Artificial intelligence across oncology specialties: current applications and emerging tools

John Kang, Kyle Lafata, Ellen Kim, Christopher Yao, Frank Lin, Tim Rattay, Harsha Nori, Evangelia Katsoulakis, Christoph Ilsuk Lee

ABSTRACT

Oncology is becoming increasingly personalised through advancements in precision in diagnostics and therapeutics, with more and more data available on both ends to create individualised plans. The depth and breadth of data are outpacing our natural ability to interpret it. Artificial intelligence (AI) provides a solution to ingest and digest this data deluge to improve detection, prediction and skill development. In this review, we provide multidisciplinary perspectives on oncology applications touched by AI—imaging, pathology, patient triage, radiotherapy, genomics-driven therapy and surgery—and integration with existing tools—natural language processing, digital twins and clinical informatics.

INTRODUCTION

In recent years, modern medical machine learning (ML) and artificial intelligence (AI) have become actualised in the clinic, with hospitals and clinics around the globe using AI assistance for a wide range of clinical applications such as diabetic retinopathy screening, stroke detection and predicting hospital readmissions.

AI promises to be a gamechanger in oncology as well, yet oncology is no stranger to moonshot promises. Soon after the initiation of the human genome project around the turn of the century, several promising laboratory-developed tests that purported to diagnose or personalise therapy in ovarian and lung cancer were developed and subsequently debunked. Despite these well-documented early setbacks, continued refinement of multigene assays has resulted in their integration into the standard of care in certain contexts, notably in breast and prostate cancer.

Many of the hard lessons learnt from entrusting decision-making to multigene laboratory tests remain, with both familiar and new questions regarding generalisability, imbalance and missing data and real-world applications. Oncology is a multidisciplinary practice with imperfect information flow between subspecialties. Our primary aim in this narrative review is to address this limitation by discussing ongoing and emerging oncology AI efforts from the perspective of practitioner researchers in clinical informatics (CI), computer science, medical oncology, medical physics, pathology, radiation oncology, radiology and surgery. For more details on AI and ML algorithms, we refer readers to prior reviews. For a more in-depth discussion about data sharing, ethics and validation, we refer readers to prior work.

CURRENT APPLICATIONS OF AI IN ONCOLOGY

Radiological imaging

Cancer imaging, especially for screening and early detection, presents the archetype for integrating deep learning and AI algorithms into clinical practice. In theory, these algorithms should be able to identify patterns within medical images imperceptible to the human eye and be able to help identify areas that may represent malignant findings. Examples of ongoing efforts for early cancer detection on medical imaging using AI include mammography screening, CT lung cancer screening and early detection on prostate MRI.

Early breast cancer detection

The most studied area for AI and cancer imaging is the use of AI algorithms for earlier breast cancer detection on both two-dimensional and three-dimensional mammography. The DREAM Digital Mammography Challenge, the large-crowdsourced effort in deep learning algorithm development for medical imaging interpretation, launched the development of multiple competing commercial AI algorithms. Thus far, early cohort studies suggest that AI improves overall accuracy when used as an adjunct tool by radiologists interpreting...
mammograms. A few studies suggest that AI as a stand-alone tool may be equivalent to human interpretative accuracy for breast cancer screening. Of note, there is a paucity of data demonstrating that there is improved accuracy and outcomes from commercially available, Food and Drug Administration (FDA)-approved AI algorithms for mammography in large, diverse patient populations. Initial interim results from a prospective trial in Sweden comparing AI-supported screening versus standard double reading suggests that AI-supported screening leads to similar cancer detection rates. Currently, many commercially available algorithms remain unreimbursed given the lack of prospective, clinical effectiveness data.

Tumour characterisation

AI algorithms are being developed to aid in the automated characterisation of intratumoral heterogeneity, potentially allowing for more precise monitoring of disease progression and treatment efficacy. Traditional tumour segmentation correlates quantitative imaging features with biological data, including genetic data and molecular signatures. AI can remove the manually laborious work of supervised tumour segmentation, providing the benefits of unsupervised biologic characterisation of tumours. With deep learning and radiomic evaluation of tumour morphology and heterogeneity, more precise monitoring of treatment response of solid tumours can be determined through serial imaging (eg, CT scan, MRI, positron emission tomography (PET)/CT scan) that surpasses current oversimplified criteria such as the Response Evaluation Criteria in Solid Tumors that relies solely on size changes. Such efforts in AI-driven advanced imaging treatment response evaluation hold promise for more precise and personalised treatment decision-making for multiple cancer types, including cancer of the lung, liver, pancreas, and head and neck, as well as for metastatic cancer.

Pathology

The field of oncologic pathology has been transformed with the introduction of whole slide image (WSI) scanners. By digitising entire glass slides, this major technological advancement has enabled the application of AI and other advanced computational methods to whole tissue specimens in a manner not possible via traditional light microscopy.

Feature-based extraction

The emerging field of pathomics is a leading example of WSI applications. Pathomics is a high-throughput approach to digital tissue phenotyping, where computational methods are employed to transcribe otherwise unstructured imaging data (ie, pixel-level WSI information) into structured imaging features and actionable knowledge. Analogous to radiomics—where quantitative features are calculated from radiology images—pathomics is based on the extraction and analysis of quantitative features derived from digitised tissue samples.

The pathomic process typically starts with deep learning algorithms that detect and segment different tissue compartments (eg, tumour, stroma, etc) and/or cell types (eg, lymphocytes, cancer cells, etc), from which quantitative features are extracted. These features collectively describe unique patterns, which can serve as digital fingerprints of the tumour microenvironment and may capture imaging characteristics that are either difficult or impossible to characterise by humans. For example, morphological features can describe the size, shape and orientation of individual nuclei; topological features can provide measures of the spatial tissue architecture and interaction between different cell types; and texture features can quantify spatially encoded pixel intensity patterns of chromatin. Illustrating examples in oncology range from pathomic-based prognostic signatures for gastric cancer, bladder cancer, renal cell carcinoma, oropharyngeal cancer and breast cancer to predictors of microsatellite instability in colorectal cancer and mutational status of BRAF-mutated melanomas.

Deep learning representations

The advantage of hand-crafted feature engineering is that features are easily interpretable. However, they are limited by human intuition. In contrast, deep learning provides an alternative form of pathomic feature extraction, where features are learnt from imaging filters in an unsupervised fashion. While deep pathomic features are difficult to interpret, they can go beyond human intuition. Although deep learning is often thought of as a complete classifier, it is primarily a kind of image representation. As stated above, deep learning is a particularly powerful tool for pixel-level object detection and segmentation (which is often the first step of a pathomics problem). Illustrating examples in oncology include segmentation of epithelial tissue in prostate cancer, cellular detection and classification in colon cancer and bone marrow, counting tumour infiltrating lymphocytes and immune cell composition and segmentation of nuclei in cervical tissue for squamous epithelium cervical intraepithelial neoplasia grading.

Patient triage

The dual tasks of prognostication and triage can drive goals of care discussion or referrals to palliative care and hospice, yet remain challenging for oncologists, who tend to overestimate patients’ survival. Several groups have performed prospective trials based on data-driven triage frameworks which we discuss here (table 1).

Predicting unplanned hospital visits

Hong et al at Duke University aimed to triage patients undergoing radiation therapy by predicting acute care visits—emergency department or hospital admissions—to improve outcomes and decrease costs. After developing a gradient tree boosting (GTB) model using tabular electronic health record (EHR) data, they performed a randomised controlled trial for patients at high risk
of acute care visits to show that two times per week visits during radiation (compared with standard one time per week clinic visits) decrease such visits from 22% to 12%.53

Predicting survival
Parikh et al at the University of Pennsylvania aimed to predict which patients with metastatic cancer were at high risk of 180-day mortality to trigger timely serious illness conversations (SIC). After fixing the event rate (the proportion of patients flagged as high risk) to 2% to avoid alert fatigue, they developed a model with approximately 50% positive predictive value (precision).54 After running a silent trial (ie, without alerting clinicians) of their GTB model confirming their prior results (table 2),55 the authors followed up with an interventional randomised stepped-wedge cluster trial.56 The clusters received either weekly emails of high-risk patients with day of text message versus usual care (versus standard one time per week clinic visits) decrease such visits from 22% to 12%.53

Table 1 Examples of prospective AI-driven clinical trials in oncology

<table>
<thead>
<tr>
<th>Trial name/dates</th>
<th>Trial design</th>
<th>Patient population</th>
<th>Clinical objective</th>
<th>Clinical action or arms</th>
<th>Model objective</th>
<th>Model methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHIELD-RT53 (2019)</td>
<td>Randomised, controlled, quality improvement</td>
<td>Patients undergoing radiation</td>
<td>Decrease acute care visits</td>
<td>Two times per week versus standard weekly clinical visit</td>
<td>Predict &gt;10% risk of acute care visit</td>
<td>Gradient boosted trees</td>
</tr>
<tr>
<td>University of Pennsylvania, Manz et al55,56 (2019)</td>
<td>Prospective silent→phase III cluster wedge</td>
<td>New encounters in outpatient clinics</td>
<td>Increase serious illness conversations</td>
<td>Weekly email of high-risk patients with day of text message versus usual care</td>
<td>Predict 6-month mortality</td>
<td>Gradient boosted trees</td>
</tr>
<tr>
<td>Stanford, Gensheimer et al60 (2020–2021)</td>
<td>Quality improvement, controlled</td>
<td>Patients in four clinics: two control, two experimental</td>
<td>Increase advance care planning</td>
<td>Weekly emails with patients predicted to survive&lt;2 years</td>
<td>Predict time-dependent survival</td>
<td>Cox proportional hazards with tabular and text data</td>
</tr>
<tr>
<td>MASA18 (2021–2022)</td>
<td>Randomised, controlled, non-inferiority screening accuracy study</td>
<td>Women eligible for routine mammography screening</td>
<td>Improve screening outcomes, including lower interval cancer rate</td>
<td>AI-supported screening versus standard double reading without AI</td>
<td>Identify cancers on mammogram images</td>
<td>Proprietary deep convolutional neural network</td>
</tr>
</tbody>
</table>

AI, artificial intelligence.

Table 2 Confusion matrix summarising performance of the prospective silent trial of the Penn gradient boosted tree model to predict patients at high risk of death within 180 days

<table>
<thead>
<tr>
<th>Total patients (24582)</th>
<th>Predict survival&gt;180 days (23963)</th>
<th>Predict survival≤180 days (619*)</th>
<th>Event rate 2.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survived ≤180 days (1022)</td>
<td>742 false negatives</td>
<td>280 true positives</td>
<td>Sensitivity 27.4%</td>
</tr>
<tr>
<td>Survived &gt;180 days (23560)</td>
<td>23221 true negatives</td>
<td>339 false positives</td>
<td>Specificity 98.6%</td>
</tr>
<tr>
<td>NPV 96.9%</td>
<td>PPV 45.2%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Anderson et al reported total patients, the number of patients who survived ≤180 days and who survived >180 days, NPV, PPV and sensitivity. The other metrics were derived with possible rounding errors.

*The ‘predict survival≤180 days’ events were fixed at 619/24582=2.5% event rate.
NPV, negative predictive value; PPV, positive predictive value.
increased to 35% compared with 3% in a control cohort of clinics. See the ‘Large language models’ section for further developments in NLP.

Radiation oncology

Organ at risk segmentation

In radiation oncology, AI-assisted autosegmentation of tumour and normal organ contours has made major advancements from the early era of intensity analysis and shape modelling to the modern techniques using deep learning. An early randomised trial by Walker et al at MD Anderson demonstrated that resident physicians using autosegmentation assistance had time savings compared with manual segmentation. Autosegmentation is also critical to facilitate online adaptive radiotherapy where real-time treatment modification occurs to account for daily changes in setup, anatomy or density. Proprietary online adaptive radiotherapy platforms incorporate AI-driven autosegmentation, deformable registration and treatment planning with ongoing trials in several disease sites, including a randomised trial in lung cancer (table 3).

Treatment de-escalation

In head and neck cancer, radiation with or without concurrent systemic therapy is a standard of care for curative intent treatment, though not without significant toxicity due to sensitive adjacent organs. Standard radiation fields may include neck regions that are not grossly involved due to concern they may be harbouring subclinical disease. Chen et al at UT Southwestern sought to use computer vision to pinpoint which nodes were involved and only target these to decrease the toxicity of elective nodal radiation. They developed and prospectively validated a hybrid radiomics and convolutional neural network (CNN) model to predict neck node involvement in INRT-AIR de-escalation trial that performed involved node radiation (INRT) with promising early results showing no solitary elective nodal failure. Building on INRT-AIR, Sher et al are using INRT with or without near marginless daily adaptation in the DARTBOARD trial (table 3). The presence of extranodal extension (ENE) in neck nodes portends more aggressive cancer, though it is very challenging to detect on imaging and validated a deep CNN model to predict pathologic ENE on CT imaging and used ECOG-ACRIN E3311 data in a quality improvement study to suggest better ability to predict pathologic ENE compared with head and neck radiologists.

Genomics-driven precision oncology

Several categories of precision oncology using clinicogenomic data are emerging, ranging from improving prognostication to biomarker selection to drug development. Although still mainly in the research stage, there are several areas where AI may improve outcomes through refining treatment selection.

Prognostic and predictive tools through clinicogenomic data modelling

Complex clinicogenomic data can be synthesised to improve assessments of natural history and predict therapy benefits. For example, a predictive model that integrates molecular data with traditional clinicopathologic features can provide more nuanced stratification of leukaemic transformation than the Revised International Prognostic Scoring System. Clinicoepigenetic signatures can also predict benefit from adjuvant chemotherapy in patients with gastric cancer, and individual liver metastases with HER2-amplified metastatic colorectal cancer treated with dual HER2-targeted therapy with potential for detecting non-responding lesions.

---

**Table 3 Examples of prospective AI-driven clinical trials for radiotherapy planning**

<table>
<thead>
<tr>
<th>Trial name/dates</th>
<th>Trial design</th>
<th>Patient population</th>
<th>Clinical objective</th>
<th>Clinical action or arms</th>
<th>Model objective</th>
<th>Model methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>INRT-AIR (2019–2022)</td>
<td>Phase II</td>
<td>Patients with oropharynx, larynx squamous cell carcinoma</td>
<td>Risk of solitary elective nodal recurrence</td>
<td>Involved nodal radiation to suspicious nodes</td>
<td>Predict individual node positivity risk</td>
<td>Hybrid radiomics and three-dimensional CNN</td>
</tr>
<tr>
<td>DARTBOARD (2022–)</td>
<td>Phase II randomised</td>
<td>Same as INRT-AIR</td>
<td>Decrease xerostomia</td>
<td>INRT±dailyART (near marginless)</td>
<td>CBCT autosegmentation via proprietary CNN architecture</td>
<td>Same as INRT-AIR±AI-driven ART deformable registration and planning</td>
</tr>
<tr>
<td>ARTIA-Lung (2022–)</td>
<td>Randomised</td>
<td>Patients with stages IIIA and IIIB NSCLC</td>
<td>Decrease composite acute grade 3+toxicity</td>
<td>ART versus standard IMRT radiotherapy</td>
<td>Same as DARTBOARD</td>
<td>AI-driven ART deformable registration and planning</td>
</tr>
</tbody>
</table>

AI, artificial intelligence; AIR, AI-based Radiomics; ART, adaptive radiotherapy; CBCT, cone-beam CT; CNN, convolutional neural network; IMRT, intensity modulated radiotherapy; INRT, involved nodal radiotherapy; NSCLC, non-small cell lung cancer.
One emerging trend is to generate complex signatures predicting therapy utility by merging multimodal data. Multimodal models can predict response to immune checkpoint inhibitors (ICI) more precisely in patients with lung cancer, through combining mutational burdens, radiomic features and PD-L1 immunohistochemistry assay. In patients with castration-resistant prostate cancer, genomics and transcriptomics data have been combined using ML to predict a response to androgen receptor antagonists.

Biomarker selection with non-destructive virtual assays
AI methods can derive surrogate biomarkers for targeting the molecular alterations of clinical significance in tumour tissues. This is helpful when universal testing is not possible or too costly, or in rare variants where a specific biomarker may not be available. Using non-destructive ‘virtual assays’ also minimises serial confirmatory assays, thus preserving biopsy tissues and resulting in better turnaround time. Several studies have demonstrated the potential of using digital pathology and radiomics in such applications. As an example, screening of mismatch repair status on H&E-stained pathology slides and other imaging methods without immunohistochemistry has shown promising results in practice. This has significant practical implications in identifying ICI response in a small number of patients with a microsatellite unstable phenotype without universal testing. Applying radiogenomic assessment on CT images can identify non-small cell lung cancer (NSCLC) harbouring EGFR mutations based on texture features from pretreatment CT and PET/CT images, suggesting the potential to screen for precise actionable molecular alterations without needing invasive biopsy. Radiogenomic analysis has demonstrated a link between CT imaging and radiation-induced changes in cell-free DNA obtained via liquid biopsy of locally advanced lung cancers.

Structured insights for drug development through systematic repurposing
AI can assist in drug development by performing high-dimensional database analysis to screen for candidate drugs and regimens for trial development. In the secondary analysis of the SHIVA01 trial cohort, AI-assisted prioritisation of targeted agents was achieved through tumour molecular profiling. Deep learning on a pharmacogenomics database such as Genomic of Drug Sensitivity in Cancer can predict drug response and patients’ survival by examining gene expression pathways to associate disease control with molecular profiles. In a retrospective analysis, the NetBio framework has shown better prediction of treatment response to ICI in select cancers compared with traditional biomarkers.

Surgery
In surgery, exploring AI’s ability to guide surgical decision-making has been of great interest. Surgical decision-making can lean heavily on visual cues and images in addition to a surgeon’s understanding of the patient’s clinical picture. For surgical applications, AI could take a virtual form or provide direct physical support (including smart operating rooms, nanorobots or patient-assistance systems). Regardless of the form, three fields have emerged as the main application of ML in surgery: surgical skills assessment, supporting intraoperative decision-making and surgical outcomes prediction.

Assessment of surgical skills
AI has been applied to laparoscopic videos and trained to assist in anatomy identification, segmenting live surgical images as high-risk areas versus low risk areas to minimise adverse events from misinterpretation of anatomy. AI algorithms can recognise the operative steps of a laparoscopic sleeve gastrectomy obtained from surgical videos with 85.6% accuracy. When additional data was captured involving instrument handling, surgeon eye tracking and motion perturbations, AI has also been used to predict surgeon skill level, procedure length and potentially patient outcomes. In fact, experienced surgeons can be differentiated from beginners within the first 10 s, a task with 90% accuracy. This has very practical relevance, as currently, assessment of surgical quality largely relies on surgical case volumes, without granular objective data from intraoperative data.

Intraoperative decision-making
The application of AI in assisting intraoperative decision-making has been of great interest. In surgical oncology, the assessment of margins is critical to achieving optimal oncological outcomes. At the same time, the aim of modern cancer surgery is to spare healthy tissue as much as possible to minimise side effects and improve long-term quality of life. With the application of ML and various enhanced imaging systems including Raman spectroscopy, there is the potential for real-time intraoperative support while in the operating room. AI has also been used to determine optimal margins for cancer resection, such as in hepatic metastasectomy. These techniques are ripe for further evaluation as part of clinical trials.

Outcomes prediction
In predicting outcomes in surgical oncology, radiomics has been applied to predict response after neoadjuvant therapy and to select patients for surgery, for example, according to quantitative features associated with micro-invasion, or to distinguish high-risk versus low-risk precancerous lesions. ML has also been used to predict major complications after cancer surgery and to identify patients who are safe to be discharged 2 days following major gastrointestinal cancer surgery.

These represent the main themes of AI application within surgery. With ongoing technological advances, surgery will move towards a smart operating room and more precise perioperative decision-making.
EMERGING TOOLS
Large language models (LLM)
NLP has been applied to detect real-world cancer outcomes such as metastatic progression from radiology reports, pathology reports and clinical notes.106-109 As discussed above, Gensheimer et al have incorporated NLP into clinical triage.108 NLP systems for clinical trial matching are in test deployment at large healthcare networks.102

Since the late 2010s, advanced NLP models called LLM have leveraged transformer-based architectures with massive datasets (on the order of terabytes of text) to make significant breakthroughs in language understanding. While not specifically trained on biomedical or clinical text, LLMs such as Generative Pre-trained Transformer 4 (GPT-4) are capable of encoding clinical knowledge, as evidenced by their remarkable performance on medical licensing examinations and challenge problems,103-105 as well as to highly specialised topics such as medical physics.106 Researchers are now experimenting with fine-tuning pretrained models to focus specifically on medical applications, as seen in Med-PaLM 2 for question-answering, RadOnc-GPT for oncology-specific tasks and LLMs for clinical trial matching.106 107 108

In some cases, models are even trained from scratch using solely medical datasets to optimise their performance in the healthcare context. A prime example of domain-specific NLP utilisation is the Clinical Bidirectional Encoder Representations from Transformers (ClinicalBERT) model, which was trained on a large corpus of deidentified intensive care unit notes and fine-tuned to predict short term readmission.109

The technological leaps provided by LLM have led to high-profile healthcare collaborations. Recent collaborations include between Epic Systems and Microsoft to integrate GPT-4 by drafting of patient communication responses and providing data visualisation recommendations and between Mayo Clinic and Google Cloud, which builds on a partnership formed in 2019.

Research in this area has grown rapidly with several research groups training their own BERT-type and GPT-type models on their own datasets for various applications, and we recommend keeping a close eye on this space.

3.1.2. Challenges of NLP in healthcare
Despite showing strong performance on benchmark tasks, significant risks still remain for the use of LLM in healthcare and oncology. One prominent concern is the presence of bias within these models, which may inadvertently perpetuate or exacerbate existing healthcare disparities. Additionally, LLM are known to experience ‘hallucinations,’ generating plausible sounding yet incorrect or unrelated information, which could potentially lead to detrimental clinical decisions. Furthermore, these models may still exhibit inaccurate yet plausible reasoning, thereby making it difficult to catch errors and omissions. It is crucial to address these challenges and ensure these technologies are used responsibly, with human oversight remaining integral to decision-making processes.

Digital twins (DTs)
A DT is a virtual replica of a physical system which is not only created to mirror the real-world system but is also capable of analysis and prediction. DTs continuously monitor patients in real time, integrating data from wearable devices, sensors and electronic health records and thus is complemented by other technologies, including transfer learning, the Internet of Things, edge computing and cloud computing.110 DTs are being explored in oncology as a promising approach to enhance cancer care and may be used in various aspects of oncology including drug discovery and personalised treatment planning.111

Drug discovery
DTs have demonstrated the potential to streamline pharmaceutical processes and generate realistic input-output predictions for biochemical reactions. Through in silico techniques, several drugs have been identified and successfully brought to market for various diseases, including anticancer agent raltitrexed.112 In silico trials are currently being investigated, initially focusing on synthetic control arms and eventually expanding to predict clinical intervention. Both the US FDA and the European Medicines Agency have taken steps to support the integration of in silico approaches into control arms. For instance, a synthetic control arm consisting of 68 patients was used to extend the coverage of targeted therapy for NSCLC specifically alectinib, across 20 European countries.113 Synthetic controls have also played a role in expanding the indication of palbociclib, a kinase inhibitor, to include men with HR-positive HER2-negative advanced or metastatic breast cancer, as well as facilitating accelerated approval for blinatumomab which treats acute lymphoblastic leukaemia.114 In a phase I trial, existing quantitative systems pharmacology model of the anti-CD20/CD3 T-cell engaging bispecific antibody, mosnetuzumab, were used to incorporate different dosing regimens and patient heterogeneity within the trial.115

Treatment planning and prognosis monitoring
DTs may be created to personalise treatment planning as they enable the simulation and optimisation of treatments by integrating patient-specific genomics, imaging and clinical information. The utilisation of NLP for large-scale labelling of CT reports presents an opportunity to advance the development of DTs in oncology. In a recent study, NLP was used to perform consecutive multi-report prediction of metastases, enabling highly detailed representations that effectively model a cancer patient’s disease progression over time.116 These approaches facilitate the generation of a comprehensive database consisting of patterns of disease spread, facilitating early detection and prediction of an individual patient’s progression. In another study, DTs of patients were generated, and clinical
Clinical informatics

There are several resources that can help non-AI expert clinicians in the interpretation and ethical application of AI tools in the clinic, and the applications within specific disease sites, both during and after clinical training.

Education

In the USA, the clinical subspecialty of CI was recognised by the American Board of Medical Specialties (ABMS) in 2011 and the first physicians were board certified in CI in 2013. CI subspecialty fellowships are open to physicians from any ABMS specialty and allow fellows to spend 2 years dedicated to studying and practicing CI. NIH National Library of Medicine informatics fellowships can also provide physicians with opportunities to gain experience in programming and application of AI tools in clinical practice. There are other pathways and less formal educational resources, including master’s degrees or certificate programmes, American Medical Informatics Association 10×10 programs and massive open online courses on ML such as those on Coursera by Andrew Ng.

Research infrastructure and community

Sharing of data and conference resources in oncology are increasing, even outside of data access statements in publications. Federated learning can facilitate AI from much larger datasets while protecting data privacy by decentralising raw data, which has the potential to speed up validation of models. Work continues to try to standardise oncology data elements—mCODE, a collaboration between ASCO, CancerLinQ and MITRE—and interoperability—FHIR by HL7. The National Cancer Institute (NCI) is a major organiser of cancer datasets like the Cancer Research Data Commons, which includes The Cancer Genome Atlas, and the NCI Data Catalog and NCI Cancer Imaging Archive. Academic groups are building free software packages and platforms such as MultiAssayExperiment and CURATE.AI. AI-driven data fusion techniques that intelligently combine data from these different source domains (eg, clinical, imaging, omics, etc) can integrate knowledge to provide insight that is greater than the sum of the parts. More than a dozen technology companies are building platforms and software as a service (SaaS) tools to try to facilitate precision oncology and data analysis, including ConcertAI, Onc.Ai, Azra AI, ArteraAI and PreciseDx. It is imperative that oncologists become comfortable with critiquing, interpreting and applying these tools in clinical practice as well as research.

CONCLUSION

Much like oncologists 10–15 years ago would have been hard pressed to predict the paradigm shifts provided by advances in targeted therapy, immunotherapy, stereotactic radiation therapy, minimally invasive surgery, digital pathology and theranostics, it seems we are at or nearing an inflection point for AI in medicine given the amount of further research, validation and clinical trials were simulated to anticipate the optimal salvage therapy following progressive disease while on pembrolizumab. For spine metastases, DT was used to simulate vertebroplasty and its impact on mechanical stability of the vertebra. Finally, so called ‘virtual imaging trials’ aim to simulate the entire radiological imaging process using realistic digital phantoms, simulated image acquisition and reconstruction and AI-driven readers/computational observer models to improve the precision and accuracy of imaging systems and downstream biomarkers.

Table 4 Summary of illustrative AI applications across domains of oncology described in this review

<table>
<thead>
<tr>
<th>Skill development</th>
<th>Adjunct second reader</th>
<th>Segmentation aid</th>
<th>Skill assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection</td>
<td>Cancer screening (randomised), radiogenomics</td>
<td>Tissue compartments, subcellular morphology, mutation status, tissue segmentation</td>
<td>Patients at high risk of mortality (randomised)</td>
</tr>
<tr>
<td>Prediction</td>
<td>Tumour progression, treatment response</td>
<td>Overall survival (randomised), hospital admission (randomised)</td>
<td>Treatment de-escalation (prospective)</td>
</tr>
<tr>
<td>Radiological imaging</td>
<td>Pathology</td>
<td>Patient triage</td>
<td>Radiation oncology</td>
</tr>
</tbody>
</table>

Applications with randomised or prospective data are noted.

AI, artificial intelligence.
of investment by hospitals, companies and governments. Table 4 summarises AI applications across fields of oncology described in this review. We see common themes for AI used in disease detection, outcome prediction and education across specialties, and over the last few years, prospective and randomised evidence is accumulating in the domains of patient triage and radiological imaging. The primary limitation of AI in oncology has been a lack of validation. In the past several years, we are seeing more prospective trials and randomised trials, though these still remain largely single institutional. As higher levels of validation improve outcomes, we expect further coverage for AI tools by payors. Further advancements in these fields supported by the rise of NLP, DTs and CI will pave the way for the actualisation of AI in oncology in the next 5–10 years.

Author affiliations
1Department of Radiation Oncology, University of Washington, Seattle, Washington, USA
2Department of Radiation Oncology, Duke University, Durham, North Carolina, USA
3Department of Radiology, Duke University, Durham, North Carolina, USA
4Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina, USA
5Department of Radiation Oncology, Brigham and Women’s Hospital, Boston, Massachusetts, USA
6Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA
7Department of Otolaryngology–Head & Neck Surgery, University of Toronto, Toronto, Ontario, Canada
8KiNoN centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia
9NHMRC Clinical Trials Centre, Camperdown, New South Wales, Australia
10Faculty of Medicine, St. Vincent’s Clinical School, University of New South Wales, Sydney, New South Wales, Australia
11Department of Genetics and Genome Biology, University of Leicester Cancer Research Centre, Leicester, UK
12Microsoft Research, Redmond, Washington, USA
13Department of Radiation Oncology, University of South Florida, Tampa, Florida, USA
14Veterans Affairs Informatics and Computing Infrastructure, Salt Lake City, Utah, USA
15Department of Radiology, University of Washington, Seattle, Washington, USA

Twitter John Kang @DrJohnKang

Contributors Jk had overall editorial responsibility and wrote the triage and radiation oncology sections and the supporting sections. Respective section topics and their primary authors were: KL for Pathology, Ekim for Clinical informatics, CY for Surgery, FL for Genomics-driven precision oncology, TR for Surgery, HN for Large language models, EKatsoulakis for Digital twins and CIL for Radiological imaging. All authors approved the final manuscript.

Funding FL reports the following grants: Medical Research Future Fund as part of the Rapid Applied Research Translation initiative (RARU000125) and Capacity Building Grant funded by Cancer Institute NSW (2021/CB0003). CIL reports grants from the National Cancer Institute (R37CA240403, R01CA262023) and Kuni Foundation. TR was supported by the UK National Institute for Health Research (NIHR; grant code CL 2017-11-002), the NIHR Leicester Biomedical Research Centre, the Association of Breast Surgery (UK) and Innovate UK. This publication presents independent research funded by the NIHR. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests EKatsoulakis reports stock in AbbVie, Pfizer, Intuitive Surgical and Candel Therapeutics. CIL reports textbook royalties from McGraw Hill, Oxford University Press and UpToDate and personal fees from the American College of Radiology for journal editorial board work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Ethics approval Not applicable.

Provenance and peer review Commissioned; externally peer reviewed.

Data availability statement No data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

ORCID ids
John Kang http://orcid.org/0000-0003-3915-8350
Ellen Kim http://orcid.org/0000-0001-3579-0192

REFERENCES


