ejHEAA
ISSN : 3032-1123
https://doi.org/10.61796/ejheaa.v1i3.418

DEVELOPING NOVEL IMAGING MODALITIES FOR EARLY CANCER DETECTION

Shatha F. Murad
Physiology Department, Medicine College, Al-Muthanna University, Al-Muthanna, Iraq

Received: Jan 22, 2024; Accepted: Feb 29, 2024; Published: March 18, 2024;

Abstract: An ever-expanding suite of cancer imaging tools is being created with the help of AI and ML. To design the best tool, it's important to include experts from other fields to determine the right use case, then test and refine the tool thoroughly before implementing it into healthcare systems. Showcasing significant advancements in the field, this interdisciplinary study. We go over the pros and downsides of using AI and ML for cancer imaging, some things to keep in mind when turning algorithms into tools for widespread use, and how to build an ecosystem that will help AI and ML expand in this field.

This is an open-acces article under the CC-BY 4.0 license

Introduction

Machine learning (ML) and artificial intelligence (AI) are quickly changing the scientific world, and many areas of medicine are not immune. Artificial intelligence (AI) is the study and development of systems that can mimic human intelligence and behaviour; machine learning (ML) is a branch of AI that focuses on teaching computers to learn from data and then use that information to generate predictions or classifications, with or without human intervention. The advent of high speed computers in the last few years has sped up the progress in these fields.

Digital medical imaging and other related fields are well-suited to be pioneers in the use of artificial intelligence and machine learning. The whole imaging process, including taking images, processing them, reporting on them, and sharing the findings, is carried out digitally, making it easier to gather data for AI and ML. Specifically, radiologists are expected to be the first to investigate and use new technologies in the field of cancer imaging since it accounts for a significant percentage of the workload in many departments. This is particularly true because these tasks can be monotonous (like reading through a mountain of normal studies to find abnormalities in cancer screening), taxing (like taking serial measurements of tumours), or repetitive (like outlining tumours for disease segmentation). In fact, the cancer imaging field is already home to a variety of commercial solutions that strive to streamline processes, cut down on mistakes, and boost diagnostic accuracy.

Problematically, many technical solutions are being created independently, which means they may not be able to make it into normal clinical use. These may have been hindered because there were few chances for experts in the field to collaborate and gain a better understanding of the clinical and data science landscape. This would have helped them identify the opportunities, risks, needs, and challenges associated with developing, testing, and adopting such tools. In order to promote innovations and advancements, it is necessary to foster collaborative interdisciplinary ecosystems, which may include commercial partners when suitable.

Results and Discussion

Here, we lay out the necessary AI and ML methods and emphasise the most promising avenues for applying these technologies to cancer imaging. We will go over the technical, professional, and clinical hurdles that come with using AI and ML for cancer imaging. From historical examples, we extrapolate the necessary technological and infrastructural advancements for artificial intelligence (AI) in cancer imaging, which will pave the way for the incorporation of ML and AI into healthcare systems and the proper education of the next generation of workers.

Imaging data in medicine: Radiomics. The majority of medical image analysis is still done by trained radiologists. These professionals can visually detect illness, tumour borders, treatment efficacy, and recurrence. When evaluating AI and ML approaches, these human talents are often utilised as reference standards. Nevertheless, there is a growing fascination in probing the constituent parts of medical images—the pixels and voxels—as imaging data. These parts are amenable to computational analysis, which might lead to the discovery of objective mathematical patterns associated with disease behaviour or outcomes.

The field known as "radiomics" uses computers to analyse specific areas inside medical pictures. The images can be two-dimensional, like a 2D X-ray or a three-dimensional computed tomography (CT) scan, or four-dimensional, like an ultrasound. They can also be scalar- or vector-valued, like a phase-contrast magnetic resonance imaging (MRI) scan, showing a relationship between the measured signal and a mathematical vector function, respectively. Using algorithms that can detect patterns in images—often beyond what the human eye can perceive—and capitalising on them to generate forecasts and so assist in clinical decision-making is the primary objective of radiomics. Many imaging characteristics are often produced by computerised image processing. To construct a mathematical model that can address the pertinent clinical question—the so-called ground truth variable—it is necessary to choose characteristics that are non-redundant, stable, and relevant. Figure 1 shows the process of choosing and evaluating radiomics characteristics to see whether they can differentiate between benign and malignant breast tumours in a given use-case. An further development in the field is the rise of radiogenomics, which combines genomics with radiomics to help in illness management via integrated diagnostics3-4

Consider a volumetric chest CT scan that includes a tumour, such as a lung nodule. In this case, the data set might be used for radiomics analysis. The normal workflow would include two steps: (1) identifying the cancer within the image; and (2) annotating the tumour with semantic attributes, typically by experienced radiologists.5. Third, the tumour must be outline or segmented. 4. Next, the features of the tumour, such as its size, mean intensity, image texture, shape, and margin sharpness, must be computed. This can be done manually or with the use of automated learning. 5. Finally, a classifier must be built to use these computed features to predict a clinical state.likelihood of a certain gene mutation, therapeutic efficacy, or total survival. 10,11

In order to make pipeline data analysis easier, many organisations are developing radiomics processing tools. The Quantitative Image Feature Pipeline12 was built at Stanford University and includes an extensible library of algorithms for quantitative imaging feature extraction and predictive modelling. It can characterise the imaging phenotype comprehensively and provides cloud-based software for making and running pipelines that generate and predict quantitative image features. Users can also use and compare image features to predict clinical and molecular features. Users may also include their own algorithms into a customisable process by uploading them as Docker containers 13.applied to cancer imaging using AI and ML methods. To guarantee data conformance or consistency, cancer imaging pre-processes and transforms patient pictures before using them as inputs to build ML algorithms and models. Whether the characteristics are specified by radiologists or generated from mathematics, these pre-processing procedures are always employed. Making sure the pictures have the same pixel-dimensions and image-section thickness is part of this process. To summarise, ML models and algorithms take in imaging data as input, create a map of the data, and then develop a mathematical function—either basic or complex—that is associated with the output, which might be anything from a clinical or scientific observation to the goal itself. When building or training an ML algorithm, it is not necessary to utilise ground truth variables. These are reference findings that have been verified by domain experts or other methods, such as pathology, laboratory testing, or clinical follow-up. Standard procedure for machine learning algorithms often involves creating a training dataset, refining it using a validation dataset, and then testing it on an independent test dataset, preferably one from a separate university, to see how well it performs.

When it comes to imaging research, some ML models are much more popular than others. The most popular version is the predictive model, which attempts to forecast y by learning the f(x) function. This assumes on which x is dependent, f is the mathematical function, and "y" stands for "target" or "output.". It is possible to try to establish a relationship between the input data x (such as an imaging characteristic) and the output y (such as gene expression) in exploratory models.

Multiple regression models may be used when dealing with continuous data. Some examples are Linear, Cox (Proportional Hazards), Regression Trees, Lasso, Ridge, ElasticNet, and others 14,15. Many classification methods may be used for discrete variables, including Naïve Bayes, Support Vector Intelligent systems, Decision Trees, Random Forests, and KNN (k-nearest neighbours) Generalised Linear methods, Bagging, and others. The use of these models has the potential to improve cancer detection, illness categorization and stratification, treatment efficacy, and overall health outcomes.

Data accessibility, computing power of machines, and further algorithm improvements all impact an ML algorithm's performance. Size of the data could dictate the ML method used. Classical ML methods like Naïve Bayes, logistic regression, Typically, smaller datasets are used while working with decision trees and support vector machines. such as less than 1000 patients, exams, or photographs, depending on the use case. Though they are more computationally intensive, more complicated ML models—like convolutional neural networks (CNN)—may be better with more datasets since they are very efficient at learning straight from pictures. Deep learning, of which CNN are a subset, is a category of machine learning techniques that use artificial neural networks. In order to solve issues, artificial neural networks mimic the connections of neurons, taking their cue from the way neurons are organised in the brain. Some ML algorithms are supervised, meaning they are trained on data that already has labels attached to it.

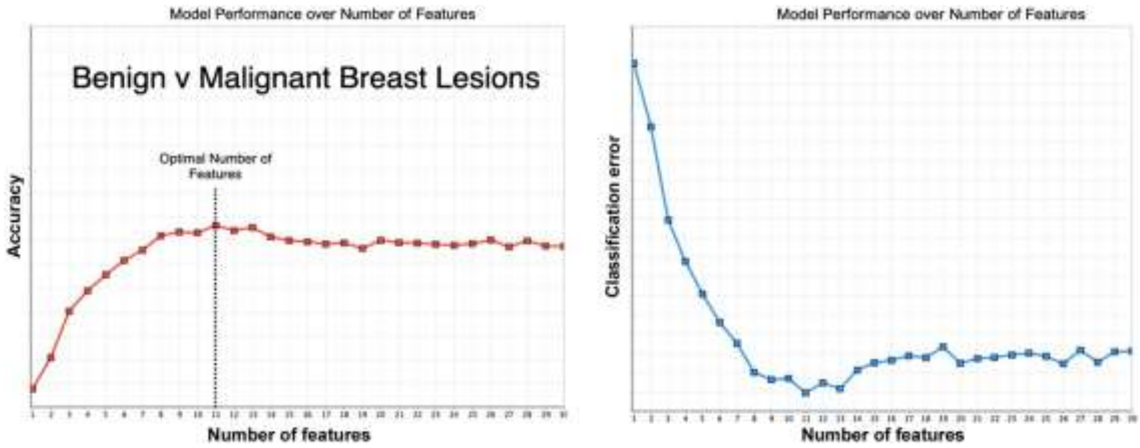


Figure 1: Radiomics feature selection. An example of a model classifier in action is shown here in its ability in order to differentiate between benign and cancerous breast tumours using imaging techniques. A recursive feature removal and reduction approach was used after a huge number of radiomic characteristics were calculated. Subsequently, variables with zero or near-zero variance, as well as those that were highly linked, were eliminated.. The model's performance, as seen above, identifies eleven characteristics that have reached saturation. The red curve on the left side shows the relationship between various characteristics and accuracy, as the blue line moves on the right side shows the relationship between the number of features and the error function of the model. By reducing the error function, this sample demonstrates great accuracy utilising eleven imaging characteristics.

On one extreme, you have supervised learning, where an algorithm is told what to do; on the other, you have unsupervised learning, where the programme learns from the data itself. These later ones are linked to more advanced CNN algorithms that can automatically find patterns in imaging data.

The interest of internet developers in automatically identifying things on photographic pictures and the massive dataset of ImageNet18 have evolved from the computer vision area as the driving forces behind CNNs.

Images have been the inspiration for several highly effective

ML architectures, such as Three versions of Inception, AlexNet, VGG-16, and 19. Several of these have found usage in medical applications via transfer learning3, which is adapting a pretrained architecture trained on ImageNet to medical imaging and making it work better for that specific task. The number of features obtained by convolutional neural networks (CNNs) often exceeds the quantity of data points or samples in ML-based cancer imaging.

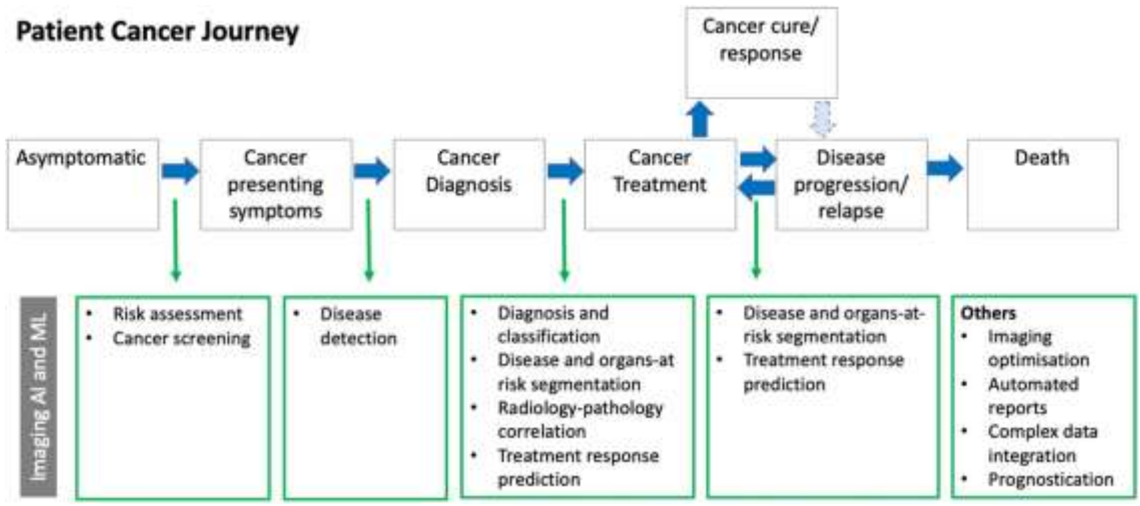


Figure 2 shows possible applications of AI and ML in cancer imaging as they pertain to a cancer patient's journey. The development of cancer presenting symptoms in an otherwise asymptomatic patient is a common occurrence that often results in a cancer diagnosis. Treatment for cancer, which may result in a successful response or even a cure, begins when the illness has been appropriately staged. But some people will progress during therapy or have a recurrence, necessitating more treatment. Some patients may sadly pass away as a result of their illness. Imaging AI and ML have many possible applications, as described and illustrated in the book, throughout the cancer journey.

The second causes overfitting, in which the model becomes well-suited to the training dataset at the expense of its performance on the dataset that was used for testing. To minimise or reduce the impact of overfitting, some popular methods include: (a) training the algorithm with more data, (b) using feature selection to reduce the initial features' dimensionality or number, and/or (d) employing a combination of learning techniques to expand the dataset, which means training the algorithm at multiple sites or institutions. Using methods such as k-fold cross-validation, which incorporates several sub-samples of the dataset, is another prevalent approach. Concerns about patient privacy and reluctance to share data persist despite the fact that more and more healthcare organisations are storing and using data on the cloud and other central locations. Due to these problems, methods of distributed or federated learning have gained popularity 19,20. Federated learning is an approach to machine learning that uses distributed learning rather than a central repository. Instead, models are sent to several institutions and trained using data specific to each site. The only data that is shared across these institutions is the model's weights. In order to make algorithms more trustworthy, there is currently a lot of focus on making them explainable and interpretable. While clinical users may not be too concerned with the nuts and bolts of ML models, they are curious in how the models arrive at their predictions and outputs, both for whole patient cohorts and for individual patients.

Clinical opportunities for AI/ML in cancer imaging. There are a number of applications for machine learning that might lead to significant improvements in cancer imaging. Figure 2 shows the usual course of a cancer patient's clinical care and draws attention to many important areas of imaging where AI systems might have a beneficial effect 22. We provide a more detailed overview of a few of them here.

Evaluation of risk: In order to make the most efficient use of cancer imaging technology, we must prioritise patients who pose the highest risk. In order to determine the likelihood of getting cancer, breast density testing is mandated in several US states. It is possible to make it such that individuals undergoing breast cancer screening get consistent density notifications by deep learning algorithms, which have shown excellent accuracy in breast density classification 23,24. Because of the importance of breast

density assessment is linked to large interobserver differences (6-85%) 25. Advancements in artificial intelligence (AI) have the potential to revolutionise risk modelling. When compared to typical breast cancer risk models on its own, a deep learning model that included both mammographic characteristics and traditional risk variables to identify high-risk individuals outperformed them 26,27. Using mammographic breast density as assessed to an AI programme, a junior radiologist, and an experienced radiologist, a very excellent agreement was reported for breast cancer risk evaluation not long ago (28).

Both cancer screening and cancer detection have attracted a lot of attention from the artificial intelligence community. Lung cancer (29–31 cases) and breast cancer (32–36 cases) are two examples of illnesses with active screening programmes that have evaluated AI systems. Research on breast cancer has revealed that AI algorithms can perform as well as human experts in reading mammograms, rank images in order of importance, and even be acceptable to women who undergo the screening process. But there isn't enough data from actual use cases to suggest widespread use of AI-based breast screening systems just yet. When directed screening tests are

neither feasible or cost-effective, opportunistic screening—which involves detecting anomalies in examinations performed for other purposes—may provide opportunities to discover more malignancies. This is in addition to systematic screening. Patients getting low-dose CT for lung cancer screening, for instance, may have their breast density evaluated on CT39 to determine their risk of breast cancer using the same pictures.

Artificial intelligence algorithms can now identify lung nodules, classify them, measure them, and even forecast whether they are cancerous. Using a deep learning model for lung nodule diagnosis and management enhanced radiologists' performance and decreased reading time 40. Without a doubt, further tumour types will be added to the list of possible applications for AI in cancer diagnosis.

Machine learning (ML) systems provide ways to improve the classification of imaging results in the field of cancer detection and classification. Malignant brain tumours may arise from a variety of sources and have varying outcomes; however, owing to the heterogeneity of the illness, tissue collection can be invasive and results in inaccurate characterization. Research has shown that AI has the ability to detect and categorise significant brain tumours, such as acoustic neuromas, pituitary adenoma, meningiomas, cerebral metastases, low grade gliomas, and acoustic neuromas, while also distinguishing them from healthy tissues 41–44. Another new use for this field is the categorization of cystic pancreatic lesions, which is useful because visually differentiating between intraductal papillary mucinous neoplasms, mucinous cystic neoplasms, and serous cystic neoplasms can be difficult (45–47) and has different consequences.

Forecasting how a patient will react to a therapy is possible with the integration of machine learning and radiomics. Included in this category are tasks such as anticipating how nasopharyngeal carcinoma will react to intensity-modulated radiation therapy (IMRT) 48, the response of non-small cell lung cancer to neoadjuvant chemotherapy (NCLC)49, and how rectal, oesophageal, and breast cancers will react to neoadjuvant treatment (NCT) 55,56. Despite its great potential, radiomics has yet to provide findings that can be applied to a broader population, which is restricting its use in clinical practice at the moment.

For the sake of patient care, quality improvement, and education, it is crucial to correlate radiological data with information found in pathology reports. To further investigate cohort-specific populations, it is feasible to use natural language processing methods to sift text-based radiology 57 and pathology 58 reports for important discoveries. Radiologists may be alerted to possible study follow-up misses using a radiology follow-up tracking engine 59 that applies a natural language processing technique to organ-level categorization of free-text pathology reports. Opportunities to integrate radiological pictures with anatomical pathological images also exist 60,61.

Segmentation of diseases: Segmentation, or the outlining of diseases, is essential to many research in artificial intelligence, machine learning, and radiomics essential for producing tumour outlines for radiation planning 62–64 and for deriving quantitative tumour data, such as tumour diameters. Clinicians may learn more about cancer response to therapy by registration of segmentations throughout time-series. Automatic disease segmentation utilising AI models has the potential to decrease the substantial inter-reader variability 65 that may result from manually tracing lesion boundaries. An expert radiologist should verify the final AI segmentation result, even at the point when DNNs are powerful enough to separate lesions.

Deep learning techniques have shown to be very effective when there is an abundance of data, which likely explains why some picture and illness categories have pretty well-developed segmentation algorithms. The data density is significantly higher for segmentation problems compared to classification problems typically considered at the per-patient level, such as in radiomics. This is because lesions or entire organs consist of hundreds or thousands of voxels, the smallest picture element, defined by the spatial resolution of the acquisition and the thickness of the image section. Disease segmentation allows for the computation of radiomic features from the entire tumour. However, a more sophisticated approach involves extracting radiomic features from habitats, which are areas within tumours that are physiologically distinct, as inferred by imaging characteristics (e.g., based on blood flow, cell density, necrosis) 66,67.

Potentially vulnerable organ segmentation: The goal of radiation treatment is to kill tumours as much as possible while avoiding healthy tissue. On the other hand, normal tissues and organs are commonly found near malignancies, making them organs-at-risk to the potentially harmful scattering effects of radiation treatment. To ensure that radiation does not harm nearby normal tissues, organs-at-risk segmentation is an essential part of the treatment process. When treating pelvic cancers 68,69, for instance, All of the usual organs that might be affected include the bladder, bowels, rectum, and hip joints. In addition to head and neck cancers 70,71, breast cancers 72, and non-small cell lung cancers 73,74, ML has also been successfully employed in organs-at-risk segmentation for radiation planning Among the many expanding uses of AI and ML in the imaging field, image optimisation is on the rise, and it's not only in cancer imaging. An oncological body scan using magnetic resonance imaging (MRI) might take anywhere from thirty to sixty minutes. Both to enhance picture quality (e.g., to create so-called super-resolution MRI images76) and to speed up image capture and/or reconstruction (i.e., to make the examination faster)75 are seeing increased use of AI and ML approaches. Reducing the amount of time it takes to get an MRI without compromising picture quality will help health organisations overcome MRI capacity challenges and increase patient throughput.

Additional applications: radiologists are looking at natural language processing to help them with repetitious work and to create automated reports. Another possible application of natural language processing as a communication tool for doctors receiving radiology reports is to alert them to actionable reports. This way, significant results may be emphasised to referrers in a timely manner.

How well AI and ML perform in the various use cases we've covered so far depends on factors like project complexity, data availability and quality, mathematical model sophistication, and algorithmic real-world testing. There is ongoing investigation and development in some of these use cases. Still, not all algorithms under development or testing end up being useful therapeutic tools. To allow the use of AI and ML in cancer imaging, it is crucial to identify the obstacles and limitations that need to be overcome.

Obstacles to using AI and ML in cancer MRI. The field of cancer imaging presents both promising potential and formidable obstacles for the advancement of artificial intelligence and machine learning. The following are some of the major clinical, professional, and technological obstacles that will be faced when beneficial mathematical algorithms are translated into more widespread clinical practice for the benefit of patients.

Clinical challenges. Developing an AI or ML technology that solves a critical clinical problem or answer a critical clinical issue is a top priority. Developers should thus be well-versed in both the clinical setting and the expected implementation environment of the AI tool. For this reason, it is common practice to consult with doctors when creating the tool.

Data comes in from many directions in the clinical realm. Clinicians are producing more and more biomedical data as a result of developments in areas such as electronic health records, advances in multi-omics technology (including genomics, proteomics, and molecular pathology), and multi-modal imaging. Therefore, successfully involving several disciplines is essential. Artificial intelligence and machine learning have the ability to combine this varied and complicated data in order to bolster customised medicine 79. On the other hand, data-driven and model-based computational approaches face additional obstacles when dealing with datasets of this size.

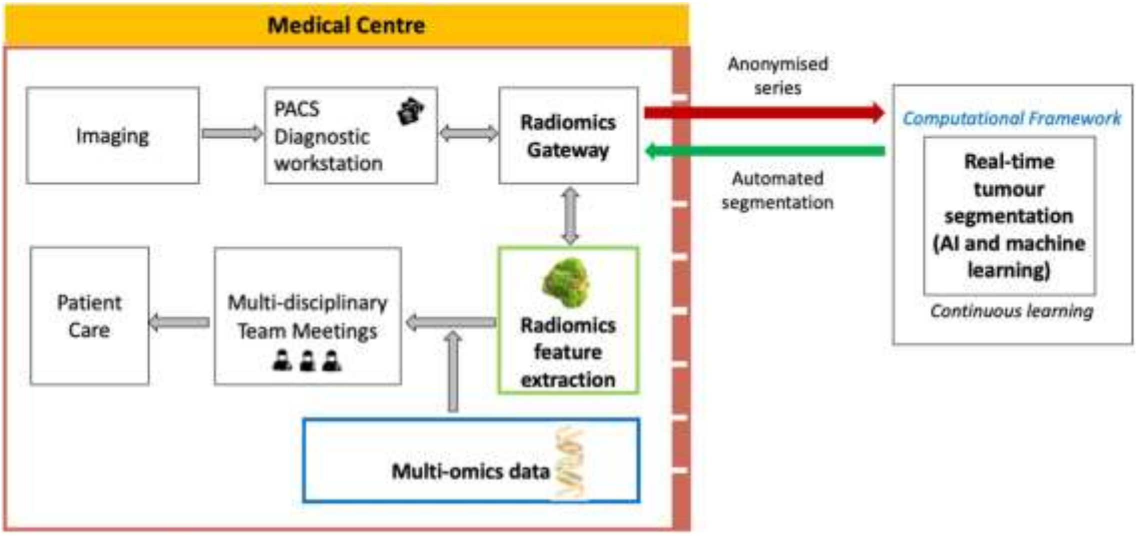
By using advanced ML and computational intelligence, AI has the ability to completely transform the field of cancer image analysis. Advanced AI techniques have the potential to revolutionise healthcare by making it more patient-centered rather than organization-centric. This might lead to more personalised solutions, lower healthcare costs, and improved clinical results. Another benefit of computerised oncological image analysis is the improvement of decision support tools and the acceleration of the shift from qualitative to quantitative image interpretation and assessment via the use of automated methods for earlier detection and enhanced lesion characterization. There are significant problems within this paradigm that can be solved with improved AI and ML. Sufficiently effective prognostic and predictive biomarkers, computer-assisted diagnosis that is both accurate and repeatable, and reliable tumour segmentation are all necessities in this field. Measuring and tracking intra- and inter-tumoural heterogeneity as the illness progresses is going to be a real challenge 82,83. To do this, we need longitudinal imaging datasets of excellent quality.

Precision oncology—the practice of tailoring cancer treatment to each individual patient by analysing their tumor's genetic profile—is one field that may benefit greatly from the use of AI and ML. With the help of integrated diagnostics 84,85—for example, radiogenomics—which integrates studies of radiomics and genomics—precision oncology has a better chance of developing powerful computational tools for studying cancer biology and forecasting treatment response (Fig. 5). To address concerns about privacy and cyber-security and to facilitate ongoing learning, the approach comprises collecting structured data on a massive scale from a variety of sources. The current biggest obstacle is getting new AI technologies into clinical practice without first conducting rigorous, well-validated clinical trials. Important for the development and implementation of AI methods in precision oncology86, this holds great promise for the future of AI in oncology as a tool for more precise patient selection and, ultimately, lower treatment costs.

Difficulties in the workplace. Professional hurdles, in addition to therapeutic ones, will certainly influence how ML is developed and used in cancer imaging. Factors that encourage the growth of ML include the ever-increasing need for imaging services, which, when combined with both short-term and long-term shortages in the workforce, may cause radiologist burnout and stress. Departments should think about modernising or reworking their IT system and process so they can test and use ML and AI technologies as they become available. Additionally, there is the matter of how the radiology staff views the possible benefits and risks of using AI and ML in healthcare settings.

An online survey was carried out with 569 radiologists from 35 different countries in order to gather information for a 2019 is organising a conference in collaboration with the International Cancer Imaging Society and the Champalimaud Foundation (Lisbon) on the use of AI and ML in cancer imaging. More over 60% thought the advantages of AI were greater than the disadvantages (Supplementary Note). The majority of people who took the survey saw AI in radiology as having mostly positive effects, such as (1) notifying radiologists of abnormal findings, (2) making work more efficient, (3) suggesting diagnoses when the radiologist is unclear, (4) acknowledging that the radiologist should be held accountable for mistakes, and (5) modifying the service model through increased patient-provider communication. The majority of responders were of the opinion that radiologist jobs would not be filled by AI and ML. More than 70% of respondents said that we should invest in education, test new tools, support large-scale picture and annotation curation, and collaborate with commercial vendors to build AI solutions that enhance workflow in order to be ready for the coming of AI.

Among the most pressing needs and topics for future AI tool development, the study highlighted the following: (1) the development of systems that can automatically monitor tumours over time to gauge treatment efficacy; (2) the enhancement of fully or partially automated



In the future, precision oncology may be able to integrate data from many omics sources to monitor tumour volume, geographic and temporal phenotypic variability (Fig. 5). When each hospital or clinic obtains and saves its own medical imaging data (in a local PACS), this architecture would make it possible to handle data from more than one institution. In order to conduct quantitative studies, a gateway for radiomics is used to interface with an external, reputable AI/ML centre that enables continuous learning. This centre is then asked to automate the segmentation of tumours in real-time. In order to address concerns about privacy and cyber-security, the medical photographs that leave the hospital are anonymized. Radiomic feature extraction and analysis are performed using the segmentation findings, which serve as virtual biopsies. In order to find correlations with clinical and multi-omics data, statistical imaging findings are combined among other sources of biological data. As precision oncology develops, this method has the potential to enhance cancer treatment in clinical practice by creating trustworthy diagnostic and prognostic tools for interdisciplinary team meetings.

A system tools for tumour segmentation; (3) tools for proforma reporting that enable prospective annotation of image data; (4) tools for concurrent identification of normal studies that allow radiologists to concentrate on abnormal examinations; and (5) tools for tumour identification system-wide.

Further, in order to provide future AI-enabled practice, imaging departments must prepare for the workforce requirements. A deeper familiarity with artificial intelligence (AI), particularly its applications in workflow management and image collection, will be necessary for radiographers and technicians. Building a platform for in-line development or testing of AI tools, a place to interact with and annotate data on images, as well as carefully selected images and data archives are all essential tasks for an informatics team.

Difficulties related to technology. Numerous cutting-edge deep learning-based AI strategies are accomplishing remarkable results. Their success may be attributed to two factors: the accessibility of massively indexed datasets annotated with precision and the robust self-learning capabilities of deep ML models. The need for subject specialists' knowledge88 makes the task of obtaining such correct annotations costly and time-consuming in biomedical research. Consequently, there has been a lot of interest in ML models that can handle imperfect annotations and lack of strong supervision. This includes things like image-level labels instead of feature-specific labels, or bounding boxes that cover an area of interest instead of exact outlining. One possible solution to the problems of data scarcity and heterogeneity is the creation of massive image databases that can be mined. Data quality and variety, in addition to sample availability, should be taken into account while gathering and creating standardised datasets. By employing transfer learning and domain adaption strategies, it may be possible to enhance the capacity to generalise across investigational settings.

It is difficult to design and find trustworthy AI imaging research. The findings of these AI models are very doubtful because of the small sample sizes (as low as 10 patients) used in these studies. This is because of the risk of overfitting, which reduces the generalizability of the results. As a general rule, while dealing with binominal classification jobs in radiomics, it is recommended to recruit 10-15 patients for each characteristic that is part of the final radiomics signature 90. The so-called test set, a dataset that does not include any instances used in training or adjusting the model, should be used to estimate performance. In addition to internal validation, external validation, which involves testing how well the model performed on many datasets collected from diverse imaging modalities or geographic patient populations, is necessary to assess the model's generalizability. Validating models on an external patient cohort that is 25-40% larger than the training sample is considered ideal.

When it comes to more complicated clinical concerns like illness risk assessment and prediction, integrative models that include data from environmental, social, and genetic sources are becoming more popular. There is a clear need for more standardisation, particularly in data gathering, to enable various use cases of AI 91, as data sparsity and non-standardized treatment techniques across institutions continue to be obstacles to constructing integrative ML models.

One way to get real-world data for evidence-generating research is to combine photographs with clinical and molecular data. There are great potential to test and assess the performance of AI technologies using retrospective data from imaging biobanks and repositories. If the degree of variability is too high and might compromise the results of a multicenter study, harmonisation techniques like ComBat 92 could be explored to standardise the imaging characteristics.

Box 1 | Data cleansing and analysis are critical steps in making AI and ML more reliable and applicable to cancer imaging.

1. Participant recruitment criteria

Consistency in the inclusion of the study population based on the presenting symptoms, results from previous tests, defining the appropriate index tests or by the selected reference standard

2. Participant sampling

To avoid or control bias in participant sampling, considerations could include the use of consecutive series of participants, use of well-defined selected data silos, clear and well-defined selection criteria; as well as adjusting for possible confounding variables

3. Data collection

What data to collect and how this is performed should be planned before participant recruitment and sampling. Where appropriate, target trial emulation may be undertaken, which is the application of design principles from randomized trials to the analysis of observational data, which may improve the quality of the observations.

4. Reference standard

The rationale and description of the reference standard should be clear

5. Technical specifications of materials and methods

Aspects of technical specifications should be well defined. These include how and when images and measurements were taken; the definition of units; cut-off thresholds; defined results categories (of both the index tests and the reference standard); description of the number, training, and expertise of persons executing and reading (original or new reporting); index tests and the reference standard; and blindness aspects (if the readers of the index tests and the reference standard were masked to other test results)

performance and applicability of the model. Radiologists may take the lead in the area by advocating observational in silico investigations and making sure that all important steps are followed, from data collection to analysis, to ensure that the findings can be reproduced. According to Box 1 93, the most important factors are as follows.

Radiomic computation and the use of radiomic features for prognostication, assessment of therapy response, and diagnosis of molecular phenotype are not without their difficulties in the context of radiomics. These difficulties stem, in part, from the fact that Image capture and reconstruction methods (94–98) and user and software-specific segmentation variations (99,100) greatly impact the radiomic feature values. If radiomics is going to reach its maximum potential, it is necessary to solve these concerns, enhance algorithms, and gain community consensus on the usage of open-source software, phantoms, and standardised approaches101.

Increasing model performance—perhaps at the cost of explain ability—has been the primary goal of AI enthusiasts, which may explain why there has been little progress in translating AI models to clinical use. An example that comes to mind deep neural networks use a black-box method. While these networks provide impressive results, it may be difficult to verify their reliability, which hinders their clinical acceptance. Prioritising AI solutions with substantial therapeutic value may potentially be hindered by a lack of multi-disciplinary interaction. Assuming AI models do not reveal how they arrive at a particular choice, There can be reluctance among healthcare professionals to include AI into routine clinical procedures.

The AI community has recently come to terms with this restriction and is focusing on creating explainable AI. Concerns about patient safety are delicately touched by the explain ability of AI models, particularly in clinical decision-support systems 102. Machine learning models are prone to patient selection bias, which causes them to underperform and make incorrect predictions in future unknown situations, due to the fact that most AI models are trained using historical, observational data. Consequently, domain experts should consistently double-check AI models' predictions and the logic behind them. Only when the models are intentionally somewhat open can the second goal be achieved. Strengthening AI models is expected to be achieved by including domain experts into model development. And repeatable and contribute to establishing credibility among consumers. It is essential to assess the AI solution's total performance in the therapeutic route environment, not just its accuracy. As part of this process, we would put these models through their paces in the real world to see how well they work, how reliable they are, and how much they cost.

What we've learned from the past: using CAD to diagnose breast cancer. While machine learning and artificial intelligence hold great promise for the imaging field, it is important to remember the mistakes made in the past when trying to use computational methods to cancer imaging (e.g., computer-assisted breast cancer detection). Beginning in the mid-1980s, researchers began to develop algorithms that could automatically identify masses and calcifications on mammograms. In 1998, the Food and Drug Administration approved the first commercial computer-aided detection (CAD) system for mammography, which was originally based on digital film. The use of CAD in clinical practice was made possible by the shift to digital mammography. Approximately 74% of mammography interpretations in the US used CAD by 2010106, thanks to the encouraging early findings from clinical studies and the implementation of Medicare funding coverage for CAD in the US. But even while commercial CAD systems' stand-alone sensitivity in enriched reader experiments is typically better than radiologists', The diagnosis accuracy of screening mammography did not significantly increase with the deployment of CAD, according to major retrospective registry-based studies 106,107, and 108.. Present commercial CAD systems provide an average of 1-2 false positive prompts every instance, which is likely to account for this disheartening outcome. Because of this false-positive prompt rate, the positive predictive value of a CAD prompt is less than 1% in an environment with low prevalence screening. There will be a propensity for radiologists to disregard the computer-generated prompts entirely, as they will have to disregard over 99% of them in order to locate the one that indicates malignancy. Traditional feature-based CAD systems have certain limits, but modern deep learning algorithms may be able to overcome some of those restrictions. Because deep learning algorithms don't need to replicate the radiologist's reading style, unsupervised training on massive datasets including millions of mammographic pictures may be able to compensate for human observers' limitations.

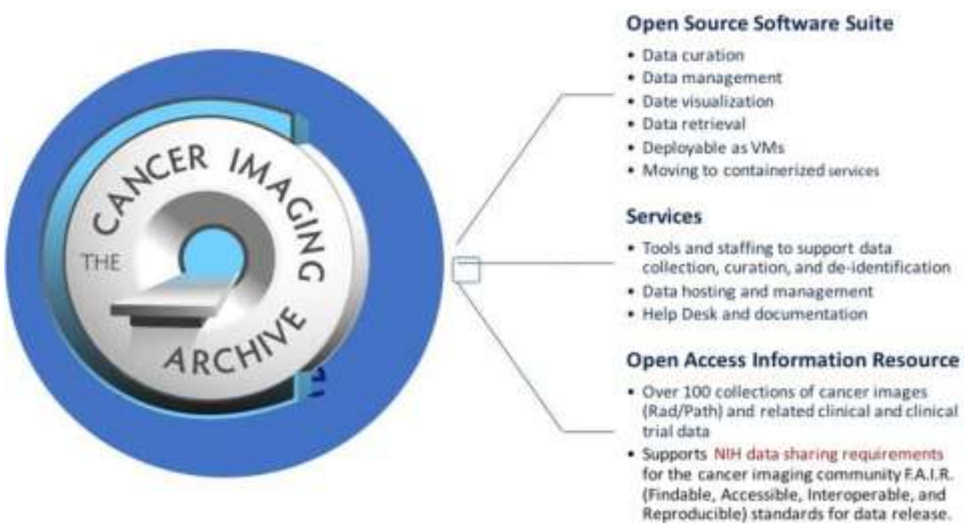


Fig. 6 Built entirely on open-source software, the Cancer Imaging Archive (TCIA) is a network of networks. Additionally, TCIA is a collection of services that aim to gather and organise high-quality clinical data and images relevant to cancer, and then make it accessible to the public. ("VMs" stands for "virtual machines").

109 mammograms. On the other hand, algorithms110 will have more work to do as detecting tasks become more automated; they may even have to establish that there is a positive impact on patient outcomes beyond just diagnosis accuracy.

Therefore, based on the important takeaways from past CAD applications in breast cancer, the next AI tool for cancer detection needs to have a high positive predictive value, meaning it will produce less cases of false positives in areas where the illness is not common, and high diagnostic accuracy in general. In order to determine the health benefits and broader advantages of these technologies, it is necessary to evaluate them in real-world settings beyond only their diagnostic performance.

Advancements in technology, infrastructure, and expertise are necessary for the use of artificial intelligence and machine learning in cancer imaging. Archive and imaging resources: In order to train and validate, supervised learning approaches need massive amounts of labelled data 103. Artificial intelligence (AI) models in cancer imaging might make use of a multitude of data sources. Such as biobanks of medical pictures in digital formats; imaging biomarkers being defined as endpoint surrogates; or even population studies 111′.By establishing probabilities of development of disease, prediction of disease, early disease diagnosis and phenotyping, grading and staging of disease, targeting therapies, monitoring treatment efficacy, predicting adverse events, imaging biobanks allow in silico evaluation and validation of the new biomarkers.

One way to collect and share enough high-quality, well selected data is via open access data repositories. Cancer image repositories that are open access are few. There is a lack of consensus on the topic of data sharing. In addition, there is a lack of uniformity among countries when it comes to data privacy, informed consent requirements, and the growing interest in the potential monetary worth of patient data, and might block data sharing. Researchers and institutions often restrict or outright ban access to key data sets because they see them as intellectual property. Data used to verify algorithms authorised for commercial usage is argued for by regulatory authorities (such as the FDA) to be kept secret114.

The Cancer Imaging Archive (TCIA), the biggest open access library of cancer images (Fig. 6), was established and is still run by the US National Cancer Institute (115,116).

One goal of TCIA is to make more high-quality cancer imaging data sets available to the public so that researchers may use them. Compliance with the F.A.I.R. (Findable, Accessible, Interoperable, and Reproducible) criteria for data release has ensured that the data is available to the public 117,118. Data warehouses are being built by other research-funded organisations within the EU and beyond. While dataset quantity is important, data variability and quality are equally crucial. The data collected should be of high enough quality and collected using consistent parameters. The results of clinical trials are more reliable because of the stringent quality controls and consistent methods used to gather the data. The primary goal of TCIA is to compile, organise, and disseminate information derived from finished clinical studies. Here, curation means making sure that data formats, metadata, and anatomical coverage are all uniform and meet international data standards; it also means making sure that any personally identifiable information about patients is anonymized.

Training ML algorithms on data that accurately depicts population variation, illness presentation, and data collecting methods is essential for their clinical use 119,120. The limited quantity and high expense of datasets for training and testing that are of high quality are caused by the manual creation of labelled data by human professionals. When working on a machine learning project, annotating the data and preparing it for future analysis and modelling might be the most time-consuming task. To improve efficiency, crowd sourcing is being tested for image annotation, which is often a bottleneck for AI and ML. At the patient level, annotations may be supplied at overall survival or disease-free survival; at the picture level, they can be provided at benign or malignant; and at the voxel level, they can be provided at lesion or non-lesion. When training automated segmentation models, radiologists manually define lesions in many image slices 121, but lesion detection algorithms often need annotations of a bounding box type, typically enclosing the lesion.

It will be important to think about ways to handle data heterogeneity when large amounts of imaging data from many locations and scanners are merged into archives. Using deep learning techniques to learn from such heterogeneous data could be a solution; this might lead to more consistent and repeatable results. In order to draw repeatable causality123 conclusions from virtual patient cohorts, studies that look back at past events using help resolve issues that are pertinent to healthcare and public policy. Using a cross-validation strategy to build and test AI models is acceptable, especially when the condition being studied is uncommon and the datasets are limited.

Concerns about the lack of transparency and explainability around the use of artificial intelligence (AI) in cancer imaging have been addressed via the use of open-source software (OSS) strategies and collaborative efforts. Open source software (OSS) refers to computer programmes whose source code is accessible to the public via a legally binding licence that grants certain rights to anyone who obtain such rights, including the ability to modify, enhance, and redistribute the software for free. Depending on the preferences of the copyright holder, there are several kinds of open source software licences available today [https:// opensource.org/approval]. In America, you may find a wide variety of licences, from the more liberal Apache 2.0 licence to the more restrictive General Public Licence (GPL). A variety of commercial contracts allow for the transformation of OSS from its non-commercial form into products that include extra services like training, documentation, warranty, and maintenance.

I open-source software (OSS), (ii) a system of governance, and (iii) a network of collaborators are the three interdependent parts of an effective open-source ecosystem. More than fifty machine learning (ML) packages are available as open source software (OSS) at this time, and they run on a wide variety of systems and languages. The TensorFlow, Keras, PyTorch, and Caffe2 packages are among the most widely used ones. The pros and cons of each one rely on the specific requirements of the target audience.

While the developers and sponsors of these open source software packages often have their own use cases and applications in mind when creating and supporting these packages, medical imaging research may benefit greatly from using these packages as a starting point. To be more effective in medical settings, however, they will require optimisation. Pattern recognition in consumer apps, for instance, is often reliant on visual characteristics and picture orientation. Diseases in medical pictures are often subtle and show up as small variations in grey value rather than graphical elements; yet, medical image patterns are often orientation-independent. These considerations highlight the need to re-train and fine-tune algorithms accessible via OSS packages using medical imaging data in order to achieve optimal performances. In conclusion, open source software (OSS) offers a realistic way for the AI community to engage, build, and test new AI tools while simultaneously addressing some of the privacy and transparency issues.

There are substantial perceived benefits of using AI solutions in healthcare124 across the whole clinical workflow, which has implications for both regulatory systems and healthcare itself. The diagnostic process for patients will be enhanced in radiology as a result, beginning with the appropriateness of imaging requests (125) and ending with the follow-up of actionable findings in radiological reports (126). Due to the persistence of substantial implementation hurdles, many improvements have not yet reached their full potential.

Software is now subject to stricter regulation as a medical device (SaMD) due to the new EU Medical Device Regulations, which went into effect in 2021. The software is certified based on how it is utilised and used in the clinical workflow. It is important to note that most AI software in imaging is now certified as a decision-support tool and should not be utilised independently for clinical or patient care. It is important to think about whether radiologists will utilise the software during primary reporting or as a second read after the original main report is provided. Just recently, an ad Throughout the healthcare industry, there are a plethora of software products that have received regulatory clearance but have yet to be used. Through ongoing training, AI products have the potential to keep improving even after their original release. Despite having CE markings or FDA approval, many items have entered the market without first undergoing independent testing. To that end, 127 a new framework for the FDA to follow is being considered to guarantee the security and efficacy of AI solutions. Premarket submissions are now required to adhere to the FDA's specified change control strategy. Both the mechanism utilised to perform these controlled changes (the algorithm change protocol) and the expected modifications (SaMD pre-speci¬fications) are part of this strategy. Manufacturers are expected to commit to transparency and real-world performance monitoring. They are also expected to report the FDA on any modifications made as part of the authorised pre-specifications and the algorithm change process.

Initiation of a Data Protection Impact Assessment is often required at the local level after software or hardware validation as a certified medical device. This is to ensure data privacy, and in Europe, this entails adhering to the General Data Protection Regulations (GDPR). Also, when considering potential IT architectures for implementation, it's wise to do a Solution Architecture Review. Since the GDPR might be interpreted differently in different countries, local regulations must also be followed when it comes to the use and storage of patient data. The need for a logical and consistent digital infrastructure and worries about patient privacy have been called "the uncomfortable truth" in medical AI 128.

The AI company's product design, the healthcare system's diversity and size, and knowledge of data transfers between healthcare providers and software processors all play a role in the software integration process with existing hospital IT infrastructure. Failure of software integration is a known barrier for adoption. In contrast to well-established companies with solid products, hospitals typically have longer integration timelines due to the complexity of their radiology informatic systems (such as PACS and VNA, which communicate with HIS and RIS) and the wide variety of data inputs (such as the non-standardized naming of imaging sequences from different scanners).

If we want AI workflows to work, we need to make sure that all imaging procedures follow the same acquisition protocol, regardless of the scanner or vendor. We also need to make sure that all radiological reports follow the same structure and use the same lexicon so that data mining can be easier. The American College of Radiology has suggested some structured reporting templates on RadReports.org. In the absence of these prerequisites, software integration may have to be planned for each modality separately, which might lead to intricate data programme errors may manifest as a result of such input data heterogeneity, depending on the level of algorithm maturity.

It is possible to approach the introduction of a new AI tool into a healthcare system with caution at first by coordinating a trial period with the provider. A "try before you buy" strategy would let customers evaluate the AI tool's reliability, usefulness, and integration with their workflow before making a purchase. Reason being, doctors won't trust the tool until it shows a high \ level of accuracy. Integrating a radiologist feedback feature into the PACS interface is one possible option. By utilising checkboxes labelled "agree," "AI overestimation," "AI underestimation," or "both over and underestimation," the radiologist may assign each AI algorithm a grade based on how well it performed. Because of this, it would be possible

Box 2 | Important factors for the selection of AI into a health system

Criteria and benchmarks

CE labelling FDA clearance UKCA marking

Incentives and motivations

Targeting a common disease

Potential for the AI algorithm to be developed into products that generate revenue Attracting better or new payors

Formulation of fair value proposition between stakeholders or partners Latitude to create/share own business model

AI tool Infrastructure fits with existing informatic systems The AI tool can be assimilated into the clinical workflow

people to report what they see as inconsistencies so that they might be investigated further. Also, be wary of vendor-developed solutions that force users to use their algorithms regardless of whether or not such methods are suitable for your individual needs. Education on the software tool's use is also necessary for the community of professionals that engage with it. While training a small set of AI users would be manageable, dealing with a big medical system's prospective users is a significant challenge.

Using approved medical devices to handle patient data in ordinary clinical practice does not need further permission. Nevertheless, manufacturers must acquire specific data permission from patients in advance if they want to use patient input to enhance their software programme. Since post-hoc sharing of such data may be prohibited, procedures should be established to identify patients who have granted consent and to withdraw it as needed.

When all the obstacles to AI deployment are removed, one question may remain unanswered: who will pay for the AI? No healthcare systems or private health insurers have covered AI use as of yet, even though AI development and testing may be supported via research grants or commercial partnerships with firms. Finding dedicated funds to assist the use of AI is difficult in the context of falling radiological procedure fees, since AI may be expensive to implement across healthcare systems. There is currently no documented proof of AI improving work efficiency, despite the fact that AI shows great potential in this area. An opportunity to establish tariff models for the usage of AI may arise with the creation of specific patient-centric services that employ AI. An initiative in the United Kingdom is testing a bone health service that will pay to screen for those who are likely to suffer from osteoporotic spinal fractures. The business case was built around the entire service, not just one AI product. The service's goal was to find patients who were at risk of osteoporotic fractures, so they could be treated early and maybe reduce healthcare costs in the long run by a smaller number of fractures. The integration of AI with value-based healthcare is shown here.

Accumulating local data to support the implementation of AI is crucial in less cohesive healthcare systems that include imaging services as part of their provider network, meaning they deliver specialised services. Metrics that demonstrate improvements in reporting accuracy, such as a decrease in patient recall rates during mammography, faster reporting times, and increased revenues, are just a few examples. It is feasible to evaluate new artificial intelligence solutions in diverse healthcare markets and with different combinations of payer models. This might lead to the ultimate widespread use of AI software tools in healthcare systems (Box 2).

The future of radiology: Radiologists need training in artificial intelligence (AI) concepts and algorithm validation so they can assess the usefulness of AI tools before incorporating them into clinical practice.

Some forces are pushing for the hasty incorporation of AI technologies into healthcare workflows. One major issue is the scarcity of qualified radiologists, which is affecting many nations' workforces. Roughly 10% of radiology positions in the UK go vacant 129. Second, the demand for and effort associated with imaging has been steadily rising over the world. There has been an annual increase of around 10% in the CT and MRI workload in the UK. Thirdly, there is an incessant push to enhance workflow efficiency by decreasing the time it takes to process images without lowering the quality of the diagnostics. Last but not least, AI is considered as a tool to assist radiologists with tedious and boring repeated duties, such as sequential tumour size measuring or cancer screening.

Conclusion

Radiologists would be able to evaluate the efficacy of AI systems if medical schools and radiology programmes included courses on algorithm design, training, testing, and validation; fundamental statistics relevant to AI and ML; data requirements and their associated challenges; and the language and primary methodology of AI and ML. In order to empower radiologists, it is necessary to teach them how to evaluate AI algorithms in a way that is both useful and rigorous in their clinical practice.

Radiologists, computer scientists, data scientists, trainee radiologists, other doctors, radiographers, medical students, and data engineers are all important players in the radiology industry, and their education will determine the future of application of AI and ML in the domain. In order to meet the unmet requirements in cancer, it is crucial to build AI tools that are both technically competent and therapeutically useful. This can only be achieved via a multidisciplinary discussion. To facilitate communication and collaboration amongst all parties involved, both locally and globally, there need to be an increase in AI-focused gatherings and conferences that draw from a variety of disciplines.

The majority of radiologists are pleased with the fast advancements in artificial intelligence (AI), especially machine learning (ML), in the field of cancer imaging, which has many positive therapeutic applications. When working on new ML techniques, the availability of imaging data is often a limiting factor;

Having said that, biobanks and open access provide the possibility of creating and use carefully selected, real-world image data. databases to circumvent these constraints. Potentially improved center-to-center communication and cooperation could result from using open-source tools for algorithm development wherever feasible. While these AI software algorithms may improve diagnostic performance, it is unclear which ones will last and which ones will be cost-effective in the long run. There is a growing consensus on the need for a more robust regulatory framework to authorise the use of AI-powered clinical tools. Because of the lack of thorough testing that these programmes often get before release, a systematic review is necessary. It is also crucial to provide all parties involved, particularly radiologists, with enough knowledge of this emerging field so that they can evaluate these technologies critically before incorporating them into their own work. The creation of practical clinical tools with the goal of improving patient care and outcomes may be facilitated by opening up opportunities for multidisciplinary cooperation..

References

[1]. Erickson, B. J., Korfiatis, P., Akkus, Z. & Kline, T. L. Machine learning for medical imaging. Radiographics 37, 505–515 (2017).

[2]. Napel, S. In Radiomics and Radiogenomics: Technical Basis and Clinical Applications (eds Napel, S. & Rubin, D. L.) 3–12 (CRC Press, 2019).

[3]. Trivizakis, E. et al. Artificial intelligence radiogenomics for advancing precision and effectiveness in oncologic care (Review). Int. J. Oncol. 57, 43–53 (2020).

[4]. Lo Gullo, R., Daimiel, I., Morris, E. A. & Pinker, K. Combining molecular and imaging

metrics in cancer: radiogenomics. Insights Imaging 11, 1 (2020).
[5]. Rubin, D. L., Ugur Akdogan, M., Altindag, C. & Alkim, E. ePAD: an image annotation and analysis platform for quantitative imaging. Tomography 5, 170–183 (2019).
[6]. Kalpathy-Cramer, J. et al. A comparison of lung nodule segmentation algorithms: methods and results from a multi-institutional study. J. Digit. Imaging 29, 476–487 (2016).
[7]. Echegaray, S., Bakr, S., Rubin, D. L. & Napel, S. Quantitative image feature engine (QIFE): an open-source, modular engine for 3D quantitative feature extraction from volumetric medical images. J. Digit. Imaging 31, 403–414 (2018).
[8]. van Griethuysen, J. J. M. et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 77, e104–e107 (2017).
[9]. Zhang, L. et al. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. Med. Phys. 42, 1341–1353 (2015).
[10]. Aerts, H. J. et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 5, 4006 (2014).
[11]. Gevaert, O. et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data–methods and preliminary results. Radiology 264, 387–396 (2012). One hundred fourteen of 180 CT image features and the PET standardized uptake value were predicted in terms of metagenes with an accuracy of 65%-86%.
[12]. Mattonen, S. A. et al. Quantitative imaging feature pipeline: a web-based tool for utilizing, sharing, and building image-processing pipelines. J. Med. Imaging 7, 042803 (2020).
[13]. Di Tommaso, P. et al. The impact of Docker containers on the performance of genomic pipelines. PeerJ. 3, e1273 (2015).
[14]. Dankers, F., Traverso, A., Wee, L. & van Kuijk, S. M. J. In Fundamentals of Clinical Data Science (eds Kubben, P.,Dumontier, M. & Dekker, A.) 101–120 (2019).
[15]. Traverso, A., Dankers, F., Osong, B., Wee, L. & van Kuijk, S. M. J. In Fundamentals of Clinical Data Science (eds Kubben, P., Dumontier, M. & Dekker, A.) 121–133 (2019).
[16]. Parmar, C. et al. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. Front. Oncol. 5, 272 (2015).
[17]. Ather, S., Kadir, T. & Gleeson, F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications. Clin. Radiol. 75, 13–19 (2020).
[18]. Deng, J. D. et al. In IEEE Conference on Computer Vision and Pattern Recognition. 248–255 (2009).
[19]. Rieke, N. et al. The future of digital health with federated learning. NPJ Digit. Med. 3, 119 (2020).
[20]. Kirienko, M. et al. Distributed learning: a reliable privacy-preserving strategy to change multicenter collaborations using AI. Eur. J. Nucl. Med. Mol. Imaging https://doi.org/10.1007/s00259-021-05339-7 (2021).
[21]. Kitamura, F. C. & Marques, O. Trustworthiness of artificial intelligence models in radiology and the role of explainability. J. Am. Coll. Radiol. 18, 1160–1162 (2021).
[22]. Bi, W. L. et al. Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J. Clin. 69, 127–157 (2019).
[23]. Mohamed, A. A. et al. A deep learning method for classifying mammographic breast density categories. Med. Phys. 45, 314–321 (2018).
[24]. Arieno, A., Chan, A. & Destounis, S. V. A review of the role of augmented intelligence in breast imaging: from automated breast density assessment to risk stratification. Am. J. Roentgenol. 212, 259–270 (2019).
[25]. Sprague, B. L. et al. Variation in mammographic breast density assessments among radiologists in clinical practice: a multicenter observational study. Ann. Intern. Med. 165, 457–464 (2016).
[26]. Yala, A., Lehman, C., Schuster, T., Portnoi, T. & Barzilay, R. A deep learning

mammography-based model for improved breast cancer risk prediction. Radiology 292, 60–66 (2019). Deep learning models that use full-field mammograms yield substantially improved risk discrimination compared with the standard Tyrer-Cuzick (version 8) risk prediction model.

[27]. Dembrower, K. et al. Comparison of a deep learning risk score and standard mammographic density score for breast cancer risk prediction. Radiology 294, 265–272 (2020).

[28]. Le Boulc'h, M. et al. Comparison of breast density assessment between human eye and automated software on digital and synthetic mammography: Impact on breast cancer risk. Diagn. Interv. Imaging 101, 811–819 (2020).

[29]. Liu, B. et al. Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: three decades' development course and future prospect. J. Cancer Res. Clin. Oncol. 146, 153–185 (2020).

[30]. Li, D. et al. The performance of deep learning algorithms on automatic pulmonary nodule detection and classification tested on different datasets that are not derived from LIDC-IDRI: a systematic review. Diagnostics 9, https:// doi.org/10.3390/diagnostics9040207 (2019). The studies reviewed reached a classification accuracy between 68-99.6% and a detection accuracy between 80.6-94%.

[31]. Schreuder, A., Scholten, E. T., van Ginneken, B. & Jacobs, C. Artificial intelligence for detection and characterization of pulmonary nodules in lung cancer CT screening: ready for practice. Transl. Lung Cancer Res. 10, 2378–2388 (2021).

[32]. Raya-Povedano, J. L. et al. AI-based strategies to reduce workload in breast cancer screening with mammography and tomosynthesis: a retrospective evaluation. Radiology 300, 57–65 (2021).

[33]. Graewingholt, A. & Duffy, S. Retrospective comparison between single reading plus an artificial intelligence algorithm and two-view digital tomosynthesis with double reading in breast screening. J. Med. Screen https:// doi.org/10.1177/0969141320984198 (2021).

[34]. Dembrower, K. et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. Lancet Digit Health 2, e468–e474 (2020).

[35]. Tran, W. T. et al. Computational radiology in breast cancer screening and diagnosis using artificial intelligence. Can. Assoc. Radiol. J. 72, 98–108 (2021).

[36]. McKinney, S. M. et al. International evaluation of an AI system for breast cancer screening. Nature 577, 89–94 (2020).

[37]. Lennox-Chhugani, N., Chen, Y., Pearson, V., Trzcinski, B. & James, J. Women's attitudes to the use of AI image readers: a case study from a national breast screening programme. BMJ Health Care Inform. 28, https://doi.org/10. 1136/bmjhci-2020-100293 (2021).

[38]. Freeman, K. et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. BMJ 374, n1872, (2021).

[39]. Chen, J. H. et al. Opportunistic breast density assessment in women receiving low-dose chest computed tomography screening. Acad. Radiol. 23, 1154–1161 (2016).

[40]. Liu, K. et al. Evaluating a fully automated pulmonary nodule detection approach and its impact on radiologist performance. Radiol. Artif. Intell. 1, e180084 (2019).

[41]. Chakrabarty, S. et al. MRI-based identification and classification of major intracranial tumor types by using a 3D convolutional neural network: a retrospective multi-institutional analysis. Radiol. Artif. Intell. 3, e200301 (2021).

[42]. Deepak, S. & Ameer, P. M. Brain tumor classification using deep CNN features via transfer learning. Comput. Biol. Med. 111, 103345 (2019).

[43]. Diaz-Pernas, F. J., Martinez-Zarzuela, M., Anton-Rodriguez, M. & Gonzalez- Ortega, D. A Deep learning approach for brain tumor classification and

[44]. segmentation using a multiscale convolutional neural network. Healthcare 9, https://doi.org/10.3390/healthcare9020153 (2021).
[45]. Nazir, M., Shakil, S. & Khurshid, K. Role of deep learning in brain tumor detection and classification (2015 to 2020): a review. Comput. Med. Imaging Graph 91, 101940 (2021).
[46]. Dmitriev, K. et al. Classification of pancreatic cysts in computed tomography images using a random forest and convolutional neural network ensemble. Med. Image Comput. Comput. Assist. Interv. 10435, 150–158 (2017).
[47]. Li, H. et al. Differential diagnosis for pancreatic cysts in CT scans using densely-connected convolutional networks. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2019, 2095–2098 (2019).
[48]. Yang, J., Guo, X., Ou, X., Zhang, W. & Ma, X. Discrimination of pancreatic serous cystadenomas from mucinous cystadenomas with CT textural features: based on machine learning. Front. Oncol. 9, 494 (2019).
[49]. Du, R. et al. Radiomics model to predict early progression of nonmetastatic nasopharyngeal carcinoma after intensity modulation radiation therapy: a multicenter study. Radiol. Artif. Intell. 1, e180075 (2019).
[50]. Khorrami, M. et al. Combination of peri- and intratumoral radiomic features on baseline CT scans predicts response to chemotherapy in lung adenocarcinoma. Radiol. Artif. Intell. 1, e180012 (2019).
[51]. Bibault, J. E. et al. Deep Learning and Radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. Sci. Rep. 8, 12611 (2018).
[52]. Delli Pizzi, A. et al. MRI-based clinical-radiomics model predicts tumor response before treatment in locally advanced rectal cancer. Sci. Rep. 11, 5379 (2021).
[53]. Shaish, H. et al. Radiomics of MRI for pretreatment prediction of pathologic complete response, tumor regression grade, and neoadjuvant rectal score in patients with locally advanced rectal cancer undergoing neoadjuvant chemoradiation: an international multicenter study. Eur. Radiol. 30, 6263–6273 (2020).
[54]. Kao, Y. S. & Hsu, Y. A meta-analysis for using radiomics to predict complete pathological response in esophageal cancer patients receiving neoadjuvant chemoradiation. In Vivo 35, 1857–1863 (2021).
[55]. Jin, X. et al. Prediction of response after chemoradiation for esophageal cancer using a combination of dosimetry and CT radiomics. Eur. Radiol. 29, 6080–6088 (2019).
[56]. DiCenzo, D. et al. Quantitative ultrasound radiomics in predicting response to neoadjuvant chemotherapy in patients with locally advanced breast cancer: Results from multi-institutional study. Cancer Med. 9, 5798–5806 (2020).
[57]. Bitencourt, A. G. V. et al. MRI-based machine learning radiomics can predict HER2 expression level and pathologic response after neoadjuvant therapy in HER2 overexpressing breast cancer. EBioMedicine 61, 103042 (2020).
[58]. Pons, E., Braun, L. M., Hunink, M. G. & Kors, J. A. Natural language processing in radiology: a systematic review. Radiology 279, 329–343 (2016).
[59]. Oliwa, T. et al. Obtaining knowledge in pathology reports through a natural language processing approach with classification, named-entity recognition, and relation-extraction heuristics. JCO Clin. Cancer Inform. 3, 1–8 (2019).
[60]. Steinkamp, J. M., Chambers, C. M., Lalevic, D., Zafar, H. M. & Cook, T. S. Automated organ-level classification of free-text pathology reports to support a radiology follow-up tracking engine. Radiol. Artif. Intell. 1, e180052 (2019).
[61]. Holzinger, A., Haibe-Kains, B. & Jurisica, I. Why imaging data alone is not enough: AI-based integration of imaging, omics, and clinical data. Eur. J. Nucl. Med. Mol. Imaging 46, 2722–2730 (2019).
[62]. Saltz, J. et al. Towards generation, management, and exploration of combined radiomics and pathomics datasets for cancer research. AMIA Jt. Summits Transl. Sci. Proc. 2017, 85–

94 (2017).

[63]. Liu, X., Li, K. W., Yang, R. & Geng, L. S. Review of deep learning based automatic segmentation for lung cancer radiotherapy. Front. Oncol. 11, 717039 (2021).

[64]. Kalantar, R. et al. Automatic segmentation of pelvic cancers using deep learning: state-of-the-art approaches and challenges. Diagnostics 11, https:// doi.org/10.3390/diagnostics11111964 (2021).

[65]. van Kempen, E. J. et al. Performance of machine learning algorithms for glioma segmentation of brain MRI: a systematic literature review and meta- analysis. Eur. Radiol. 31, 9638–9653 (2021).

[66]. Dinkel, J. et al. Inter-observer reproducibility of semi-automatic tumor diameter measurement and volumetric analysis in patients with lung cancer. Lung Cancer 82, 76–82 (2013). By using computer-assisted size assessment in primary lung tumor, interobserver-variability can be reduced to about half to one-third compared to standard manual measurements.

[67]. Napel, S., Mu, W., Jardim-Perassi, B. V., Aerts, H. & Gillies, R. J. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. Cancer 124, 4633–4649 (2018).

[68]. Rundo, L. et al. Tissue-specific and interpretable sub-segmentation of whole tumour burden on CT images by unsupervised fuzzy clustering. Comput. Biol. Med. 120, 103751 (2020).

[69]. Savenije, M. H. F. et al. Clinical implementation of MRI-based organs-at-risk auto-segmentation with convolutional networks for prostate radiotherapy. Radiat. Oncol. 15, 104 (2020).

[70]. Chen, X. et al. A deep learning-based auto-segmentation system for organs-at- risk on whole-body computed tomography images for radiation therapy. Radiother. Oncol. 160, 175–184 (2021).

[71]. Vrtovec, T., Mocnik, D., Strojan, P., Pernus, F. & Ibragimov, B. Auto- segmentation of organs at risk for head and neck radiotherapy planning: from atlas-based to deep learning methods. Med. Phys. 47, e929–e950 (2020).

[72]. Chan, J. W. et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. Med. Phys. 46, 2204–2213 (2019).

[73]. Chung, S. Y. et al. Clinical feasibility of deep learning-based auto- segmentation of target volumes and organs-at-risk in breast cancer patients after breast-conserving surgery. Radiat. Oncol. 16, 44 (2021).

[74]. Feng, X., Qing, K., Tustison, N. J., Meyer, C. H. & Chen, Q. Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images. Med. Phys. 46, 2169–2180 (2019).

[75]. Zhu, J. et al. Comparison of the automatic segmentation of multiple organs at risk in CT images of lung cancer between deep convolutional neural network- based and atlas-based techniques. Acta Oncol. 58, 257–264 (2019).

[76]. Shanbhogue, K. et al. Accelerated single-shot T2-weighted fat-suppressed (FS) MRI of the liver with deep learning-based image reconstruction: qualitative and quantitative comparison of image quality with conventional T2-weighted FS sequence. Eur. Radiol. https://doi.org/10.1007/s00330-021-08008-3 (2021). Deep learning image reconstruction demonstrated superior image quality, improved respiratory motion and other ghosting artefacts, and increased lesion conspicuity with comparable liver-to-lesion contrast compared to conventional sequence.

[77]. Chaudhari, A. S. et al. Diagnostic accuracy of quantitative multicontrast 5-minute knee MRI using prospective artificial intelligence image quality enhancement. Am. J. Roentgenol. 216, 1614–1625 (2021).

[78]. Monshi, M. M. A., Poon, J. & Chung, V. Deep learning in generating radiology reports: a

survey. Artif. Intell. Med. 106, 101878 (2020).
[79]. Nakamura, Y. et al. Automatic detection of actionable radiology reports using bidirectional encoder representations from transformers. BMC Med. Inform. Decis. Mak. 21, 262 (2021).
[80]. Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence. Nat. Med. 25, 44–56 (2019).
[81]. Seyhan, A. A. & Carini, C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? J. Transl. Med. 17, 114 (2019).
[82]. Brady, S. M., Highnam, R., Irving, B. & Schnabel, J. A. Oncological image analysis. Med. Image Anal. 33, 7–12 (2016).
[83]. Jimenez-Sanchez, A. et al. Heterogeneous tumor-immune microenvironments among differentially growing metastases in an ovarian cancer patient. Cell 170, 927–938.e920 (2017).
[84]. Martin-Gonzalez, P. et al. Integrative radiogenomics for virtual biopsy and treatment monitoring in ovarian cancer. Insights Imaging 11, 94 (2020).
[85]. Bukowski, M. et al. Implementation of eHealth and AI integrated diagnostics with multidisciplinary digitized data: are we ready from an international perspective. Eur. Radiol. 30, 5510–5524 (2020).
[86]. Mun, S. K., Wong, K. H., Lo, S. B., Li, Y. & Bayarsaikhan, S. Artificial intelligence for the future radiology diagnostic service. Front. Mol. Biosci. 7, 614258 (2020).
[87]. Allen, B. Jr. et al. A road map for translational research on artificial intelligence in medical imaging: from the 2018 National Institutes of Health/ RSNA/ACR/The Academy Workshop. J. Am. Coll. Radiol. 16, 1179–1189 (2019).
[88]. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015).
[89]. Litjens, G. et al. A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60–88 (2017).
[90]. Rajchl, M. et al. DeepCut: object segmentation from bounding box annotations using convolutional neural networks. IEEE Trans. Med. Imaging 36, 674–683 (2017).
[91]. Chalkidou, A., O'Doherty, M. J. & Marsden, P. K. False discovery rates in PET and CT studies with texture features: a systematic review. PLoS ONE 10, e0124165 (2015).
[92]. Zanfardino, M. et al. Bringing radiomics into a multi-omics framework for a comprehensive genotype-phenotype characterization of oncological diseases. J. Transl. Med. 17, 337 (2019).
[93]. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127 (2007).
[94]. Hernan, M. A. & Robins, J. M. Using big data to emulate a target trial when a randomized trial is not available. Am. J. Epidemiol. 183, 758–764 (2016).
[95]. Schwier, M. et al. Repeatability of multiparametric prostate MRI radiomics features. Sci. Rep. 9, 9441 (2019).
[96]. Orlhac, F., Frouin, F., Nioche, C., Ayache, N. & Buvat, I. Validation of a method to compensate multicenter effects affecting CT radiomics. Radiology 291, 53–59 (2019).
[97]. Berenguer, R. et al. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. Radiology 288, 407–415 (2018). Many radiomics features were found to be redundant and nonreproducible, indicating the need for careful feature selection.
[98]. Hagiwara, A., Fujita, S., Ohno, Y. & Aoki, S. Variability and standardization of quantitative imaging: monoparametric to multiparametric quantification, radiomics, and artificial intelligence. Invest. Radiol. 55, 601–616 (2020).
[99]. Fedorov, A. et al. An annotated test-retest collection of prostate multiparametric MRI. Sci. Data 5, 180281 (2018).
[100]. Kalpathy-Cramer, J. et al. Radiomics of lung nodules: a multi-institutional study of robustness and agreement of quantitative imaging features. Tomography 2, 430–437

(2016).
[101]. McNitt-Gray, M. et al. Standardization in quantitative imaging: a multicenter comparison of radiomic features from different software packages on digital reference objects and patient data sets. Tomography 6, 118–128 (2020).
[102]. Zwanenburg, A. et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 295, 328–338 (2020).
[103]. Shortliffe, E. H. & Sepulveda, M. J. Clinical decision support in the era of artificial intelligence. JAMA 320, 2199–2200 (2018).
[104]. Giger, M. L., Chan, H. P. & Boone, J. Anniversary paper: History and status of CAD and quantitative image analysis: the role of Medical Physics and AAPM. Med. Phys. 35, 5799–5820 (2008).
[105]. Helvie, M. A. et al. Sensitivity of noncommercial computer-aided detection system for mammographic breast cancer detection: pilot clinical trial. Radiology 231, 208–214 (2004).
[106]. Birdwell, R. L., Ikeda, D. M., O'Shaughnessy, K. F. & Sickles, E. A. Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. Radiology 219, 192–202 (2001).
[107]. Kohli, A. & Jha, S. Why CAD failed in mammography. J. Am. Coll. Radiol. 15, 535–537 (2018).
[108]. Lehman, C. D. et al. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA Intern. Med. 175, 1828–1837 (2015).
[109]. Fenton, J. J. et al. Influence of computer-aided detection on performance of screening mammography. N. Engl. J. Med. 356, 1399–1409 (2007).
[110]. Rodriguez-Ruiz, A. et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology 290, 305–314 (2019).
[111]. Jaremko, J. L. et al. Canadian association of radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. Can. Assoc. Radiol. J. 70, 107–118 (2019).
[112]. Radiology, E. S. o. ESR position paper on imaging biobanks. Insights Imaging
[113]. 6, 403–410 (2015).
[114]. Guinney, J. & Saez-Rodriguez, J. Alternative models for sharing confidential biomedical data. Nat. Biotechnol. 36, 391–392 (2018).
[115]. Negrouk, A. & Lacombe, D. Does GDPR harm or benefit research participants? An EORTC point of view. Lancet Oncol. 19, 1278–1280 (2018).
[116]. Gallas, B. D. et al. Evaluating imaging and computer-aided detection and diagnosis devices at the FDA. Acad. Radiol. 19, 463–477 (2012).
[117]. Prior, F. et al. The public cancer radiology imaging collections of The Cancer Imaging Archive. Sci. Data 4, 170124 (2017).
[118]. Clark, K. et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging 26, 1045–1057 (2013). TCIA contains 30.9 million radiology images representing data collected from approximately 37,568 subjects; it outlines the curation and publication methods employed by TCIA and makes available 15 collections of cancer imaging data.
[119]. Wilkinson, M. D. et al. A design framework and exemplar metrics for FAIRness. Sci. Data 5, 180118 (2018).
[120]. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 3, 160018 (2016).
[121]. Prior, F. et al. Open access image repositories: high-quality data to enable machine learning research. Clin. Radiol. 75, 7–12 (2020).
[122]. Vayena, E., Blasimme, A. & Cohen, I. G. Machine learning in medicine: addressing ethical challenges. PLoS Med. 15, e1002689 (2018).

[123]. Müller, H., Kalpathy-Cramer, J. & Seco de Herrera, A. G. Information retrieval evaluation in a changing wolrd. 41 (2019).
[124]. von Elm, E. et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 370, 1453–1457 (2007).
[125]. Castro, D. C., Walker, I. & Glocker, B. Causality matters in medical imaging.
[126]. Nat Commun 11, 3673 (2020).
[127]. Langlotz, C. P. Will artificial intelligence replace radiologists? Radiol. Artif. Intell. 1, e190058 (2019).
[128]. Bizzo, B. C., Almeida, R. R., Michalski, M. H. & Alkasab, T. K. Artificial intelligence and clinical decision support for radiologists and referring providers. J. Am. Coll. Radiol. 16, 1351–1356 (2019).
[129]. Lou, R., Lalevic, D., Chambers, C., Zafar, H. M. & Cook, T. S. Automated detection of radiology reports that require follow-up imaging using natural language processing feature engineering and machine learning classification. J. Digit. Imaging 33, 131–136 (2020).
[130]. US Food and Drugs Adminstration. Machine Learning (AI/ML)-based Software as a Medical Device (SaMD). (2019).
[131]. Panch, T., Mattie, H. & Celi, L. A. The "inconvenient truth" about AI in healthcare. NPJ Digit. Med. 2, 77 (2019).
[132]. Clinical Radiology. UK workforce census 2020 report. (Royal College of Radiologists, 2020).