

EDITORIAL

## Balancing between A and I

Emma Skovgaard Pedersen<sup>a</sup> , Christoffer Johansen<sup>b</sup>  and Mette Kielsholm Thomsen<sup>a,b</sup> 

<sup>a</sup>Department of Clinical Epidemiology, Aarhus University & Aarhus University Hospital, Aarhus, Denmark; <sup>b</sup>Danish Cancer Society National Cancer Survivorship and Late Effects Research Center (CASTLE), Department of Oncology, Copenhagen University Hospital - Rigshospitalet, Copenhagen Denmark

Artificial intelligence (AI) in cancer is fast evolving, and the next step on the evolutionary ladder is incorporating AI methods for risk stratification in precision oncology, but we need to call for balance between methodological rigor and clinical relevance. Currently, studies focus on either one, but rarely both aspects.

Ageing populations, increasing cancer incidence, and lack of health care personnel and resources [1] mean that technological assistance from AI applications are becoming ever more important, as potential tools to aid efficacy of diagnostic and treatment pathways, as well as prioritisation of services. This potential has been met by an increase in publications on AI in oncology; during the past two decades, the publication activity has moved from less than 10 papers per year up to more than a thousand papers annually, illustrating the interest in exploring the utility of these techniques in all aspects of cancer care (Figure 1).

AI areas of success currently span across cancer diagnosis, pre-treatment and treatment. AI tools have proven useful in aiding visual inspection of images in diagnosis and treatment of cancer patients, for example, during colonoscopies and for interpretation of mammogram images [2–4]. A review of clinical trials found that AI-tools increased adenoma detection during colonoscopies, with the increase driven by an improved detection of small adenomas, rather than advanced adenomas, underscoring that this tool can act as a supplement to colonoscopist expertise rather than a substitution [5]. In pre-treatment, an ongoing trial is investigating AI-assisted plans for perioperative care to optimise patients before surgery [6]. In treatment, AI tools may already be able to ease the workload of clinicians; an AI algorithm for contouring of scans prior to proton treatment of squamous cell carcinoma of the pharynx or larynx was shown to be more consistent within patients and were able to adhere more strictly to national guidelines than contours made by oncologists [7]. To plan and monitor treatment, AI methods are accelerating developments within circulating tumour DNA (ctDNA) for assessment of therapeutic response, by enhancing signal enrichment, making ctDNA monitoring applicable for personalised treatment strategies, and across more cancers than was previously possible [8].

### ARTICLE HISTORY

Received 7 March 2025

Accepted 24 April 2025

Published 7 May 2025

### KEYWORDS

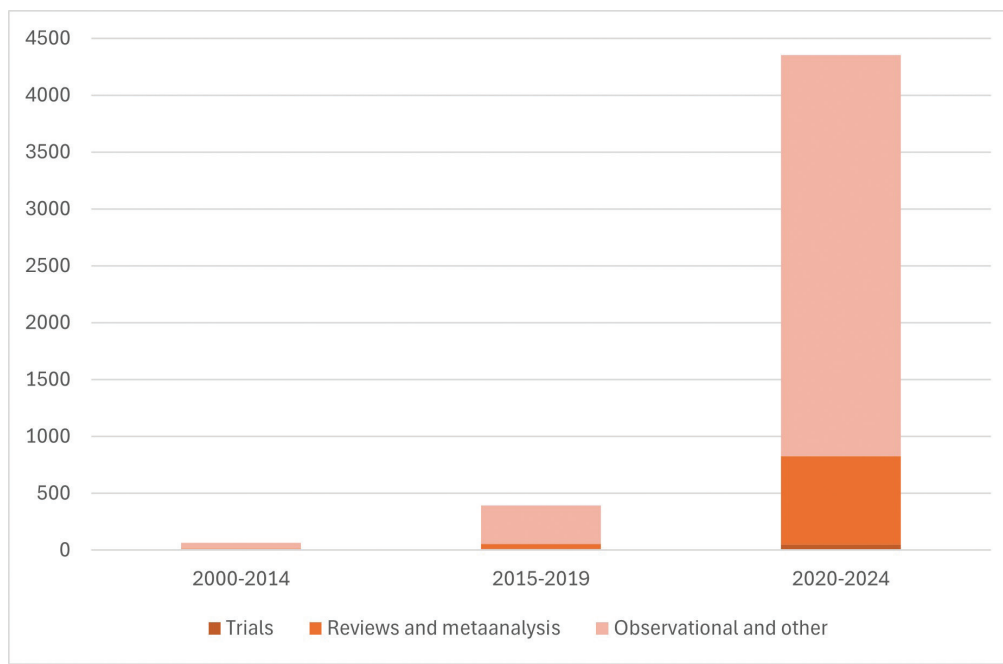
Artificial intelligence; machine learning; precision medicine; diagnostic imaging; cancer therapy

While successful implementation of the above-mentioned AI tools into clinical practice constitutes a research area in itself [4, 9, 10], another major AI research area in cancer is developing high-performing prediction models using machine learning models, a subtype of AI [11]. Machine learning models have the potential to aid risk-stratification for screening, diagnosis, treatment planning, and prognostic prediction [12]. So far, this is mainly an academic pursuit, with a small minority of published prediction models achieving regulatory approval and subsequently being implemented into clinical practice [13]. Thus, the largest unharvested potential of AI in cancer, in our opinion, lies within prediction models for personalised oncology.

To enhance the use of AI prediction models in cancer research, we have identified two overarching sources of imbalance, that are currently tipping studies into the no-man's-land between data science research and clinical medicine. These two challenges pertain to *expertise* and *alignment*.

### Expertise – in cancer and data science

Studies utilising machine learning for prediction models in cancer tend to suffer from one of two problems: First, studies performed by data scientists will often have a skilfully built model and include thorough documentation of the employed machine learning algorithms, but the models may not be clinically applicable or relevant, due to a lack of medical expertise, and results on actual model features and risk stratification utility are not sufficiently reported. A lack of epidemiological insight or understanding of the clinical reality may lead to issues such as conditioning on the future – for example, Zhang et al., included



**Figure 1.** From 65 PubMed hits until 2014 to more than 4000 in 2020–2024: Reports with AI/ML and cancer in the title, by publication period and type of report.

information on any second or third tumour to predict survival time from diagnoses of first tumour to death [14], resulting in inflated model performance. Another example is not acknowledging that in health care data, registrations are there for a reason, which means that healthy individuals are underrepresented in the development of AI tools [15, 16]. Medical records are a rich data source, but we should not blindly plunge into the vast amounts of data, without considering if the target population is represented by the data material.

Second, prediction studies led by clinicians may have a strong clinical foundation and relevance, but lack of technical and data expertise means that the resulting models may not be developed sufficiently to achieve adequate performance, and that reporting of methods is not sufficiently transparent [17]. These studies are at the same time especially prone to issues originating from lack of understanding of data structures and data management [18].

### Alignment – of purpose, results, and interpretation

A general theme in studies applying machine learning for prediction models in cancer is a lack of alignment between the purpose, results reported, and interpretation. When machine learning models are used for prediction, results tend to focus on metrics for model performance and comparison of various types of machine learning algorithms. Performance metrics are of course vital in judging which model is better [18], however, model performance by itself has limited clinical value. More transparency is needed in terms of which patient, tumour and treatment characteristics that lead to the prediction of low or high risk of the given outcome, or preferably the reporting of predictors included and a readily applicable model equation if possible and meaningful. Model output should be reported explicitly and in detail in the main text of papers reporting on

machine learning models, rather than being relegated to supplementary materials [19]. We appreciate that explicit listing of features may wrongfully imply causal inference to some readers; thus, clear and precise wording of model specifications is required.

Fitting a prediction model is only one step on the road to utilising AI in clinical settings. However, it is not rare to see a study developing a prediction model and concluding that said model should promptly be implemented in practice. It can be helpful, as others have pointed out, to liken the process of AI development for health care with that of drug development [20]. In other words, the development of the specific prediction model (i.e. the phase 1 study) is followed by processes of evaluation, deployment, and monitoring of daily use (i.e. phase 2, 3 and 4), before the AI tool can be integrated into clinical settings. As such, all four phases are necessary to judge the clinical value, and we cannot go from phase 1 to concluding on the clinical utility of AI models. For this, pragmatic validation and pilot tests are necessary steps [21].

### The road ahead for AI in cancer research

To advance AI and machine learning in prediction studies within cancer, a road that is unlikely to be straightforward lies ahead. Researchers, healthcare engineers, and tech-savvy clinicians alike must largely pave the way as they go, in the absence of up-to-date, and interdisciplinary scientific guidelines. Technologies are evolving quicker than the usual pace of systematic overviews and collaborative efforts. The result of this situation is an abundance of publications and unimplemented models built by well-meaning researchers in a race against obsolescence, wishing for a data-driven crystal ball to foretell the destiny of cancer patients [22]. No matter our optimism for technological advancements, it is crucial that we work together

across disciplines to ensure the best possible balance between validity of models and interpretability and applicability of results.

### Disclosure statements

The authors report there are no competing interests to declare.

### Data availability statement

N/A.

### Ethics declarations & trial registry information

N/A.

### Author contributions

MKT and ESP searched the existing literature and drafted the manuscript. All authors reviewed, edited and approved the final manuscript.

### References

- [1] Global cancer burden growing, amidst mounting need for services [Internet]. [cited 2025 Mar 6]. Available from: <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services>
- [2] Singh G, Kamalja A, Patil R, Karwa A, Tripathi A, Chavan P. A comprehensive assessment of artificial intelligence applications for cancer diagnosis. *Artif Intell Rev*. 2024;57(7):179. <https://doi.org/10.1007/s10462-024-10783-6>
- [3] Eisemann N, Bunk S, Mukama T, Baltus H, Elsner SA, Gomille T, et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. *Nat Med*. 2025;31(2):1–8. <https://doi.org/10.1038/s41591-024-03408-6>
- [4] Ng AY, Oberije CJG, Ambrózay É, Szabó E, Serfőző O, Karpati E, et al. Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. *Nat Med*. 2023;29(12):3044–9. <https://doi.org/10.1038/s41591-023-02625-9>
- [5] Hassan C, Spadaccini M, Mori Y, Foroutan F, Facciorusso A, Gkolfakis P, et al. Real-time computer-aided detection of colorectal neoplasia during colonoscopy : a systematic review and meta-analysis. *Ann Intern Med*. 2023;176(9):1209–20. <https://doi.org/10.7326/M22-3678>
- [6] Study Details. The AIDPRO-CRC Trial. *ClinicalTrials.gov* [Internet]. [cited 2025 Mar 6]. Available from: <https://clinicaltrials.gov/study/NCT06645015>
- [7] Nielsen CP, Lorenzen EL, Jensen K, Sarup N, Brink C, Smulders B, et al. Consistency in contouring of organs at risk by artificial intelligence vs oncologists in head and neck cancer patients. *Acta Oncol (Madr)*. 2023;62(11):1418–25. <https://doi.org/10.1080/0284186X.2023.2256958>
- [8] Widman AJ, Shah M, Frydendahl A, Halmos D, Khamnei CC, Øgaard N, et al. Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment. *Nat Med*. 2024;30(6):1655–66. <https://doi.org/10.1038/s41591-024-03040-4>
- [9] Swart R, Boersma L, Fijten R, Van Elmpt W, Cremers P, Jacobs MJG. Implementation strategy for artificial intelligence in radiotherapy: can implementation science help? *JCO Clin Cancer Inform*. 2024;8:e2400101. <https://doi.org/10.1200/CCI.24.00101>
- [10] Friedewald SM, Sieniek M, Jansen S, Mahvar F, Kohlberger T, Schacht D, et al. Triaging mammography with artificial intelligence: an implementation study. *Breast Cancer Res Treat*. 2025;211:1–10. <https://doi.org/10.1007/s10549-025-07616-7>
- [11] Farina E, Nabhen JJ, Dacoregio MI, Batalini F, Moraes FY. An overview of artificial intelligence in oncology. *Future Sci OA*. 2022;8(4):FSO787. <https://doi.org/10.2144/foa-2021-0074>
- [12] Sebastian AM, Peter D. Artificial intelligence in cancer research: trends, challenges and future directions. *Life*. 2022;12(12):1991. <https://doi.org/10.3390/life12121991>
- [13] Markowitz F. All models are wrong and yours are useless: making clinical prediction models impactful for patients. *NPJ Precis Oncol*. 2024;8(1):1–3. <https://doi.org/10.1038/s41698-024-00553-6>
- [14] Zhang IY, Hart GR, Qin B, Deng J. Long-term survival and second malignant tumor prediction in pediatric, adolescent, and young adult cancer survivors using Random Survival Forests: a SEER analysis. *Sci Rep*. 2023;13(1):1–12. <https://doi.org/10.1038/s41598-023-29167-x>
- [15] Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. A review of challenges and opportunities in machine learning for health. *AMIA Summits on Translational Science Proceedings*. American Medical Informatics Association, Washington DC, USA. 2020. p. 191.
- [16] Haue AD, Hjaltelin JX, Holm PC, Placido D, Brunak S. Artificial intelligence-aided data mining of medical records for cancer detection and screening. *Lancet Oncol*. 2024;25(12):e694–703. [https://doi.org/10.1016/S1470-2045\(24\)00277-8](https://doi.org/10.1016/S1470-2045(24)00277-8)
- [17] Ellis RJ, Sander RM, Limon A. Twelve key challenges in medical machine learning and solutions. *Intell Based Med*. 2022;6:100068. <https://doi.org/10.1016/j.ibmed.2022.100068>
- [18] Perez-Lopez R, Ghaffari Laleh N, Mahmood F, Kather JN. A guide to artificial intelligence for cancer researchers. *Nat Rev Cancer*. 2024;24(6):427–41. <https://doi.org/10.1038/s41568-024-00694-7>
- [19] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Calster B Van, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. <https://doi.org/10.1136/bmj-2023-078378>
- [20] Patel MR, Balu S, Pencina MJ. Translating AI for the clinician. *JAMA*. 2024;332(20):1701–2. <https://doi.org/10.1001/jama.2024.21772>
- [21] Cabitza F, Campagner A, Balsano C. Bridging the ‘last mile’ gap between AI implementation and operation: ‘data awareness’ that matters. *Ann Transl Med*. 2020;8(7):501. <https://doi.org/10.21037/atm.2020.03.63>
- [22] Seneviratne MG, Shah NH, Chu L. Bridging the implementation gap of machine learning in healthcare. *BMJ Innov*. 2020;6(2):45–7. <https://doi.org/10.1136/bmjinnov-2019-000359>