driven approaches for optimizing clinical matching, reducing the cost of diagnostic testing, and improving the prediction of clinical phenotypes hold great promise for enhancing the clinical management of patients and maximize the value of precision oncology. However, all stakeholders must be aware of the important challenges that need to be overcome to reach these goals. This editorial aims at raising awareness about these challenges.

Artificial intelligence (AI), and more concretely its machine learning (ML) branch, can process large-scale and heterogeneous data sets to discern medically relevant patterns. ML is hence showing promise for making a positive impact in patient care. Pioneering works in the field of computer vision, and specifically in digital pathology, have demonstrated the enormous potential of ML models for improving the accuracy of diagnostic protocols using automated methods that require little human input. For instance, deep learning models can learn to predict molecular features of tumor samples, such as driver mutations or histological subtyping, based on features of histopathological images^{1,2}. Approaches such as this could be scaled-up and incorporated in clinical pathways to streamline the time-consuming task of histological examination, and the generalization across tumor types and potential for accelerating clinical diagnoses through assisting expert pathologists have been supported by recent studies^{3–5}. The field of diagnostic radiology is another area where the integration of ML techniques might have an impact on early cancer detection and diagnosis. Models trained on sufficiently large and well-annotated data sets of mammography or chest radiographs have demonstrated potential to predict the risk of developing breast⁶ or lung⁷ cancers by extracting information that is often imperceptible to the expert human eye, and to identify prognostic imaging biomarkers that can be associated with long-term patient outcome⁸. In non-invasive molecular testing, random forest algorithms applied on circulating microRNAs⁹ can accurately diagnose glioblastoma preoperatively, and other

types of ML analyses applied to DNA methylation signatures measured in plasma cell-free nucleic acids achieve robust performance for tumor detection and classification¹⁰. To date, however, these models have been tested in retrospective studies, and their real-world impact for accelerating clinical workflows and increasing the rates of early cancer detection remains to be achieved prospectively.

On another clinical front, AI-powered decision support systems are increasingly getting better at optimizing treatment decision making for cancer patients. In addition to their role in drug discovery efforts, ML models integrating data on tumor growth kinetics, molecular profiling, and pharmacological properties could provide accurate prospective recommendations on the most effective therapeutic approaches for individual cancer patients—be it through drug repurposing, by identifying synergistic drug combinations, or in optimizing dosing schedules to maximize therapeutic efficacy. For this to be achieved, population-scale data sets containing well-annotated clinical and molecular information need to be made available^{11,12}. Issues such as tumor heterogeneity and sampling bias, tumor evolution dynamics, or incomplete information on electronic medical records, among many other factors, represent major roadblocks for using and integrating these sources of data to build ML models, and are only partially solved by approaches increasing training set size (number of patients) at the expense of relevance (e.g., including patients with other cancer subtypes or treated with related drugs) or robustness of the data (e.g., merging clinical trials with different dosage schedules or even exploiting less-controlled data such as electronic health records). Even in the uncommon situations where large, carefully curated, and relevant data sets with low levels of noise are available, several fundamental methodological challenges currently limit ML approaches, including reliable inference of testable causality¹³, identifying the most suitable algorithm and features to solve a given problem¹⁴, learning from high-dimensional data sets¹⁵, or accurately estimating model generalization¹⁶. A related medical challenge is found when predicting patient response or resistance to anti-cancer treatments, not only for cytotoxic chemotherapies that form the backbone of most treatment regimens, but also for targeted therapies that either elicit unacceptable toxicities or achieve modest benefit even in the presence of the associated genomic marker. Here too, ML models are particularly promising for improving these predictions by identifying an optimal combination of patient features, which may include a range of non-genetic tumor features, if provided with appropriate input training data¹⁷. To overcome the limitation imposed by the paucity of large, well-annotated patient data sets, some studies have attempted to predict clinical phenotypes using ML models trained on response data from preclinical experimental systems of patient-derived xenografts¹⁸ or large-scale in vitro drug response studies¹⁹. However, despite the high value of preclinical models for drug development^{20,21}, their general usefulness for precision oncology remains to be proven²². The problem of predicting optimal treatment approaches and response patterns remains unsolved, despite substantial efforts, as these models have not so far influenced clinical care in a

significant way. Notwithstanding, this nascent field may hold great potential for advancing precision oncology.

The success of AI systems across medical domains will also depend on establishing a roadmap that delineates the step-wise integration of digital tools in the clinic²³. There have been several initiatives proposing “best practice” guidelines to ensure that AI methods are developed and deployed in a way that maximizes the benefit for patients. On the model development front, these include recommendations for providing sufficient methodological detail on the development of algorithms, and encouraging sharing of data sets and code to enhance the transparency in the reporting of AI algorithms in medicine^{11,12,24,25} and allow other researchers to determine the rigor, quality, reproducibility, and generalizability of the findings^{26–28}. Studies also need to adopt standardized guidelines for describing and reporting aspects related to the purpose and context of the “clinical need” that is being addressed, the quality of data used to train the models—including issues related to power calculation, labeling, model biases, etc.—measures of performance, outputs and framework for integration in clinical pathways and workflows, among others. Many of these issues have been approached by recent checklists intending to promote a standard and transparent reporting of clinical interventions involving ML approaches^{29–31}. In addition, successful deployment of AI systems in the clinic will largely depend on the trust that end-users—clinicians and patients alike—have in the recommendations provided. To achieve this, it will be necessary to educate users on the working principles and degree of interpretability of AI systems, and to carefully define and test the interfaces that enable human–computer collaboration.

The application of AI to precision oncology is still in its infancy. In recent years we have witnessed a proliferation of proof-of-concept studies that offer a glimpse of what the next generation of precision oncology could look like. We have outlined here several challenges that need to be overcome for AI to make strides in medicine. We have also argued that the prospective application of these AI systems needs to follow a responsible path in collaboration with all stakeholders. Expectations are justifiably high given these initial studies, but true progress can only come from a deeper understanding of the discussed limitations. We look forward to seeing how AI may enhance precision oncology approaches and improve patient care around the world in the years to come.

Received: 23 April 2021; Accepted: 21 July 2021;

Published online: 18 August 2021

Pedro J. Ballester^{1,2,3,4,6} and Javier Carmona^{5,6}

¹Cancer Research Center of Marseille, INSERM U1068, Marseille, France. ²Institut Paoli-Calmettes, Marseille, France. ³Aix-Marseille Université, Marseille, France. ⁴CNRS UMR7258, Marseille, France.

⁵Nature Medicine, New York, NY, USA. ⁶These authors contributed equally: Pedro J. Ballester, Javier Carmona. ✉email: pedro.ballester@inserm.fr; javier.carmona@us.nature.com

REFERENCES

- Coudray, N. et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat. Med.* **24**, 1559–1567 (2018).
- Campanella, G. et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat. Med.* **25**, 1301–1309 (2019).
- Chen, M. et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. *npj Precis. Oncol.* **4**, 14 (2020).
- Hollon, T. C. et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat. Med.* **26**, 52–58 (2020).
- Tschanl, P. et al. Human–computer collaboration for skin cancer recognition. *Nat. Med.* **26**, 1229–1234 (2020).
- Yala, A., Lehman, C., Schuster, T., Portnoi, T. & Barzilay, R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* **292**, 60–66 (2019).
- Ardila, D. et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **25**, 954–961 (2019).
- Lu, M. T. et al. Deep learning to assess long-term mortality from chest radiographs. *JAMA Netw. Open* **2**, e197416 (2019).
- Ebrahimkhani, S. et al. Deep sequencing of circulating exosomal microRNA allows non-invasive glioblastoma diagnosis. *npj Precis. Oncol.* **2**, 1–9 (2018).
- Shen, S. Y. et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* **563**, 579–583 (2018).
- Clinical Cancer Genome Task Team of the Global Alliance for Genomics and Health, Lawler, M. et al. Sharing clinical and genomic data on cancer—the need for global solutions. *N. Engl. J. Med.* **376**, 2006–2009 (2017).
- Rutella, S. et al. Society for Immunotherapy of Cancer clinical and biomarkers data sharing resource document: volume i – conceptual challenges. *J. Immunother. Cancer* **8**, 1389 (2020).
- Azuaje, F. Artificial intelligence for precision oncology: beyond patient stratification. *npj Precis. Oncol.* **3**, 1–5 (2019).
- Chen, D. et al. Deep learning and alternative learning strategies for retrospective real-world clinical data. *npj Digit. Med.* **2**, 43 (2019).
- Bomane, A., Gonçalves, A. & Ballester, P. Paclitaxel response can be predicted with interpretable multi-variate classifiers exploiting DNA-methylation and miRNA data. *Front. Genet.* **10**, 1041 (2019).
- Tsamardinos, I., Greasidou, E. & Borboudakis, G. Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation. *Mach. Learn.* **107**, 1895–1922 (2018).
- Adam, G. et al. Machine learning approaches to drug response prediction: challenges and recent progress. *npj Precis. Oncol.* **4**, 1–10 (2020).
- Li, J., Ye, C. & Mansmann, U. R. Comparing patient-derived xenograft and computational response prediction for targeted therapy in patients of early-stage large cell lung cancer. *Clin. Cancer Res.* **22**, 2167–2176 (2016).
- Geeleher, P. et al. Discovering novel pharmacogenomic biomarkers by imputing drug response in cancer patients from large genomics studies. *Genome Res.* **27**, 1743–1751 (2017).
- Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* **16**, 531–543 (2017).
- Sausville, E. A. & Burger, A. M. Contributions of human tumor xenografts to anticancer drug development. *Cancer Res.* **66**, 3351–3354 (2006).
- Willyard, C. The mice with human tumours: growing pains for a popular cancer model. *Nature* **560**, 156–157 (2018).
- Wiens, J. et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat. Med.* **25**, 1337–1340 (2019).
- Hulsen, T. et al. From big data to precision medicine. *Front. Med.* **6**, 34 (2019).
- Mbuagbaw, L., Foster, G., Cheng, J. & Thabane, L. Challenges to complete and useful data sharing. *Trials* **18**, 71 (2017).
- Kakarmath, S. et al. Best practices for authors of healthcare-related artificial intelligence manuscripts. *npj Digit. Med.* **3**, 1–3 (2020).
- Mongan, J., Moy, L. & Kahn, C. E. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. *Radiol. Artif. Intell.* **2**, e200029 (2020).
- Haibe-Kains, B. et al. Transparency and reproducibility in artificial intelligence. *Nature* **586**, E14–E16 (2020).
- Liu, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat. Med.* **26**, 1364–1374 (2020).
- Rivera, S. C. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat. Med.* **26**, 1351–1363 (2020).
- Norgeot, B. et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat. Med.* **26**, 1320–1324 (2020).

COMPETING INTERESTS

P. J. B. is an Associate Editor for *npj Precision Oncology*. J. C. is an Advisory Editor for *npj Precision Oncology*.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021