



# Artificial intelligence across oncology specialties: current applications and emerging tools

John Kang <sup>1</sup>, Kyle Lafata,<sup>2,3,4</sup> Ellen Kim <sup>5,6</sup>, Christopher Yao,<sup>7</sup> Frank Lin,<sup>8,9,10</sup> Tim Rattay,<sup>11</sup> Harsha Nori,<sup>12</sup> Evangelia Katsoulakis,<sup>13,14</sup> Christoph IIsuk Lee<sup>15</sup>

**To cite:** Kang J, Lafata K, Kim E, *et al.* Artificial intelligence across oncology specialties: current applications and emerging tools. *BMJ Oncology* 2024;**3**:e000134. doi:10.1136/bmjonc-2023-000134

Received 10 July 2023  
Accepted 03 January 2024

## 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,<sup>1</sup> stroke detection<sup>2</sup> and predicting hospital readmissions.<sup>3</sup>

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.<sup>4</sup> Despite these well-documented early setbacks,<sup>5</sup> continued refinement of multigene assays has resulted in their integration into the standard of care in certain contexts, notably in breast and prostate cancer.<sup>6–8</sup>

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.<sup>4</sup> 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.<sup>9–12</sup> For a more in-depth discussion about data sharing, ethics and validation, we refer readers to prior work.<sup>13</sup>

## 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.<sup>14</sup> Thus far, early cohort studies suggest that AI improves overall accuracy when used as an adjunct tool by radiologists interpreting



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

### Correspondence to

Dr John Kang; jkang3@uw.edu

mammograms.<sup>15</sup> A few studies suggest that AI as a stand-alone tool may be equivalent to human interpretative accuracy for breast cancer screening.<sup>16</sup> 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.<sup>17</sup> 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.<sup>18</sup> 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.<sup>19</sup> 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.<sup>20</sup> 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,<sup>21 22</sup> liver,<sup>23</sup> pancreas<sup>24 25</sup> and head and neck,<sup>26–28</sup> as well as for metastatic cancer.<sup>29</sup>

#### 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.<sup>30</sup>

#### 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.<sup>31</sup> 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.<sup>31</sup>

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,<sup>32</sup> bladder cancer,<sup>33</sup> renal cell carcinoma,<sup>34</sup> oropharyngeal cancer<sup>35</sup> and breast cancer<sup>36</sup> to predictors of microsatellite instability in colorectal cancer<sup>37</sup> and mutational status of BRAF-mutated melanomas.<sup>38</sup>

#### 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,<sup>39–41</sup> it is primarily a kind of image representation.<sup>42</sup> 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,<sup>43</sup> cellular detection and classification in colon cancer<sup>44</sup> and bone marrow,<sup>45</sup> counting tumour infiltrating lymphocytes<sup>46</sup> and immune cell composition<sup>47</sup> and segmentation of nuclei in cervical tissue for squamous epithelium cervical intraepithelial neoplasia grading.<sup>48</sup>

#### 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.<sup>49 50</sup> Several groups have performed prospective trials based on data-driven triage frameworks which we discuss here (table 1).<sup>51</sup>

#### 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,<sup>52</sup> they performed a randomised controlled trial for patients at high risk

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

Trial name/ dates	Trial design	Patient population	Clinical objective	Clinical action or arms	Model objective	Model methods
SHIELD-RT <sup>53</sup> (2019)	Randomised, controlled, quality improvement	Patients undergoing radiation	Decrease acute care visits	Two times per week versus standard weekly clinical visit	Predict >10% risk of acute care visit	Gradient boosted trees
University of Pennsylvania, Manz <i>et al</i> <sup>55 56</sup> (2019)	Prospective silent→phase III cluster wedge	New encounters in outpatient clinics	Increase serious illness conversations	Weekly email of high-risk patients with day of text message versus usual care	Predict 6-month mortality	Gradient boosted trees
Stanford, Gensheimer <i>et al</i> <sup>60</sup> (2020–2021)	Quality improvement, controlled	Patients in four clinics: two control, two experimental	Increase advance care planning	Weekly emails with patients predicted to survive<2 years	Predict time- dependent survival	Cox proportional hazards with tabular and text data
MASAI <sup>18</sup> (2021– 2022)	Randomised, controlled, non-inferiority screening accuracy study	Women eligible for routine mammography screening	Improve screening outcomes, including lower interval cancer rate	AI-supported screening versus standard double reading without AI	Identify cancers on mammogram images	Proprietary deep convolutional neural network

AI, artificial intelligence.

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%.<sup>53</sup>

#### 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).<sup>54</sup> After running a silent trial (ie, without alerting clinicians) of their GTB model confirming their prior results (table 2),<sup>55</sup> the authors followed up with an interventional randomised stepped-wedge cluster trial.<sup>56</sup> The clusters received either a behavioural nudge (weekly emails including up to six high-risk patients) or usual care. In the subset of high-risk encounters, nudges increased SIC from 3.6% to 15.2%,

more than tripling the increase seen in all encounters. Of note, the model sensitivity was approximately 25%, but event rate can be tuned to allow more high-risk predictions at the cost of higher risk of alert fatigue due to more false positives.<sup>57</sup>

Gensheimer *et al* at Stanford used clinical notes in addition to EHR, inpatient billing and registry data to train a Cox proportional hazards model to predict overall survival for metastatic solid tumours.<sup>58</sup> Text was modelled through a bag of words approach by tallying the top 100 000 one to two words phrases. In a subsequent study, life expectancy estimation was compared between this text-driven model, oncologists and a baseline performance status-based model, with the natural language processing (NLP)-driven model shown to be superior.<sup>59</sup> In a quality improvement interventional trial, the rate of advance care planning in clinics that received weekly emails of high-risk patients (predicted survival<2 years)

**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<sup>16</sup>

Total patients (24 582)	Predict survival>180 days (23 963)	Predict survival≤180 days (619*)	Event rate 2.5%
Survived ≤180 days (1022)	742 false negatives	280 true positives	Sensitivity 27.4%
Survived >180 days (23 560)	23221 true negatives	339 false positives	Specificity 98.6%
	NPV 96.9%	PPV 45.2%	

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 619/24 582=2.5% event rate.

NPV, negative predictive value; PPV, positive predictive value.

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

Trial name/ dates	Trial design	Patient population	Clinical objective	Clinical action or arms	Model objective	Model methods
INRT-AIR <sup>66</sup> (2019–2022)	Phase II	Patients with oropharynx, larynx squamous cell carcinoma	Risk of solitary elective nodal recurrence	Involved nodal radiation to suspicious nodes	Predict individual node positivity risk	Hybrid radiomics and three-dimensional CNN
DARTBOARD <sup>67</sup> (2022–)	Phase II randomised	Same as INRT- AIR	Decrease xerostomia	INRT±daily ART (near marginless)	CBCT autosegmentation via proprietary CNN architecture <sup>64</sup>	Same as INRT- AIR±AI-driven ART deformable registration and planning <sup>64</sup>
ARTIA-Lung <sup>65</sup> (2022–)	Randomised	Patients with stages IIIA and IIIB NSCLC	Decrease composite acute grade 3+toxicity	ART versus standard IMRT radiotherapy	Same as DARTBOARD	AI-driven ART deformable registration and planning <sup>64</sup>

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.

increased to 35% compared with 3% in a control cohort of clinics.<sup>60</sup> 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.<sup>61</sup> 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.<sup>62</sup> 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.<sup>63</sup> Proprietary online adaptive radiotherapy platforms incorporate AI-driven autosegmentation, deformable registration and treatment planning<sup>64</sup> with ongoing trials in several disease sites, including a randomised trial in lung cancer (table 3).<sup>65</sup>

#### 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.<sup>66</sup> Building on INRT-AIR, Sher *et al* are using INRT with or without near marginless daily adaptation in the DARTBOARD trial (table 3).<sup>67</sup>

The presence of extranodal extension (ENE) in neck nodes portends more aggressive cancer, though it is very challenging to detect on imaging.<sup>68</sup> Kann *et al* developed and validated a deep CNN model to predict pathologic ENE on CT imaging<sup>69</sup> 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.<sup>70</sup>

### 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.<sup>71</sup> Clinicogenomic signatures can also predict benefit from adjuvant chemotherapy in patients with gastric cancer,<sup>72</sup> and individual liver metastases with HER2-amplified metastatic colorectal cancer treated with dual HER2-targeted therapy with potential for detecting non-responding lesions.<sup>73</sup>



One emerging trend is to generate complex signatures predicting therapy utility by merging multimodal data.<sup>74</sup> 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.<sup>75</sup> In patients with castration-resistant prostate cancer, genomics and transcriptomics data have been combined using ML to predict a response to androgen receptor antagonists.<sup>76</sup>

#### Biomarker selection with non-destructive virtual assays

AI methods can derive surrogate biomarkers for targeting the molecular alterations of clinical significance in tumour tissues.<sup>77</sup> 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.<sup>78–80</sup> 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.<sup>81</sup> 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.<sup>82 83</sup>

#### 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.<sup>84</sup> 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<sup>85</sup> 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.<sup>86</sup>

#### 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.<sup>87</sup> AI algorithms can recognise the operative steps of a laparoscopic sleeve gastrectomy obtained from surgical videos with 85.6% accuracy.<sup>88</sup> 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.<sup>88</sup> In fact, experienced surgeons can be differentiated from beginners within the first 10 s, a task with 90% accuracy.<sup>89</sup> 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.<sup>90</sup> AI has also been used to determine optimal margins for cancer resection, such as in hepatic metastasectomy.<sup>91</sup> 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,<sup>92</sup> or to distinguish high-risk versus low-risk precancerous lesions.<sup>93</sup> ML has also been used to predict major complications after cancer surgery<sup>94</sup> and to identify patients who are safe to be discharged 2 days following major gastrointestinal cancer surgery.<sup>95</sup>

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.<sup>96–101</sup> As discussed above, Gensheimer *et al* have incorporated NLP into clinical triage.<sup>58</sup> NLP systems for clinical trial matching are in test deployment at large healthcare networks.<sup>102</sup>

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,<sup>103–105</sup> as well as to highly specialised topics such as medical physics.<sup>106</sup> 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.<sup>102 107 108</sup>

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.<sup>109</sup>

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.<sup>110</sup> 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.<sup>111</sup>

#### 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.<sup>112</sup> *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.<sup>113</sup> 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.<sup>114</sup> In a phase I trial, existing quantitative systems pharmacology model of the anti-CD20/CD3 T-cell engaging bispecific antibody, mosunetuzumab, were used to incorporate different dosing regimens and patient heterogeneity within the trial.<sup>115</sup>

#### 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 multireport prediction of metastases, enabling highly detailed representations that effectively model a cancer patient’s disease progression over time.<sup>116</sup> 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

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

	Radiological imaging	Pathology	Patient triage	Radiation oncology	Genomic-driven oncology	Surgery
Detection	Cancer screening (randomised), radiogenomics	Tissue compartments, subcellular morphology, mutation status, tissue segmentation	Patients at high risk of mortality (randomised)	Organ segmentation, nodal extracapsular extension	Virtual biomarker, mutation detection	Anatomy segmentation; operative steps; surgical skill
Prediction	Tumour progression, treatment response		Overall survival (randomised), hospital admission (randomised)	Treatment de-escalation (prospective)	Systemic therapy benefit, drug screening	Optimal margins; patient selection; complications; discharge planning
Skill development	Adjunct second reader			Segmentation aid		Skill assessment

Applications with randomised or prospective data are noted. AI, artificial intelligence.

trials were simulated to anticipate the optimal salvage therapy following progressive disease while on pembrolizumab.<sup>117</sup> For spine metastases, DT was used to simulate vertebroplasty and its impact on mechanical stability of the vertebra.<sup>118</sup> 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 on DTs.<sup>119</sup> Further research, validation and clinical trials are needed to fully establish the effectiveness and integration of DTs into routine clinical practice in oncology.

### 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<sup>120 121</sup> 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.<sup>122</sup> 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<sup>123</sup> 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.<sup>124</sup> 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.<sup>125</sup> 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 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 evidence lead to improved 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

- <sup>1</sup>Department of Radiation Oncology, University of Washington, Seattle, Washington, USA
- <sup>2</sup>Department of Radiation Oncology, Duke University, Durham, North Carolina, USA
- <sup>3</sup>Department of Radiology, Duke University, Durham, North Carolina, USA
- <sup>4</sup>Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina, USA
- <sup>5</sup>Department of Radiation Oncology, Brigham and Women's Hospital, Boston, Massachusetts, USA
- <sup>6</sup>Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA
- <sup>7</sup>Department of Otolaryngology-Head & Neck Surgery, University of Toronto, Toronto, Ontario, Canada
- <sup>8</sup>Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia
- <sup>9</sup>NHMRC Clinical Trials Centre, Camperdown, New South Wales, Australia
- <sup>10</sup>Faculty of Medicine, St Vincent's Clinical School, University of New South Wales, Sydney, New South Wales, Australia
- <sup>11</sup>Department of Genetics and Genome Biology, University of Leicester Cancer Research Centre, Leicester, UK
- <sup>12</sup>Microsoft Research, Redmond, Washington, USA
- <sup>13</sup>Department of Radiation Oncology, University of South Florida, Tampa, Florida, USA
- <sup>14</sup>Veterans Affairs Informatics and Computing Infrastructure, Salt Lake City, Utah, USA
- <sup>15</sup>Department 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 (RARUR00125) and Capacity Building Grant funded by Cancer Institute NSW (2021/CBG003). 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.

**Patient consent for publication** Not applicable.

**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/0000-0001-9379-0192>

#### REFERENCES

- 1 Miliard M. *Google, Verily Using AI to Screen for Diabetic Retinopathy in India*. Healthcare IT News 2019. Available: <https://www.healthcareitnews.com/news/asia/google-verily-using-ai-screen-diabetic-retinopathy-india> [accessed 14 May 2023].
- 2 Morey JR, Zhang X, Yaeger KA, *et al*. Real-world experience with artificial intelligence-based triage in transferred large vessel occlusion stroke patients. *Cerebrovasc Dis* 2021;50:450–5.
- 3 Hwang AB, Schuepfer G, Pietrini M, *et al*. External validation of EPIC's risk of unplanned readmission model, the LACE+ index and Sqlape as predictors of unplanned hospital readmissions: a monocentric, retrospective, diagnostic cohort study in Switzerland. *PLoS ONE* 2021;16:e0258338.
- 4 Kang J, Rancati T, Lee S, *et al*. Machine learning and radiogenomics: lessons learned and future directions. *Front Oncol* 2018;8:228.
- 5 Micheel C, Nass SJ, Omenn GS, *et al*. Committee on the review of Omics-based tests for predicting patient outcomes in clinical trials. In: *Evolution of Translational Omics: Lessons Learned and the Path Forward*. National Academies Press, 2012.
- 6 Lehman CD, Wellman RD, Buist DSM, *et al*. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;175:1828–37.
- 7 Sparano JA, Gray RJ, Makower DF, *et al*. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N Engl J Med* 2018;379:111–21.
- 8 Zhao SG, Chang SL, Spratt DE, *et al*. Development and validation of a 24-gene predictor of response to postoperative radiotherapy in prostate cancer: a matched, retrospective analysis. *Lancet Oncol* 2016;17:1612–20.
- 9 Kang J, Schwartz R, Flickinger J, *et al*. Machine learning approaches for predicting radiation therapy outcomes: a clinician's perspective. *Int J Radiat Oncol Biol Phys* 2015;93:1127–35.
- 10 Kann BH, Hosny A, Aerts H. Artificial intelligence for clinical oncology. *Cancer Cell* 2021;39:916–27.
- 11 Bhinder B, Gilvary C, Madhukar NS, *et al*. Artificial intelligence in cancer research and precision medicine. *Cancer Discov* 2021;11:900–15.
- 12 Shimizu H, Nakayama KI. Artificial intelligence in oncology. *Cancer Sci* 2020;111:1452–60. 10.1111/cas.14377 Available: <https://onlinelibrary.wiley.com/doi/10.1111/cas.14377>
- 13 Kang J, Thompson RF, Aneja S, *et al*. National Cancer Institute workshop on artificial intelligence in radiation oncology: training the next generation. *Practical Radiation Oncology* 2021;11:74–83.
- 14 Schaffter T, Buist DSM, Lee CI, *et al*. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open* 2020;3:e200265.
- 15 Houssami N, Kirkpatrick-Jones G, Noguchi N, *et al*. Artificial intelligence (AI) for the early detection of breast cancer: a scoping review to assess AI's potential in breast screening practice. *Expert Rev Med Devices* 2019;16:351–62.
- 16 Anderson AW, Marinovich ML, Houssami N, *et al*. Independent external validation of artificial intelligence algorithms for automated interpretation of screening mammography: a systematic review. *J Am Coll Radiol* 2022;19(2 Pt A):259–73.
- 17 Lee CI, Houssami N, Elmore JG, *et al*. Pathways to breast cancer screening artificial intelligence algorithm validation. *The Breast* 2020;52:146–9.



- 18 Lång K, Josefsson V, Larsson A-M, *et al.* Artificial intelligence-supported screen reading versus standard double reading in the mammography screening with artificial intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *Lancet Oncol* 2023;24:936–44.
- 19 Paudyal R, Shah AD, Akin O, *et al.* Artificial intelligence in CT and MR imaging for oncological applications. *Cancers (Basel)* 2023;15:2573.
- 20 Jaffe CC. Measures of response: RECIST, WHO, and new alternatives. *JCO* 2006;24:3245–51.
- 21 Owens C, Hindocha S, Lee R, *et al.* The lung cancers: staging and response, CT, 18F-FDG PET/CT, MRI, DWI: review and new perspectives. *Br J Radiol* 2023;96:20220339.
- 22 Lafata KJ, Hong JC, Geng R, *et al.* Association of pre-treatment radiomic features with lung cancer recurrence following stereotactic body radiation therapy. *Phys Med Biol* 2019;64:025007.
- 23 Fahmy D, Alksas A, Elnakib A, *et al.* The role of radiomics and AI technologies in the segmentation, detection, and management of hepatocellular carcinoma. *Cancers (Basel)* 2022;14:6123.
- 24 Rigioli F, Hoye J, Lerebours R, *et al.* CT Radiomic features of superior mesenteric artery involvement in pancreatic ductal adenocarcinoma: a pilot study. *Radiology* 2021;301:610–22.
- 25 Rigioli F, Hoye J, Lerebours R, *et al.* Exploratory analysis of mesenteric-portal axis CT radiomic features for survival prediction of patients with pancreatic ductal adenocarcinoma. *Eur Radiol* 2023;33:5779–91.
- 26 Lafata KJ, Chang Y, Wang C, *et al.* Intrinsic radiomic expression patterns after 20 Gy demonstrate early metabolic response of oropharyngeal cancers. *Medical Physics* 2021;48:3767–77. 10.1002/mp.14926 Available: <https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.14926>
- 27 Wang C, Liu C, Chang Y, *et al.* Dose-distribution-driven PET image-based outcome prediction (DDD-PIOP): a deep learning study for oropharyngeal cancer IMRT application. *Front Oncol* 2020;10:1592.
- 28 Ji H, Lafata K, Mowery Y, *et al.* Post-radiotherapy PET image outcome prediction by deep learning under biological model guidance: a feasibility study of oropharyngeal cancer application. *Front Oncol* 2022;12:895544.
- 29 Rompianesi G, Pegoraro F, Ceresa CD, *et al.* Artificial intelligence in the diagnosis and management of colorectal cancer liver metastases. *WJG* 2022;28:108–22.
- 30 Bera K, Schalper KA, Rimm DL, *et al.* Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019;16:703–15.
- 31 Lafata KJ, Wang Y, Konkel B, *et al.* Radiomics: a primer on high-throughput image phenotyping. *Abdom Radiol (NY)* 2022;47:2986–3002.
- 32 Chen D, Fu M, Chi L, *et al.* Prognostic and predictive value of a pathomics signature in gastric cancer. *Nat Commun* 2022;13:6903.
- 33 Chen S, Jiang L, Zheng X, *et al.* Clinical use of machine learning-based pathomics signature for diagnosis and survival prediction of bladder cancer. *Cancer Sci* 2021;112:2905–14.
- 34 Chen S, Jiang L, Gao F, *et al.* Machine learning-based pathomics signature could act as a novel Prognostic marker for patients with clear cell renal cell carcinoma. *Br J Cancer* 2022;126:771–7.
- 35 Ali S, Lewis J, Madabhushi A. Spatially aware cell cluster(spacc1) graphs: predicting outcome in oropharyngeal P16+ tumors. *Med Image Comput Assist Interv MICCAI Int Conf Med Image Comput Assist Interv* 2013;16(Pt 1):412–9.
- 36 Heindl A, Sestak I, Naidoo K, *et al.* Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *J Natl Cancer Inst* 2018;110.
- 37 Cao R, Yang F, Ma S-C, *et al.* Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer. *Theranostics* 2020;10:11080–91.
- 38 Kim RH, Nomikou S, Coudray N, *et al.* Deep learning and pathomics analyses reveal cell nuclei as important features for Mutation prediction of BRAF-mutated melanomas. *J Invest Dermatol* 2022;142:1650–8.
- 39 Campanella G, Hanna MG, Geneslaw L, *et al.* Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 2019;25:1301–9.
- 40 Song Z, Zou S, Zhou W, *et al.* Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 2020;11:4294.
- 41 Wetstein SC, de Jong VMT, Stathonikos N, *et al.* Deep learning-based breast cancer grading and survival analysis on whole-slide Histopathology images. *Sci Rep* 2022;12:15102.
- 42 Barisoni L, Lafata KJ, Hewitt SM, *et al.* Digital Pathology and computational image analysis in Nephropathology. *Nat Rev Nephrol* 2020;16:669–85.
- 43 Bulten W, Bándi P, Hoven J, *et al.* Epithelium segmentation using deep learning in H&E-stained prostate specimens with immunohistochemistry as reference standard. *Sci Rep* 2019;9:864.
- 44 Sirinukunwattana K, Ahmed Raza SE, *et al.* Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans Med Imaging* 2016;35:1196–206.
- 45 Wang C-W, Huang S-C, Lee Y-C, *et al.* Deep learning for bone marrow cell detection and classification on whole-slide images. *Med Image Anal* 2022;75:102270.
- 46 Linder N, Taylor JC, Colling R, *et al.* Deep learning for detecting tumour-infiltrating lymphocytes in Testicular germ cell tumours. *J Clin Pathol* 2019;72:157–64.
- 47 Aprupe L, Litjens G, Brinker TJ, *et al.* Robust and accurate quantification of biomarkers of immune cells in lung cancer micro-environment using deep convolutional neural networks. *PeerJ* 2019;7:e6335.
- 48 Sornapudi S, Stanley RJ, Stoecker WV, *et al.* Deep learning nuclei detection in digitized histology images by superpixels. *J Pathol Inform* 2018;9:5.
- 49 Taniyama TK, Hashimoto K, Katsumata N, *et al.* Can oncologists predict survival for patients with progressive disease after standard Chemotherapies. *Curr Oncol* 2014;21:84–90.
- 50 Kim YJ, Yoon SJ, Suh S-Y, *et al.* Performance of clinician prediction of survival in oncology outpatients with advanced cancer. *PLOS ONE* 2022;17:e0267467.
- 51 Kang J, Chowdhry AK, Pugh SL, *et al.* Integrating artificial intelligence and machine learning into cancer clinical trials. *Semin Radiat Oncol* 2023;33:386–94.
- 52 Hong JC, Niedzwiecki D, Palta M, *et al.* Predicting emergency visits and hospital admissions during radiation and chemoradiation: an internally validated pretreatment machine learning algorithm. *JCO Clin Cancer Inform* 2018;2:1–11.
- 53 Hong JC, Eclov NCW, Dalal NH, *et al.* System for high-intensity evaluation during radiation therapy (SHIELD-RT): a prospective randomized study of machine learning-directed clinical evaluations during radiation and chemoradiation. *J Clin Oncol* 2020;38:3652–61.
- 54 Parikh RB, Manz C, Chivers C, *et al.* Machine learning approaches to predict 6-month mortality among patients with cancer. *JAMA Netw Open* 2019;2:e1915997.
- 55 Manz CR, Chen J, Liu M, *et al.* Validation of a machine learning algorithm to predict 180-day mortality for outpatients with cancer. *JAMA Oncol* 2020;6:1723.
- 56 Manz CR, Parikh RB, Small DS, *et al.* Effect of integrating machine learning mortality estimates with behavioral nudges to clinicians on serious illness conversations among patients with cancer: a stepped-wedge cluster randomized clinical trial. *JAMA Oncol* 2020;6:e204759.
- 57 Kang J, Morin O, Hong JC. Closing the gap between machine learning and clinical cancer care—first steps into a larger world. *JAMA Oncol* 2020;6:1731–2.
- 58 Gensheimer MF, Henry AS, Wood DJ, *et al.* Automated survival prediction in metastatic cancer patients using high-dimensional electronic medical record data. *J Natl Cancer Inst* 2019;111:568–74.
- 59 Gensheimer MF, Aggarwal S, Benson KRK, *et al.* Automated model versus treating physician for predicting survival time of patients with metastatic cancer. *J Am Med Inform Assoc* 2021;28:1108–16.
- 60 Gensheimer MF, Gupta D, Patel MI, *et al.* Use of machine learning and lay care coaches to increase advance care planning conversations for patients with metastatic cancer. *JCO Oncol Pract* 2023;19:e176–84.
- 61 Harrison K, Pullen H, Welsh C, *et al.* Machine learning for auto-segmentation in radiotherapy planning. *Clinical Oncology* 2022;34:74–88.
- 62 Walker GV, Avran M, Tao R, *et al.* Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. *Radiotherapy and Oncology* 2014;112:321–5.
- 63 Heukelum J, Fuller CD. Head and neck cancer adaptive radiation therapy (ART): conceptual considerations for the informed clinician. *Semin Radiat Oncol* 2019;29:258–73.
- 64 Archambault Y, Boylan C, Bullock D, *et al.* Making on-line adaptive radiotherapy possible using artificial intelligence and machine learning for efficient daily re-planning. *Med Phys Int* 2020;8:77–86.
- 65 Varian, a Siemens Healthineers Company. Daily adaptive vs non-adaptive external beam radiation therapy with concurrent chemotherapy for locally advanced non-small cell lung cancer:

- a prospective randomized trial of an individualized approach for toxicity reduction (ARTIA-lung). *ClinicaltrialsGov* 2022. Available: <https://clinicaltrials.gov/ct2/show/NCT05488626>
- 66 Sher DJ, Avksholt V, Moon D, *et al.* Recurrence and quality-of-life following involved node radiotherapy for head and neck squamous cell carcinoma: initial results from the phase II INRT-air trial. *Int J Radia Oncol\*Biolo\*Phy* 2021;111:e398.
  - 67 Sher D. DARTBOARD: a prospective randomized phase II study of daily adaptive radiotherapy to better organ-at-risk doses in head and neck cancer. *ClinicaltrialsGov*; 2022. Available: <https://clinicaltrials.gov/ct2/show/NCT04883281>
  - 68 Sahin O, Wahid KA, Multidisciplinary OPC ENE Group. Multi-specialty expert physician identification of Extranodal extension in computed tomography scans of oropharyngeal cancer patients: prospective blinded human inter-observer performance evaluation. *Oncology* 2023.02.25.23286432. [Preprint] 2023.
  - 69 Kann BH, Hicks DF, Payabvash S, *et al.* Multi-institutional validation of deep learning for pretreatment identification of extranodal extension in head and neck squamous cell carcinoma. *J Clin Oncol* 2020;38:1304–11.
  - 70 Kann BH, Likitlersuang J, Ye Z, *et al.* Screening for extranodal extension with deep learning: evaluation in ECOG-ACRIN E3311, a randomized de-escalation trial for HPV-associated oropharyngeal carcinoma. *Int J Radia Oncol\*Biolo\*Phy* 2022;114:S26–7.
  - 71 Nazha A, Komrokji R, Meggendorfer M, *et al.* Personalized prediction model to risk stratify patients with myelodysplastic syndromes. *JCO* 2021;39:3737–46.
  - 72 Sundar R, Barr Kumarakulasinghe N, Huak Chan Y, *et al.* Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomised phase III SAMIT trial. *Gut* 2022;71:676–85.
  - 73 Giannini V, Rosati S, Defeudis A, *et al.* Radiomics predicts response of individual Her2-amplified colorectal cancer liver metastases in patients treated with Her2-targeted therapy. *Int J Cancer* 2020;147:3215–23.
  - 74 Lipkova J, Chen RJ, Chen B, *et al.* Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* 2022;40:1095–110.
  - 75 Vanguri RS, Luo J, Aukerman AT, *et al.* Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer* 2022;3:1151–64.
  - 76 de Jong AC, Danyi A, van Riet J, *et al.* Predicting response to enzalutamide and abiraterone in metastatic prostate cancer using whole-omics machine learning. *Nat Commun* 2023;14:1968.
  - 77 Cifci D, Foersch S, Kather JN. Artificial intelligence to identify genetic alterations in conventional Histopathology. *J Pathol* 2022;257:430–44.
  - 78 Echle A, Ghaffari Laleh N, Quirke P, *et al.* Artificial intelligence for detection of microsatellite instability in colorectal cancer—a multicentric analysis of a pre-screening tool for clinical application. *ESMO Open* 2022;7:100400.
  - 79 Gerwert K, Schörner S, Großberueschkamp F, *et al.* Fast and label-free automated detection of microsatellite status in early colon cancer using artificial intelligence integrated infrared imaging. *European Journal of Cancer* 2023;182:122–31.
  - 80 Schrammen PL, Ghaffari Laleh N, Echle A, *et al.* Weakly supervised annotation-free cancer detection and prediction of genotype in routine Histopathology. *J Pathol* 2022;256:50–60.
  - 81 Nair JKR, Saeed UA, McDougall CC, *et al.* Radiogenomic models using machine learning techniques to predict EGFR mutations in non-small cell lung cancer. *Can Assoc Radiol J* 2021;72:109–19.
  - 82 Lafata KJ, Corradetti MN, Gao J, *et al.* Radiogenomic analysis of locally advanced lung cancer based on CT imaging and Intratreatment changes in cell-free DNA. *Radiol Imaging Cancer* 2021;3:e200157.
  - 83 Corradetti MN, Torok JA, Hatch AJ, *et al.* Dynamic changes in circulating tumor DNA during chemoradiation for locally advanced lung cancer. *Adv Radiat Oncol* 2019;4:748–52.
  - 84 Petak I, Kamal M, Dirner A, *et al.* A computational method for prioritizing targeted therapies in precision oncology: performance analysis in the Shiva01 trial. *NPJ Precis Oncol* 2021;5:59.
  - 85 Sakellaropoulos T, Vougas K, Narang S, *et al.* A deep learning framework for predicting response to therapy in cancer. *Cell Rep* 2019;29:3367–73.
  - 86 Kong J, Ha D, Lee J, *et al.* Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nat Commun* 2022;13:3703.
  - 87 Madani A, Namazi B, Altieri MS, *et al.* Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg* 2022;276:363–9.
  - 88 Hashimoto DA, Rosman G, Witkowski ER, *et al.* Computer vision analysis of intraoperative video: automated recognition of operative steps in laparoscopic sleeve gastrectomy. *Ann Surg* 2019;270:414–21.
  - 89 French A, Lendvay TS, Sweet RM, *et al.* Predicting surgical skill from the first N seconds of a task: value over task time using the Isogony principle. *Int J Comput Assist Radiol Surg* 2017;12:1161–70.
  - 90 Haifler M, Pence I, Sun Y, *et al.* Discrimination of malignant and normal kidney tissue with short wave infrared dispersive Raman spectroscopy. *J Biophotonics* 2018;11:e201700188.
  - 91 Bertsimas D, Margonis GA, Sujichantararat S, *et al.* Using artificial intelligence to find the optimal margin width in hepatectomy for colorectal cancer liver metastases. *JAMA Surg* 2022;157:e221819.
  - 92 Zheng J, Chakraborty J, Chapman WC, *et al.* Preoperative prediction of microvascular invasion in hepatocellular carcinoma using quantitative image analysis. *J Am Coll Surg* 2017;225:778–788e1.
  - 93 Attiyeh MA, Chakraborty J, Gazit L, *et al.* Preoperative risk prediction for intraductal papillary mucinous neoplasms by quantitative CT image analysis. *HPB* 2019;21:212–8.
  - 94 Deng H, Eftekhari Z, Carlin C, *et al.* Development and validation of an explainable machine learning model for major complications after cytoreductive surgery. *JAMA Netw Open* 2022;5:e2212930.
  - 95 van de Sande D, van Genderen ME, Verhoef C, *et al.* Optimizing discharge after major surgery using an artificial intelligence-based decision support tool (DESIRE): an external validation study. *Surgery* 2022;172:663–9.
  - 96 Kehl KL, Elmarakeby H, Nishino M, *et al.* Assessment of deep natural language processing in ascertaining oncologic outcomes from radiology reports. *JAMA Oncol* 2019;5:1421–9.
  - 97 Kehl KL, Xu W, Lepisto E, *et al.* Natural language processing to ascertain cancer outcomes from medical oncologist notes. *JCO Clin Cancer Inform* 2020;4:680–90.
  - 98 Kehl KL, Xu W, Gusev A, *et al.* Artificial intelligence-aided clinical annotation of a large multi-cancer genomic dataset. *Nat Commun* 2021;12:7304.
  - 99 Banerjee I, Bozkurt S, Caswell-Jin JL, *et al.* Natural language processing approaches to detect the timeline of metastatic recurrence of breast cancer. *JCO Clin Cancer Inform* 2019;3:1–12.
  - 100 Sanyal J, Tariq A, Kurian AW, *et al.* Weakly supervised temporal model for prediction of breast cancer distant recurrence. *Sci Rep* 2021;11:9461.
  - 101 Karimi YH, Blayney DW, Kurian AW, *et al.* Development and use of natural language processing for identification of distant cancer recurrence and sites of distant recurrence using unstructured electronic health record data. *JCO Clin Cancer Inform* 2021;5:469–78.
  - 102 Wong C, Zhang S, Gu Y, *et al.* Scaling clinical trial matching using large language models: a case study in oncology. 2023. Available: <http://arxiv.org/abs/2308.02180>
  - 103 Bhayana R, Bleakney RR, Krishna S. GPT-4 in radiology: improvements in advanced reasoning. *Radiology* 2023;307:e230987.
  - 104 Nori H, King N, McKinney SM, *et al.* Capabilities of GPT-4 on medical challenge problems. 2023. Available: <https://arxiv.org/abs/2303.13375v2>
  - 105 Singhal K, Azizi S, Tu T, *et al.* Large language models encode clinical knowledge. *Nature* 2023;620:172–80.
  - 106 Holmes J, Liu Z, Zhang L, *et al.* Evaluating large language models on a highly-specialized topic, radiation oncology physics. *Front Oncol* 2023;13:1219326.
  - 107 Singhal K, Tu T, Gottweis J, *et al.* Towards expert-level medical question answering with large language models. *arXiv.org*. Available: <https://arxiv.org/abs/2305.09617v1> [Accessed 6 Jun 2023].
  - 108 Liu Z, Wang P, Li Y, *et al.* Radonc-GPT: a large language model for radiation oncology. 2023. Available: [10.48550/arXiv.2309.10160](https://arxiv.org/abs/2309.10160)
  - 109 Huang K, Altosaar J, Ranganath R. Clinicalbert: modeling clinical notes and predicting hospital readmission. *ArXiv190405342 Cs*; 2020. Available: <http://arxiv.org/abs/1904.05342> [Accessed 28 May 2021].
  - 110 Moztarzadeh O, Jamshidi MB, Sargolzaei S, *et al.* Metaverse and healthcare: machine learning-enabled digital twins of cancer. *Bioengineering (Basel)* 2023;10:455.
  - 111 Sahal R, Alsamhi SH, Brown KN. Personal digital twin: a close look into the present and a step towards the future of personalised healthcare industry. *Sensors (Basel)* 2022;22:15.

- 112 Shaker B, Ahmad S, Lee J, *et al*. In silico methods and tools for drug discovery. *Comput Biol Med* 2021;137:104851.
- 113 Davies J, Martinec M, Delmar P, *et al*. Comparative effectiveness from a single-arm trial and real-world data: alectinib versus Ceritinib. *J Comp Eff Res* 2018;7:855–65.
- 114 Wedam S, Fashoyin-Aje L, Bloomquist E, *et al*. FDA approval summary: palbociclib for male patients with metastatic breast cancer. *Clin Cancer Res* 2020;26:1208–12.
- 115 Susilo ME, Li C-C, Gadkar K, *et al*. Systems-based digital twins to help characterize clinical dose-response and propose predictive biomarkers in a phase I study of Bispecific antibody, Mosunetuzumab, in NHL. *Clin Transl Sci* 2023;16:1134–48.
- 116 Batch KE, Yue J, Darcovich A, *et al*. Developing a cancer digital twin: supervised metastases detection from consecutive structured radiology reports. *Front Artif Intell* 2022;5:826402.
- 117 Qi T, Cao Y. Virtual clinical trials: a tool for predicting patients who may benefit from treatment beyond progression with pembrolizumab in non-small cell lung cancer. *CPT Pharmacom & Syst Pharma* 2023;12:236–49. 10.1002/psp4.12896 Available: <https://ascpt.onlinelibrary.wiley.com/toc/21638306/12/2>
- 118 Ahmadian H, Mageswaran P, Walter BA, *et al*. A digital twin for simulating the vertebroplasty procedure and its impact on mechanical stability of vertebra in cancer patients. *Numer Methods Biomed Eng* 2022;38. 10.1002/cnm.3600 Available: <https://onlinelibrary.wiley.com/toc/20407947/38/6>
- 119 Abadi E, Segars WP, Tsui BMW, *et al*. Virtual clinical trials in medical imaging: a review. *J Med Imag* 2020;7:1.
- 120 Shreve JT, Khanani SA, Haddad TC. Artificial intelligence in oncology: current capabilities, future opportunities, and ethical considerations. *Am Soc Clin Oncol Educ Book* 2022;42:1–10.
- 121 Lahmi L, Mamzer M-F, Burgun A, *et al*. Ethical aspects of artificial intelligence in radiation oncology. *Semin Radiat Oncol* 2022;32:442–8.
- 122 Detmer DE, Lumpkin JR, Williamson JJ. Defining the medical subspecialty of clinical informatics. *J Am Med Infor Assoc* 2009;16:167–8.
- 123 Kim E, Bitterman DS, Kann BH, *et al*. Hidden in plain sight: clinical Informaticians are the oncology subspecialists you did not know you needed. *Clinical Oncology* 2022;34:135–40.
- 124 Blasiak A, Khong J, Kee T. CURATE.AI: optimizing personalized medicine with artificial intelligence. *SLAS Technol* 2020;25:95–105.
- 125 Wang Y, Li X, Konanur M, *et al*. Towards optimal deep fusion of imaging and clinical data via a model-based description of fusion quality. *Med Phys* 2023;50:3526–37. 10.1002/mp.16181 Available: <https://aapm.onlinelibrary.wiley.com/toc/24734209/50/6>