Received:
03 October 2019

Revised:
12 January 2020

Accepted:
13 January 2020

<https://doi.org/10.1259/bjr.20190855>

Cite this article as:

El Naqa I, Haider MA, Giger ML, Ten Haken RK. Artificial Intelligence: reshaping the practice of radiological sciences in the 21st century. *Br J Radiol* 2020; **93**: 20190855.

BJR 125th ANNIVERSARY: REVIEW ARTICLE

Artificial Intelligence: reshaping the practice of radiological sciences in the 21st century

¹ISSAM EL NAQA, PhD, ²MASOOM A HAIDER, MD, ³MARYELLEN L GIGER, PhD and ¹RANDALL K TEN HAKEN, PhD

¹Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, USA

²Department of Medical Imaging and Lunenfeld-Tanenbaum Research Institute, University of Toronto, Toronto, ON, Canada

³Department of Radiology, University of Chicago, Chicago, IL, USA

Address correspondence to: Issam El Naqa

E-mail: ielnaqa@med.umich.edu

ABSTRACT

Advances in computing hardware and software platforms have led to the recent resurgence in artificial intelligence (AI) touching almost every aspect of our daily lives by its capability for automating complex tasks or providing superior predictive analytics. AI applications are currently spanning many diverse fields from economics to entertainment, to manufacturing, as well as medicine. Since modern AI's inception decades ago, practitioners in radiological sciences have been pioneering its development and implementation in medicine, particularly in areas related to diagnostic imaging and therapy. In this anniversary article, we embark on a journey to reflect on the learned lessons from past AI's chequered history. We further summarize the current status of AI in radiological sciences, highlighting, with examples, its impressive achievements and effect on re-shaping the practice of medical imaging and radiotherapy in the areas of computer-aided detection, diagnosis, prognosis, and decision support. Moving beyond the commercial hype of AI into reality, we discuss the current challenges to overcome, for AI to achieve its promised hope of providing better precision healthcare for each patient while reducing cost burden on their families and the society at large.

INTRODUCTION

Humans have been always fascinated in the development of *intelligent machines* that can emulate their cognition and serve their needs.¹ Homer, in the Greek mythology wrote about Hephaestus, a blacksmith, who manufactured mechanical servants to help their masters. Al-Jazri, a 13th century artisan and mathematician is accredited with building the first *humanoid (robot)* that performed music and washed hands among other human-like functions.² However, artificial intelligence (AI) extends beyond robotics into processing information and invoking cognitive computing systems. For instance, the Turing machine, is a computing device that can simulate an algorithmic logic much like a human. Hence, it has long been argued that an intelligent machine, a machine that can think like humans, needs to pass the Turing test; *i.e.* when the responses from the computer and the human become indistinguishable.³ Although it remains argumentative whether such a test is indicative of the cognitive implications of true human-like intelligence, the notion of AI is satisfactory, as it indicates an aiding tool that can support human needs and improve their quality of life, thus dispelling fictional fears of a machine takeover! This

is particularly true in the areas for practitioners of radiological sciences (diagnostic and therapy), who have been the vanguards of new technologies in medicine including early adoption of AI development and implementation as will be highlighted in this *BJR* 125th anniversary article, which is quite fulfilling since AI has currently become front and center in radiological sciences.

WHAT IS ARTIFICIAL INTELLIGENCE?

Formally, AI as a field was conceived at the 1956 Dartmouth Summer Research Project (Dartmouth College, Hanover, 1956) by John McCarthy and colleagues, who coined the term “artificial intelligence.” According to McCarthy “AI is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.”^{4,5} Hence, AI currently can be thought of as representing a wide range of technologies that involve developing machines that can perform tasks that are characteristic of human's (or other beings') intelligence, *e.g.* playing a boardgame or autonomously driving a vehicle. The subcategory of AI that is primarily

Figure 1. Venn diagram of the relationship between AI, ML, and DL. AI, artificial intelligence; DL, deep learning; ML, machine learning.

dedicated to the development of intelligent computer programs (algorithms) is denoted as *machine learning* (ML). The term was coined by another noted attendee of the Dartmouth workshop, IBM engineer Arthur Samuel, whom described ML as “a *field of study that gives computers the ability to learn without being explicitly programmed*.⁵” Traditionally, ML has relied on feeding computer-extracted human-engineered patterns (features) derived from the raw data, e.g. using computer vision methods, to an algorithm to perform a designated learning task; a process colloquially referred to now as *shallow learning* (SL). This is in contrast to a special subcategory of ML that allows for combined data representation (e.g. feature extraction) and task learning (e.g. classification or detection) known as *deep learning* (DL). Conceptually, DL comprises representation learning methods that are fed with raw data and can automatically discover the features needed for detection or classification using any ML approach.⁶ However, pragmatically this has been shown by Hinton and colleagues to be most successfully implemented currently with deep neural networks methods,^{7,8} in which the frontend layers of the neural network learn the data representation and then pass this information into backend layers to perform the classification or detection tasks, all in one framework. Hence, in the case of a standard convolutional neural network (CNN), one of the most commonly used architectures for DL, the convolutional layers are tasked with learning the appropriate pattern representation (features) from raw image data and passing these features into the fully connected layers to perform the classification task.^{7,8} A schematic of the relationship between AI, ML, and DL is shown in Figure 1. For instance, if the goal is to develop a computer-aided diagnosis (CADx) system that is comparable to a radiologist for distinguishing a benign from a malignant nodule on an imaging scan, then the hardware and software infrastructure

of the CADx as a whole would constitute the AI-aid system. The collective computer set of rules, *i.e.* the algorithm, that is used to learn from annotated data, which nodule is benign or malignant would be the ML algorithm (classifier) that is being trained to classify these nodules. If the ML algorithm design requires extraction of human-engineered features, the learning process would be designated as an ML/SL procedure while if the algorithm operates directly on the raw imaging data and is able to automatically learn the data representation, then the learning process is designated as an ML/DL procedure. ML/DL has the intrinsic advantage of mitigating the statistical bias of human-engineered feature selection processes that are associated with ML/SL, leading to superior performances in many applications, e.g. imaging analytics.

Furthermore, ML algorithms that require labeled data (annotations) for training, *i.e.* (input, output) pairs, undergo what is denoted as *supervised learning*, *i.e.* learning with a teacher. In contrast, if only unlabeled data are used as input data, this ML training is referred to as *unsupervised learning*, *i.e.* inferring patterns within the data. ML training in which part of the data is labelled and the rest unlabeled is referred to as *semi-supervised learning* and is generally thought to be more representative of human learning. If the algorithm learns to map inputs into optimized actions, this is denoted as *reinforcement learning*, *i.e.* goal-oriented tasks. These algorithms currently represent the main categories of ML, with supervised learning being the most common type in radiological sciences with applications ranging from detection, to diagnosis, to therapeutic interventions. However, several techniques are emerging to relieve the burden and cost of data labeling in supervised learning including: the semi-supervised approach mentioned above, *transfer learning* (using knowledge from other domains, such as natural images when learning medical ones), *active learning* (an interactive approach with human being involved), and more recently *weak supervised learning*, where the labels are assumed to be imprecise or noisy.⁹ Unsupervised learning is typically used for clustering or data reduction tasks while reinforcement learning is applied for optimizing sequential decision-making processes, in clinical management for instance. A summary of some of the common applications of these different ML/DL categories in radiological sciences is given in Table 1.

Historically, applications of ML in medicine demonstrated decision-support capabilities in detecting disease onset or analyzing the relationship between treatment and response. Ledley and Lusted in their seminal 1959 paper, anticipated the role of a probabilistic logic-based approach to understand and support physicians’ reasoning¹²⁴. An early major AI initiative was the MYCIN project at Stanford in the 1970s, which was a rule-based system to identify bacteria types that may cause infectious diseases¹²⁵, achieving an acceptability rating of 65% from a panel of experts¹²⁶. However, several radiological applications in medical imaging preceded MYCIN, with Winsberg *et al.* reporting, in 1967, on computer detection of radiographic abnormalities in mammograms with computer analysis¹⁰ and with Lodwick *et al.* presenting a roentgenograms concept for analyzing bone and lung cancer images^{47,48} and Meyers *et al.*

Table 1. Sample radiological applications of ML algorithms in radiological sciences

ML category	Applications	References in text
Supervised learning	Detection of objects in images	10–22
	Auto contouring and segmentation	23–46
	Diagnosis of cancer lesions	11,12,14,47–62
	Image prognosis	21,63–65
	Pathological analysis	66–71
	Image analysis for risk assessment	72–75
	Image quality	76–81
	Quality and assurance	82–86
	Treatment planning and management	87–102
	Response prediction	101,103–113
Weak supervised learning	EMR and imaging integration	114,115
Unsupervised learning	Clustering for tumor segmentation	116,117
	Dimensionality reduction for classification	102,118
	Error detection and quality assurance	119
	Image registration	120
Reinforcement learning	Treatment adaptation	121,122
	Planning optimization	123

ML, machine learning.

developing an automated computer analysis of cardiothoracic ratios.⁷²

Through pioneering efforts at the University of Chicago in the 1980s, computer-aided diagnosis (CAD) became a major subject in medical imaging and diagnostic radiology, providing radiologists with computer output as a “second opinion” to aid in making final decisions.^{11–13,49–51} These CAD systems utilized image feature-based analysis for the detection of microcalcifications in mammogram images¹⁴ and lung nodules in digital chest radiographs.⁵² It is important and interesting to note that the subtle change from computer readers to computer aids allows for a more acceptable approach to incorporating AI into clinical practice, possibly leaving the final decision to the human clinician. However, current issues surrounding physician liability and medical AI use are complex and many challenges are yet to be resolved. Therefore, new guidance on assigning liability for injuries that may arise from interaction between algorithms and practitioners (including appropriate use of the AI system by the human) is needed.^{127,128}

These CAD applications were extended in the 1990s by application of artificial neural networks (ANN) generally¹⁵ and convolutional neural networks (CNNs) specifically^{16,17,53} as reviewed in the literature.^{11–13,50,51,73} In the 2000s, CAD applications mainly utilized ML methods based on support vector machines (SVM) algorithms^{18,54,55} paralleling advances in machine learning at the time. A similar trend was noted in therapeutic radiology, with applications in mid 1990s using ANNs for automating treatment planning evaluation,⁸⁷ beam orientation customization,⁸⁸ and standardization (knowledge-based planning).⁸⁹ In the 2000s,

these applications focused on predicting radiotherapy response with ANNs^{103–105} and SVMs.^{106–108}

The promise of AI in medicine in general and radiological sciences in particular have been hyped in the past, but, despite progress, failed on predictions in the 1970s that by the year 2000 computers would act as a powerful extension of the physician's intellect.^{129,130} This may be partly due to clinicians not using AI as intended or as FDA approval may have originally required. Although AI had its fair share of boom and busts, denoted as *AI winters*, applications of AI in medicine are witnessing a new but more cautious spring with applications spanning all aspects of practice. A summary of publications of AI in radiology and radiation oncology in the past 30 years is shown in Figure 2 using PubMed with search terms (“Radiology”

Figure 2. AI related publications in radiology and radiation oncology over the past 30 years. AI, artificial intelligence.

OR “Radiation Oncology”) AND (“Artificial Intelligence” OR “Machine Learning”). The figure highlights the rapid increase in AI consideration, especially in more recent years.

Radiological science practitioners, have been at the forefront utilizing AI, leading its implementations, particularly in areas related to diagnostic imaging and therapy. This anniversary article highlights the current status of AI in the radiological sciences, noting its achievements and existing challenges. It is intended to provide a realistic vision of its role, not as a competitor to human cognition but as a partner in delivering improved healthcare by combining AI/ML software with the best human clinician knowledge (*i.e.* human-in-the-loop) to permit delivering care that outperforms what either can do alone and thus improving both credibility and performance.^{131–134}

AI today for radiological sciences

In the past, the U.S. Food and Drug Administration (FDA) has approved AI systems for clinical use, including the R2 ImageChecker system, which received FDA approval in 1998¹⁹ for use as an aid in the detection of breast abnormalities on screening mammograms (CADE), *i.e.* as a second reader. Interestingly, the R2 ImageChecker system included a CNN.^{17,19} The approval launched the first phase of FDA-approved systems for detection of lesions. More recently, the FDA is approving multiple platforms that incorporate AI, signaling a new era in medicine and radiology.¹³⁵ Among these, the first system cleared was the QuantX Advanced system to aid in breast cancer diagnosis (CADx), developed originally by Giger and colleagues.^{56,57,63,116,136} Today, with the success of deep convolutional neural networks in the classification of natural objects in digital photograph images, the field of diagnostic imaging with its digital workflow and petabytes of digital medical images is becoming a natural target for AI studies. CT and MRI scans typically involve hundreds of

images per exam; reviewing the sheer volume of these images imposes limitations on the productivity of practicing radiologists. Thus, a natural application of AI is to assist the radiologist in the detection and classification of abnormalities on medical images to improve productivity without loss of diagnostic accuracy. Furthermore, there is a focus on quality and efficiency in healthcare systems due to rising costs. Radiologists are highly skilled and expensive professionals, so any tool that can maximize quality and efficiency is of great interest. ML/DL also can improve image quality by improving current image reconstruction and filtering of modalities such as MRI,⁷⁶ CT⁷⁷ and ultrasound⁷⁸ as well as potentially improve their spatial resolution. Finally, precision medicine requires derivation of reproducible imaging biomarkers.¹³⁷ There is potential of ML/DL to help in this task as well as will be discussed in the following.

Applications in computer-aided detection, diagnosis (CADE,x) and decision support

Applications in diagnostics radiology have witnessed steady growth and demonstrated remarkable progress in this new era of AI spanning its workflow as shown in Figure 3.^{66,138–141} In the following, we highlight some representative cases. Note that along this workflow, AI can aid in setting acquisition parameters, reconstructing images, assessing image quality of acquired images, preprocessing and “hanging” images for interpretation, along with the decision-making tasks of detection, diagnosis, patient management, and treatment response.

Case triage is an example of how ML can improve radiologists’ workflow, *i.e.* CADt. An AI algorithm identifies exams with a higher likelihood of containing critical findings and puts them at the top of a reporting list for the radiologist to reduce turnaround and potentially improve outcomes. Chilamkurthy *et al.*

Figure 3. Permission has been obtained to reproduce this image. The schematics highlights AI role in radiology.¹³⁸

recently demonstrated using deep Residual Network (ResNet), a popular CNN architecture that can have deeper layers than standard CNNs without being subject to degradation by preserving identity connection (short-cuts) through skipping some layers, an area under the ROC curve of greater than 0.9 for a variety of critical findings such as haemorrhage and calvarial fractures on unenhanced head CT in a retrospective study involving 313,318 scans from about 20 centers.²⁰

Initial applications of ML in medical image *segmentation* have focused on unsupervised learning techniques such as clustering methods.^{67,116} More recent studies have shown the promise of supervised learning with CNN's in segmentation of organs largely empowered by *transfer learning* from natural images databases (e.g. Google's ImageNet), and no doubt we will see further development in this arena.^{23,24,117} The most commonly used DL algorithm for biomedical image segmentation is based on the U-Net architecture, which consists of two symmetric CNNs, one is an encoding (contracting) path and the other is a decoding (expanding) path, and hence the U shape.²⁵ Fatty infiltration of the liver and the potential for liver damage related to non-alcoholic steatohepatitis is an important health concern. In a recent study Graffy *et al.* used a modified three-dimensional U-Net for liver segmentation to assess the distribution of fatty infiltration of the liver in 11,669 CT scans in 9552 adults and published the distribution hepatic steatosis, demonstrating excellent performance compared to manual assessment.⁶⁸ The feasibility of routine addition of liver steatosis measurement using a fully automated tool can thus be envisioned to add value to routine CT reports and development as a useful population screening tool.

When dealing with different medical imaging modalities (CT, MR, positron emission tomography, single-photon emission CT, ultrasound, etc.), an important task is to bring these images into the same frame of reference, a process denoted as image *registration*. Varying algorithms for rigid and deformable registration have been proposed in the literature using intensity or feature-based approaches^{142,143} and more recently ML/DL algorithms.¹⁴⁴ For example, de Vos *et al.*, developed an unsupervised learning technique to train CNNs for image registration using intensity-based metrics. They demonstrated its performance for registration of cardiac cine MRI and the registration of chest CT with comparable results to conventional image registration while being orders of magnitude faster.¹²⁰

ML/DL algorithms can be used to develop *prognostic and predictive* tools in cancer therapeutics beyond the RECIST criterion with its known shortcomings. Hosny *et al.* demonstrated the feasibility of prognostication using CT in non-squamous cell lung cancer assessment achieving area under the curves (AUCs) of 0.7–0.71 in surgical and radiation treated cohorts with a CNN.²¹ While Hylton *et al.*⁶⁴ demonstrated the benefit of a semi-manual functional tumor volume from MRI to predict recurrence-free survival-results from the ACRIN 6657/CALGB 150007 I-SPY 1 TRIAL on neoadjuvant chemotherapy for breast cancer. Drukker *et al.*⁶⁵ showed that an automated computer-extracted most-enhancing tumor volume radiomic feature could perform as

well. Recent applications have utilized weak supervised learning for integration of electronic health records with imaging data to improve detection and diagnosis. For instance, Wang *et al.* demonstrated this approach for classification and localization of common thoracic diseases from chest X-ray images¹¹⁴ and Zhu *et al.* for detection of pulmonary nodules in lung CT images.¹¹⁵

Maximizing *image quality* is the goal of manufacturers of imaging equipment with the ultimate aim being to improve diagnostic accuracy. These improvements can come with added costs, radiation exposure, and often increases in time on the scanner. Post-acquisition processing of images or use of ML/DL approaches has shown promise in improving image quality and potentially improving patient throughput. Improving spatial resolution of images of the knee on MRI⁷⁹ and reduction of the dose of intravenous Gd-based contrast agents for brain MRI are examples of this potential.⁸⁰ Similar efforts have been demonstrated in thoracic CT using DL to assess acceptable image quality prior to interpretation.⁸¹

In prostate cancer, there is a major shift in the paradigm of diagnosis, which involves the use of multiparametric MRI (mpMRI) for cancer detection prior to biopsy. It is anticipated that there will be a marked increase in demand for prostate mpMRI. CNNs have shown promise in detection of prostate cancer on MRI showing near human performance and they can potentially be used to aid the radiologist in reporting and improve efficiency and consistency of reporting across the radiology community, thus reducing costs.⁷⁴ Further work to show robustness of these algorithms across institutions and MRI equipments is still needed (Figure 4).

Many of the original advancements in AI in radiology have occurred in breast imaging,^{58,59,73,75,118,137,145} where CNNs have contributed to breast cancer risk assessment, detection, and diagnosis. Much research has focused on the formats of the input and the proper usage of outputs, as “garbage in, garbage out” can stymie success. For example, in Figure 5, Antropova *et al.*⁵⁹ demonstrated how CNN performance varies depending on the input format of breast MRIs. In addition, recent investigations have shown that the fusion of human-engineered imaging features (radiomics) with DL extraction methods can yield statistically significant improvements as compared to either one separately.⁶⁰

ML advances are also contributing to cancer discovery where, through the benefit of multidisciplinary, multi-institution collaborations, the computer-extracted tumor information (*i.e.* image-based phenotypes) is mapped to other “-omics” such as histopathology and genomics.^{61,69–71,109} Figure 6 demonstrates significant associations that were found between breast MRI radiomic phenotypes with genomic features as part of the NCI TCGA/TCIA collaboration. In cancer discovery, once relationships between image features and genomics (radiogenomics) are found to yield an image-based “cancer signature,” these computer-extracted radiomic features can be used as “virtual” biopsies for when an actual biopsy is not practical, such as in cancer screening, monitoring response to therapy, or assessing risk of recurrence.^{62,137}

Figure 4. Permission has been obtained to reproduce this image. Explainable AI: A Class activation map from Ensemble Learning Architecture with DeepConvolutional Neural Network. An ensemble deepconvolutional neural network gave the probability of cancer in this patient of 83%. The patient has a clinically significant Gleason grade Group II cancer. (A) An apparent diffusion coefficient map from MRI annotated by a radiologist showing the cancer outlined in blue. (B) A gradient-weighted class activation map (Grad-Cam) shows that the area of highest activation matches the site of tumor giving confidence that the DCNN is deriving the high probability of cancer based on the features of the cancer itself (reprinted from Cao *et al.*⁷⁴). AI, artificial intelligence; DCNN, deep convolutional neural network.

Applications in radiation oncology

The treatment of cancer using high energy ionizing radiation, known as *radiotherapy*, involves a large set of complex processes spanning consultation to treatment planning to delivery with a large amount of heterogeneous information (*e.g.* clinical, dosimetric, imaging, and biologics), human-machine interaction, optimization and decision making. Hence, ML/DL can help automate many of these processes while allowing for personalization of treatment via application of its predictive analytics^{146–148} as depicted in the schematics of Figure 7.

Automation applications of ML/DL in radiotherapy has been an area of tremendous progress and it includes but is not limited to: (1) auto-contouring (segmentation) of organs, tissues and the targeted cancer,^{22,26–46} which is a major bottleneck in terms of treatment planning efficiency and the availability of such ML/DL tools can reduce this burden significantly; (2) dose prediction,^{90–92} which is an area where AI can improve efficiency in radiotherapy and mitigate prior authorization delays; (3) knowledge-based treatment planning,^{93–97,123,149} which is important for mitigating common errors, particularly in under served or disadvantaged settings with limited resources; (4) radiation physics (machine or patient) quality assurance (QA) to improve safe radiation administration^{82–86,119}; and (4) motion (respiratory or cardiac) management during radiation delivery^{98,99} to enhance tumor targeting and limit uninvolved tissue exposure.

Predictive analytics of ML/DL in radiotherapy includes: (1) image-guidance^{100,150} to ensure that the planning objectives of tumor targeting and normal tissue avoidance are maintained; (2) response modeling of treatment outcomes such as tumor control or toxicities^{101,102,106,107,110–113}; and treatment adaptation to account for geometrical changes¹⁵¹ or adapting prescription to improve outcomes.^{121,122} Details about these and other

applications are reviewed in the literature.¹⁵² An example for adaptive radiotherapy to improve outcomes is shown in Figure 8, which combines changes in imaging (radiomics) with dosimetric (dosimetrics) and biological factors (genomics) to personalize response to radiotherapy. In this case, for a population of non-small cell lung cancer patients, PET-FDG as detected by mid-treatment was used for escalating the radiation dose and was shown to improve local tumor control at 2 year follow-up.¹⁵³ However, the dose adaptations in such studies were based on clinicians' subjective assessment. Alternatively, to objectively assess the adaptive dose per fraction, a 3-component deep reinforcement learning (DRL) approach with neural network architectures was developed. The architecture was composed of (1) a generative adversarial network (GAN) to learn patient population characteristics necessary for DRL training from a relatively limited sample size; (2) a radiotherapy artificial environment was reconstructed by a deep neural network utilizing both original and synthetic data (by generative adversarial network) to estimate the transition probabilities for adaptation of personalized radiotherapy patients' treatment courses; and (3) a deep Q-network (DQN) applied to the radiotherapy artificial environment for choosing the optimal dose in a response-adapted treatment setting. The DRL was trained on large-scale patient characteristics including clinical, genetic, and imaging radiomics features in addition to tumor and lung dosimetric variables. In comparison with the clinical protocol,¹⁵³ although on average slightly less aggressive, the DRL achieved a root mean-squared error = 0.5 Gy for dose escalation recommendation. Interestingly, on occasion, the DRL also seemed to suggest better decisions than the clinical ones in terms of mitigating toxicity risks and improving local control as seen in Figure 8.¹²¹

Figure 5. Full breast MRI images and ROIs for (a) a MIP image of the second postcontrast subtraction MRI, (b) a center slice of the second post-contrast MRI, (c) a central slice of the second postcontrast subtraction MRI, and (d) ROC curves showing the performance of three classifiers. The classifiers were trained on CNN features extracted from ROIs selected on the three input formats. (reprinted from Antropova *et al.*⁵⁹). CNN, convolutional neural network; MIP, maximum intensity projection; ROC, receiver operating characteristic.

AI necessary infrastructure and resources

Access to large well-annotated data sets remains the number one requirement for advancement of AI in radiological sciences. The majority of currently implemented AI approaches rely upon supervised learning techniques. Thus, radiologists face the daunting task of annotating these large data sets. The process of curation and annotation can be helped through mining of radiology reports using natural language processing (NLP), semi- or weak-supervised learning to help reduce the number of cases that must be manually annotated. There have been enormous efforts to collect, curate, and share large data sets through the NCI's TCIA,¹⁵⁴ which intakes clinical trial studies as well as investigators' research data sets.

Storage of large data sets scrubbed of patient information is not part of the normal archival process for patient records. *Data lakes* specifically designed for research need to be created that can handle large imaging and therapy data sets. This requires investment and hiring of data science teams to manage these data lakes. Furthermore, data provenance and governance policies that address concerns of institutional review boards

and the broader ethical issues around the use of large patient data sets for research and sharing of large amounts of data with for profit companies also need to be developed.

In addition to curation of data, there is a need to curate or QA the data analytic pipelines developed to train, test and validate an ML/DL algorithm. Live data scenarios or even changes in versions of software, hardware platforms and open source libraries can all affect the predicted output. Being able to reproduce and share analytic pipelines in the community is vital to develop confidence in the results of AI studies.

For the radiologist/oncologist having an intuitive sense of how the classifier reached its conclusion is of utmost importance particularly as we are in the early days of AI testing and adoption. This is especially true when the AI-based analysis result in a conclusion discrepant with the physician own's opinion. Unfortunately, many DL algorithms suffer from the black box stigma.¹⁵⁵ Several approaches to derive proxy models, develop attention maps, provide disentangled representation or learn with known operators have been emerging to create a more

Figure 6. Overview of identified statistically significant associations between radiomic phenotypes and genomic features. Each node is a genomic feature or a radiomic phenotype. Each line is an identified statistically significant association. Associations were deemed as statistically significant if the (multiple comparison) adjusted p -values ≤ 0.05 . (reprinted from Zhu *et al.*⁷¹).

interpretable/explainable AI paradigm.^{156–160} Eventually, if AI performs consistently and at a better than human performance level then the need for fully explainable AI may diminish. However, AI researchers and industry also have recognized this need by developing new tools to quantify uncertainties,¹⁶¹

to understand feature attributions,¹⁶² or to investigate model's behavior such as the *What-If* tool in TensorFlow by Google, for instance.

Figure 7. The schematics highlights AI role in radiation oncology. Adapted from Meyer *et al.*¹⁴⁶ AI, artificial intelligence; PET, positron emission tomography; QA, quality assurance; CBCT, cone beam CT.

Figure 8. A deep reinforcement learning for automated radiation adaptation in lung cancer. (a) Three different neural networks are used for synthetic data generation (GAN), the radiotherapy environment (DNN), and Deep Q-network (DQN). (b) Performance results showing DQN (black solid line) vs clinical decision (blue dashed line) with RMSE error = 0.5 Gy. Evaluation against eventual outcomes of good (green dots), bad (red dots), and potentially good decisions (orange dots) are shown suggesting not only comparable but also at instances better overall performance by the DQN (reprinted from Tseng *et al.*¹²¹). DNN, deep neural network; DQN, Deep Q-network; GAN, generative adversarial network; RMSE, root mean-squared error.

Current academic and commercial tools

Tools for annotation of images are available, however, they are not necessarily optimized for use in image analysis in the AI era. There is a growing need for common annotation platforms that are in the open source community. These tools should provide access to segmentations using standard formats, so that these can be fed into analytic pipelines and shared across the community. 3D slicer¹⁶³ is an example of an open source viewing and annotation tool. CLARA is a software collection of developer toolkits built on NVIDIA's computing platform aimed at accelerating computation, artificial intelligence, and advanced visualization being offered to the research community.¹⁶⁴

Future of AI in radiological sciences

Significant progress in computational hardware, with parallel processing capabilities (*e.g.* Graphics Processing Unit, Tensor Processing Unit, Intelligent Processing Unit etc.), has accelerated the calculation of large and sophisticated algorithms from weeks to days and few hours only. This is coupled with the availability of open-library software platforms (*e.g.* TensorFlow,¹⁶⁵ Pytorch,¹⁶⁶ etc.) with notable success rivaling that of humans when applied to large databases such as ImageNet,¹⁶⁷ a large-scale database with more than one million annotated natural images (animals, plants, persons, devices, instruments etc) with about 20,000 text descriptors. However, an *ImageNet*-like medical imaging database is yet to be realized with caveats related to data sharing, provenance, patient privacy and the

nature of medical imaging acquisition that not only varies in technologies and parameters but also shifts over time with new developments. Moreover, issues related to learning bias^{168,169} and adversarial examples^{170,171} need to be accounted for too. For instance, an ML algorithm developed for predicting the risk of pneumonia counter intuitively suggested that patients with pneumonia and asthma to be at a lower risk of death than patients with pneumonia but without asthma.¹⁶⁹ Similar controversial examples were noted in the case of skin cancer risk prediction, where the presence of a ruler in the image may be a cue for the ML algorithm of high risk¹⁷² or the appearance of a tube in a chest X-ray being indicative of severe lung disease.¹⁷³ These examples and others stress the importance of data quality and context when training and applying ML/DL algorithms.

The development of trustworthy AI is currently a major goal in the AI community for medical applications with efforts led by the European Union and other international bodies that advocate for lawful, ethical and robust application from technological and societal perspectives. Towards this goal, there has been shifts toward developing more explainable/interpretable AI algorithms,¹⁷⁴ which would allow for better transparency, oversight, and accountability. The radiological clinic of the future can be thought of to be composed of explainable AI platforms with streamlined access to data lakes, assisting clinicians to improve efficiency as well as achieve better diagnostic or therapeutic decisions.

CONCLUSION

The past few years have witnessed a tremendous rise in AI applications to a wide range of areas in radiological sciences (diagnostic and therapy) including automation of segmentation, improving image quality, and developing decision-support systems for personalization of detection and treatment. AI is expected to alter the way patients are being managed and how doctors reach their clinical decisions. Appropriately developed and used, such diagnostics are expected to be faster, cheaper, and more accurate than ever. However, one has to be mindful of the characteristics and the limitations of this disruptive technology as well. AI methods require large amounts of data for their training and validation, which also beg the questions of computerized trust, data sharing, and privacy concerns. With the pre-existing domain knowledge, the merits of human-engineered analytics or features standing for accumulative knowledge based on numerous observations should be incorporated, hybridized, and inherently infused with modern deep learning architectures. Aside from commercial hype, the cautious deployment and

assistance of the state-of-the-art AI technologies, imaging informatics hold tremendous potentials to provide better precision and personalized healthcare for patients while reducing the cost burden on society.

ACKNOWLEDGMENT

IEN and RTH acknowledge support from National Institute of Health (NIH) grants P01 CA059827, R37-CA222215, and R01-CA233487. MLG acknowledge support from NIH U01 CA195564. The authors would like to thank Mr Steven Kronenberg for help with graphics.

CONFLICTS OF INTEREST

IEN: serves as scientific advisor for Endectra LLC and receives NIH funding. MAH: None. MLG: is a stockholder in R2 Technology/Hologic and a cofounder and equity holder in Quantitative Insights (now Qlarity Imaging) and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Riverain Medical, Mitsubishi, and Toshiba. RTH: receives NIH funding.

REFERENCES

- Bruce GB. A (very) brief history of artificial intelligence. *AI Magazine* 2005; **26**.
- Rosheim ME. *Robot Evolution: The Development of Anthrobotics*; Wiley; 1994.
- Turing, AM. Computing machinery and intelligence. *Mind* 1950; **236**: 433–60.
- McCarthy J. What is Artificial Intelligence. 2007. Available from: <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>.
- Samuel AL. Some studies in machine learning using the game of checkers. *IBM J. Res. & Dev.* 1959; **3**: 210–29. doi: <https://doi.org/10.1147/rd.33.0210>
- Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986; **323**: 533–6. doi: <https://doi.org/10.1038/323533a0>
- Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science* 2006; **313**: 504–7. doi: <https://doi.org/10.1126/science.1127647>
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436–44. doi: <https://doi.org/10.1038/nature14539>
- Zhou Z-H. A brief introduction to weakly supervised learning. *National Science Review* 2018; **5**: 44–53. doi: <https://doi.org/10.1093/nsr/nwx106>
- Winsberg F, Elkin M, Macy J, Bordaz V, Weymouth W, et al. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology* 1967; **89**: 211–5. doi: <https://doi.org/10.1148/89.2.211>
- Doi K. Artificial intelligence and neural networks in radiology: Application to computer-aided diagnostic schemes. In: Hendee W, Trueblood J, eds. *Digital Imaging*: AAPM Medical Physics Monograph; 1993. pp. 301–22.
- Sonka M, Fitzpatrick M. eds. Computer-aided diagnosis in mammography. In: *Handbook of Medical Imaging*; 2000. pp. 915–1004.
- Swett H, Giger M, Doi K. Computer vision and decision support. In: Hendee, Wells P, eds. *in Perception of Visual Information*: Springer-Verlag; 1993. pp. 272–315.
- Chan HP, Doi K, Galhotra S, Vyborny CJ, MacMahon H, Jokich PM, et al. Image feature analysis and computer-aided diagnosis in digital radiography. I. automated detection of microcalcifications in mammography. *Med Phys* 1987; **14**: 538–48. doi: <https://doi.org/10.1118/1.596065>
- Wu Y, Doi K, Giger ML, Nishikawa RM. Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks. *Med Phys* 1992; **19**: 555–60. doi: <https://doi.org/10.1118/1.596845>
- Chan HP, Lo SC, Sahiner B, Lam KL, Helvie MA. Computer-aided detection of mammographic microcalcifications: pattern recognition with an artificial neural network. *Med Phys* 1995; **22**: 1555–67. doi: <https://doi.org/10.1118/1.597428>
- Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA, et al. Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. *Med Phys* 1994; **21**: 517–24. doi: <https://doi.org/10.1118/1.597177>
- El-Naqa I, Yang Y, Wernick MN, Galatsanos NP, Nishikawa RM. A support vector machine approach for detection of microcalcifications. *IEEE Trans Med Imaging* 2002; **21**: 1552–63. doi: <https://doi.org/10.1109/TMI.2002.806569>
- Roehrig Jet al. Clinical results with r2 imagechecker system. In: *Digital Mammography*: Kluwer Academic Publishers; 1998. pp. 395–400.
- Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *The Lancet* 2018; **392**: 2388–96. doi: [https://doi.org/10.1016/S0140-6736\(18\)31645-3](https://doi.org/10.1016/S0140-6736(18)31645-3)
- Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med* 2018; **15**: e1002711. doi: <https://doi.org/10.1371/journal.pmed.1002711>
- Liang S, Tang F, Huang X, Yang K, Zhong T, Hu R, et al. Deep-learning-based detection and segmentation of organs at risk in nasopharyngeal carcinoma computed tomographic images for radiotherapy planning. *Eur Radiol* 2019; **29**: 1961–7. doi: <https://doi.org/10.1007/s00330-018-5748-9>
- Vorontsov E, Tang A, Roy D, Pal CJ, Kadoury S. Metastatic liver tumour

- segmentation with a neural network-guided 3D deformable model. *Med Biol Eng Comput* 2017; **55**: 127–39. doi: <https://doi.org/10.1007/s11517-016-1495-8>
24. Weston AD, Korfiatis P, Kline TL, Philbrick KA, Kostandy P, Sakinis T, et al. Automated abdominal segmentation of CT scans for body composition analysis using deep learning. *Radiology* 2019; **290**: 669–79. doi: <https://doi.org/10.1148/radiol.2018181432>
 25. Ronneberger O. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015;.
 26. Cardenas CE, Yang J, Anderson BM, Court LE, Brock KB. Advances in Auto-Segmentation. *Semin Radiat Oncol* 2019; **29**: 185–97. doi: <https://doi.org/10.1016/j.semradonc.2019.02.001>
 27. Balagopal A, Kazemifar S, Nguyen D, Lin M-H, Hannan R, Owangi A, et al. Fully automatic organ segmentation in male pelvic CT images. *Phys Med Biol* 2018; **63**: 245015. doi: <https://doi.org/10.1088/1361-6560/aaf11c>
 28. Chan JW, Kearney V, Haaf S, Wu S, Bogdanov M, Reddick M, et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. *Med Phys* 2019; **46**: 2204–13. doi: <https://doi.org/10.1002/mp.13495>
 29. Dong X et al. Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network. *Radiother Oncol* 2019;.
 30. Dong X, Lei Y, Wang T, Thomas M, Tang L, Curran WJ, et al. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Med Phys* 2019; **46**: 2157–68. doi: <https://doi.org/10.1002/mp.13458>
 31. Fechter T, Adebahr S, Baltas D, Ben Ayed I, Desrosiers C, Dolz J, et al. Esophagus segmentation in CT via 3D fully convolutional neural network and random walk. *Med Phys* 2017; **44**: 6341–52. doi: <https://doi.org/10.1002/mp.12593>
 32. Feng X, Qing K, Tustison NJ, Meyer CH, Chen Q. Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images. *Med Phys* 2019; **46**: 2169–80. doi: <https://doi.org/10.1002/mp.13466>
 33. Fu Y, Mazur TR, Wu X, Liu S, Chang X, Lu Y, et al. A novel MRI segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy. *Med Phys* 2018; **45**: 5129–37. doi: <https://doi.org/10.1002/mp.13221>
 34. Guo Z, Guo N, Gong K, Zhong Shun'an, Li Q. Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network. *Phys Med Biol* 2019; **64**: 205015. doi: <https://doi.org/10.1088/1361-6560/ab440d>
 35. He K et al. Pelvic organ segmentation using distinctive curve guided fully Convolutional networks. *IEEE Trans Med Imaging* 2018;.
 36. Ibragimov B, Toesca D, Chang D, Koong A, Xing L. Combining deep learning with anatomical analysis for segmentation of the portal vein for liver SBRT planning. *Phys Med Biol* 2017; **62**: 8943–58. doi: <https://doi.org/10.1088/1361-6560/aa9262>
 37. Kearney V, Chan JW, Wang T, Perry A, Yom SS, Solberg TD, et al. Attention-enabled 3D boosted convolutional neural networks for semantic CT segmentation using deep supervision. *Phys Med Biol* 2019; **64**: 135001. doi: <https://doi.org/10.1088/1361-6560/ab2818>
 38. Ma Z, Zhou S, Wu X, Zhang H, Yan W, Sun S, et al. Nasopharyngeal carcinoma segmentation based on enhanced convolutional neural networks using multi-modal metric learning. *Phys Med Biol* 2019; **64**: 025005. doi: <https://doi.org/10.1088/1361-6560/aaf5da>
 39. Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. *Med Phys* 2017; **44**: 6377–89. doi: <https://doi.org/10.1002/mp.12602>
 40. Mylonas A, Keall PJ, Booth JT, Shieh C-C, Eade T, Poulsen PR, et al. A deep learning framework for automatic detection of arbitrarily shaped fiducial markers in intrafraction fluoroscopic images. *Med Phys* 2019; **46**: 2286–97. doi: <https://doi.org/10.1002/mp.13519>
 41. van der Heyden B, Wohlfahrt P, Eekers DBP, Richter C, Terhaag K, Troost EGC, et al. Dual-Energy CT for automatic organs-at-risk segmentation in brain-tumor patients using a multi-atlas and deep-learning approach. *Sci Rep* 2019; **9**: 4126. doi: <https://doi.org/10.1038/s41598-019-40584-9>
 42. van Dijk LV, Van den Bosch L, Aljabar P, Peressutti D, Both S, J.H.M. Steenbakkers R, et al. Improving automatic delineation for head and neck organs at risk by deep learning contouring. *Radiotherapy and Oncology* 2019; . doi: <https://doi.org/10.1016/j.radonc.2019.09.022>
 43. Yang J, Veeraraghavan H, Armato SG, Farahani K, Kirby JS, Kalpathy-Kramer J, et al. Autosegmentation for thoracic radiation treatment planning: a grand challenge at AAPM 2017. *Med Phys* 2018; **45**: 4568–81. doi: <https://doi.org/10.1002/mp.13141>
 44. Zaffino P, Pernelle G, Mastmeyer A, Mehrtaash A, Zhang H, Kikinis R, et al. Fully automatic catheter segmentation in MRI with 3D convolutional neural networks: application to MRI-guided gynecologic brachytherapy. *Phys Med Biol* 2019; **64**: 165008. doi: <https://doi.org/10.1088/1361-6560/ab2f47>
 45. Cardenas CE, Anderson BM, Aristophanous M, Yang J, Rhee DJ, McCarroll RE, et al. Auto-delineation of oropharyngeal clinical target volumes using 3D convolutional neural networks. *Phys Med Biol* 2018; **63**: 215026. doi: <https://doi.org/10.1088/1361-6560/aae8a9>
 46. Cardenas CE, McCarroll RE, Court LE, Elgohari BA, Elhalawani H, Fuller CD, et al. Deep learning algorithm for Auto-Delineation of high-risk oropharyngeal clinical target volumes with built-in dice similarity coefficient parameter optimization function. *Int J Radiat Oncol Biol Phys* 2018; **101**: 468–78. doi: <https://doi.org/10.1016/j.ijrobp.2018.01.114>
 47. Lodwick GS, Keats TE, Dorst JP. The coding of roentgen images for computer analysis as applied to lung cancer. *Radiology* 1963; **81**: 185–200. doi: <https://doi.org/10.1148/81.2.185>
 48. Lodwick GS, Haun CL, Smith WE, Keller RF, Robertson ED, et al. Computer diagnosis of primary bone tumors. *Radiology* 1963; **80**: 273–5. doi: <https://doi.org/10.1148/80.2.273>
 49. Doi K. Computer-Aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Graph* 2007; **31**(4-5): 198–211. doi: <https://doi.org/10.1016/j.compmedimag.2007.02.002>
 50. Haus A, Yaffe M. eds. Future of breast imaging. In: *Computer-aided diagnosis, in AAPM/RSNA Categorical Course on the Technical Aspects of Breast Imaging*; 1992 . pp. 257–270.
 51. Giger ML. Computer-Aided diagnosis in radiology. *Acad Radiol* 2002; **9**: 1–3. doi: [https://doi.org/10.1016/S1076-6332\(03\)80289-1](https://doi.org/10.1016/S1076-6332(03)80289-1)
 52. Giger ML, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography. 3. automated detection of nodules in peripheral lung fields. *Med Phys* 1988; **15**: 158–66. doi: <https://doi.org/10.1118/1.596247>
 53. Sahiner B, Chan HP, Petrick N, Wei D, Helvie MA, Adler DD, et al. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE*

- Trans Med Imaging* 1996; **15**: 598–610. doi: <https://doi.org/10.1109/42.538937>
54. El-Naqa I, Yang Y, Galatsanos NP, Nishikawa RM, Wernick MN. A similarity learning approach to content-based image retrieval: application to digital mammography. *IEEE Trans Med Imaging* 2004; **23**: 1233–44. doi: <https://doi.org/10.1109/TMI.2004.834601>
 55. Oh JH, Yang Y, El Naqa I. Adaptive learning for relevance feedback: application to digital mammography. *Med Phys* 2010; **37**: 4432–44. doi: <https://doi.org/10.1118/1.3460839>
 56. Chen W, Giger ML, Newstead GM, Bick U, Jansen SA, Li H, et al. Computerized assessment of breast lesion malignancy using DCE-MRI robustness study on two independent clinical datasets from two manufacturers. *Acad Radiol* 2010; **17**: 822–9. doi: <https://doi.org/10.1016/j.acra.2010.03.007>
 57. Yuan Y, Giger ML, Li H, Bhooshan N, Sennett CA. Multimodality computer-aided breast cancer diagnosis with FFDm and DCE-MRI. *Acad Radiol* 2010; **17**: 1158–67. doi: <https://doi.org/10.1016/j.acra.2010.04.015>
 58. Giger ML, 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* 2008; **35**: 5799–820. doi: <https://doi.org/10.1118/1.3013555>
 59. Antropova N, Abe H, Giger ML. Use of clinical MRI maximum intensity projections for improved breast lesion classification with deep convolutional neural networks. *J. Med. Imag.* 2018; **5**: 1. doi: <https://doi.org/10.1117/1.JMI.5.1.014503>
 60. Antropova N, Huynh BQ, Giger ML. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. *Med Phys* 2017; **44**: 5162–71. doi: <https://doi.org/10.1002/mp.12453>
 61. Guo W, Li H, Zhu Y, Lan L, Yang S, Drukker K, et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. *J Med Imaging* 2015; **2**: 041007. doi: <https://doi.org/10.1117/1.JMI.2.4.041007>
 62. Sheth D, Giger ML. Artificial intelligence in the interpretation of breast cancer on MRI. *J Magn Reson Imaging* 2019; **36**: 1. doi: <https://doi.org/10.1002/jmri.26878>
 63. Bhooshan N, Giger ML, Jansen SA, Li H, Lan L, Newstead GM, et al. Cancerous breast lesions on dynamic contrast-enhanced Mr images: computerized characterization for image-based prognostic markers. *Radiology* 2010; **254**: 680–90. doi: <https://doi.org/10.1148/radiol.09090838>
 64. Hylton NM, Gatsonis CA, Rosen MA, Lehman CD, Newitt DC, Partridge SC, et al. Neoadjuvant chemotherapy for breast cancer: functional tumor volume by MR imaging predicts recurrence-free Survival-Results from the ACRIN 6657/CALGB 150007 I-SPY 1 trial. *Radiology* 2016; **279**: 44–55. doi: <https://doi.org/10.1148/radiol.2015150013>
 65. Drukker K, Li H, Antropova N, Edwards A, Papaioannou J, Giger ML, et al. Most-enhancing tumor volume by MRI radiomics predicts recurrence-free survival "early on" in neoadjuvant treatment of breast cancer. *Cancer Imaging* 2018; **18**: 12. doi: <https://doi.org/10.1186/s40644-018-0145-9>
 66. Wilkie JR, Giger ML, Engh CA, Hopper RH, Martell JM. Radiographic texture analysis in the characterization of trabecular patterns in periprosthetic osteolysis. *Acad Radiol* 2008; **15**: 176–85. doi: <https://doi.org/10.1016/j.acra.2007.08.009>
 67. Zheng J, El Naqa I, Rowold FE, Pilgram TK, Woodard PK, Saffitz JE, et al. Quantitative assessment of coronary artery plaque vulnerability by high-resolution magnetic resonance imaging and computational biomechanics: a pilot study ex vivo. *Magn Reson Med* 2005; **54**: 1360–8. doi: <https://doi.org/10.1002/mrm.20724>
 68. Graffy PM, Sandfort V, Summers RM, Pickhardt PJ. Automated liver fat quantification at Nonenhanced abdominal CT for population-based steatosis assessment. *Radiology* 2019; **293**: 334–42. doi: <https://doi.org/10.1148/radiol.2019190512>
 69. Li H, Zhu Y, Burnside ES, Huang E, Drukker K, Hoadley KA, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ Breast Cancer* 2016; **2**: 16012. doi: <https://doi.org/10.1038/nbjbcancer.2016.12>
 70. Burnside ES, Drukker K, Li H, Bonaccio E, Zuley M, Ganott M, et al. Using computer-extracted image phenotypes from tumors on breast magnetic resonance imaging to predict breast cancer pathologic stage. *Cancer* 2016; **122**: 748–57. doi: <https://doi.org/10.1002/cncr.29791>
 71. Zhu Y, Li H, Guo W, Drukker K, Lan L, Giger ML, et al. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. *Sci Rep* 2015; **5**: 17787. doi: <https://doi.org/10.1038/srep17787>
 72. Meyers PH, Nice CM, Becker HC, Nettleton WJ, Sweeney JW, Meckstroth GR, et al. Automated computer analysis of radiographic images. *Radiology* 1964; **83**: 1029–34. doi: <https://doi.org/10.1148/83.6.1029>
 73. Giger ML, Karssemeijer N, Schnabel JA. Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer. *Annu Rev Biomed Eng* 2013; **15**: 327–57. doi: <https://doi.org/10.1146/annurev-bioeng-071812-152416>
 74. Cao R, Mohammadian Bajgiran A, Afshari Mirak S, Shakeri S, Zhong X, Enzmann D, et al. Joint prostate cancer detection and Gleason score prediction in mp-MRI via FocalNet. *IEEE Trans Med Imaging* 2019; **38**: 2496–506. doi: <https://doi.org/10.1109/TMI.2019.2901928>
 75. Li H, Giger ML, Huynh BQ, Antropova NO, et al. Deep learning in breast cancer risk assessment: evaluation of convolutional neural networks on a clinical dataset of full-field digital mammograms. *J Med Imaging* 2017; **4**: 1. doi: <https://doi.org/10.1117/1.JMI.4.4.041304>
 76. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik* 2019; **29**: 102–27. doi: <https://doi.org/10.1016/j.zemedi.2018.11.002>
 77. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol* 2019; **29**: 6163–71. doi: <https://doi.org/10.1007/s00330-019-06170-3>
 78. Liu S, Wang Y, Yang X, Lei B, Liu L, Li SX, et al. Deep learning in medical ultrasound analysis: a review. *Engineering* 2019; **5**: 261–75. doi: <https://doi.org/10.1016/j.eng.2018.11.020>
 79. Chaudhari AS, Fang Z, Kogan F, Wood J, Stevens KJ, Gibbons EK, et al. Super-resolution musculoskeletal MRI using deep learning. *Magnetic Resonance in Medicine* 2018; **80**: 2139–54. doi: <https://doi.org/10.1002/mrm.27178>
 80. Gong E, Pauly JM, Wintermark M, Zaharchuk G. Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI. *J Magn Reson Imaging* 2018; **48**: 330–40. doi: <https://doi.org/10.1002/jmri.25970>
 81. Lee JH et al. Assessment of diagnostic image quality of computed tomography (CT) images of the lung using deep learning. *SPIE Medical Imaging*. 2018; **10573**.
 82. Carlson JNK, Park JM, Park S-Y, Park JI, Choi Y, Ye S-J, et al. A machine learning approach to the accurate prediction of multi-leaf collimator positional errors. *Phys*

- Med Biol* 2016; **61**: 2514–31. doi: <https://doi.org/10.1088/0031-9155/61/6/2514>
83. Kalet AM, Gennari JH, Ford EC, Phillips MH. Bayesian network models for error detection in radiotherapy plans. *Phys Med Biol* 2015; **60**: 2735–49. doi: <https://doi.org/10.1088/0031-9155/60/7/2735>
 84. Li Q, Chan MF. Predictive time-series modeling using artificial neural networks for linac beam symmetry: an empirical study. *Ann N Y Acad Sci* 2017; **1387**: 84–94. doi: <https://doi.org/10.1111/nyas.13215>
 85. Valdes G, Morin O, Valenciaga Y, Kirby N, Pouliot J, Chuang C, et al. Use of TrueBeam developer mode for imaging QA. *J Appl Clin Med Phys* 2015; **16**: 322–33. doi: <https://doi.org/10.1120/jacmp.v16i4.5363>
 86. Valdes G, Scheuermann R, Hung CY, Olszanski A, Bellerive M, Solberg TD, et al. A mathematical framework for virtual IMRT QA using machine learning. *Med Phys* 2016; **43**: 4323–34. doi: <https://doi.org/10.1118/1.4953835>
 87. Willoughby TR, Starkschall G, Janjan NA, Rosen II, et al. Evaluation and scoring of radiotherapy treatment plans using an artificial neural network. *Int J Radiat Oncol Biol Phys* 1996; **34**: 923–30. doi: [https://doi.org/10.1016/0360-3016\(95\)02120-5](https://doi.org/10.1016/0360-3016(95)02120-5)
 88. Rowbottom CG, Webb S, Oldham M. Beam-orientation customization using an artificial neural network. *Phys Med Biol* 1999; **44**: 2251–62. doi: <https://doi.org/10.1088/0031-9155/44/9/312>
 89. Wells DM, Niederer J. A medical expert system approach using artificial neural networks for standardized treatment planning. *Int J Radiat Oncol Biol Phys* 1998; **41**: 173–82. doi: [https://doi.org/10.1016/S0360-3016\(98\)00035-2](https://doi.org/10.1016/S0360-3016(98)00035-2)
 90. Nguyen D, Long T, Jia X, Lu W, Gu X, Iqbal Z, et al. A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. *Sci Rep* 2019; **9**: 1076. doi: <https://doi.org/10.1038/s41598-018-37741-x>
 91. Willems Set al. *Feasibility of CT-Only 3D Dose Prediction for VMAT Prostate Plans Using Deep Learning*. Cham: Springer International Publishing; 2019.
 92. Liu Z, Fan J, Li M, Yan H, Hu Z, Huang P, et al. A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy. *Med Phys* 2019; **46**: 1972–83. doi: <https://doi.org/10.1002/mp.13490>
 93. Moore KL, Brame RS, Low DA, Mutic S. Experience-Based quality control of clinical intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys* 2011; **81**: 545–51. doi: <https://doi.org/10.1016/j.ijrobp.2010.11.030>
 94. Wu B, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Jacques R, et al. Data-Driven approach to generating achievable Dose–Volume histogram objectives in intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys* 2011; **79**: 1241–7. doi: <https://doi.org/10.1016/j.ijrobp.2010.05.026>
 95. Wu B, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Chuang M, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys* 2009; **36**: 5497–505. doi: <https://doi.org/10.1118/1.3253464>
 96. Kisling K, Zhang L, Simonds H, Fakie N, Yang J, McCarroll R, et al. Fully automatic treatment planning for external-beam radiation therapy of locally advanced cervical cancer: a tool for low-resource clinics. *J Glob Oncol* 2019; **5**: 1–9. doi: <https://doi.org/10.1200/JGO.18.00107>
 97. Shiraishi S, Moore KL. Knowledge-Based prediction of three-dimensional dose distributions for external beam radiotherapy. *Med Phys* 2016; **43**: 378–87. doi: <https://doi.org/10.1118/1.4938583>
 98. Ruan D, Keall P. Online prediction of respiratory motion: multidimensional processing with low-dimensional feature learning. *Phys Med Biol* 2010; **55**: 3011–25. doi: <https://doi.org/10.1088/0031-9155/55/11/002>
 99. Isaksson M, Jalden J, Murphy MJ. On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. *Med Phys* 2005; **32**: 3801–9. doi: <https://doi.org/10.1118/1.2134958>
 100. Guidi G, Maffei N, Meduri B, D'Angelo E, Mistretta GM, Ceroni P, et al. A machine learning tool for re-planning and adaptive RT: a multicenter cohort investigation. *Phys Med* 2016; **32**: 1659–66. doi: <https://doi.org/10.1016/j.ejmp.2016.10.005>
 101. Zhou Z, Folkert M, Cannon N, Iyengar P, Westover K, Zhang Y, et al. Predicting distant failure in early stage NSCLC treated with SBRT using clinical parameters. *Radiother Oncol* 2016; **119**: 501–4. doi: <https://doi.org/10.1016/j.radonc.2016.04.029>
 102. Dawson LA, Biersack M, Lockwood G, Eisbruch A, Lawrence TS, Ten Haken RK, et al. Use of principal component analysis to evaluate the partial organ tolerance of normal tissues to radiation. *Int J Radiat Oncol Biol Phys* 2005; **62**: 829–37. doi: <https://doi.org/10.1016/j.ijrobp.2004.11.013>
 103. Gulliford SL, Webb S, Rowbottom CG, Corne DW, Dearnaley DP. Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate. *Radiother Oncol* 2004; **71**: 3–12. doi: <https://doi.org/10.1016/j.radonc.2003.03.001>
 104. Munley MT, Lo JY, Sibley GS, Bentel GC, Anscher MS, Marks LB, et al. A neural network to predict symptomatic lung injury. *Phys Med Biol* 1999; **44**: 2241–9. doi: <https://doi.org/10.1088/0031-9155/44/9/311>
 105. Su M, Miften M, Whiddon C, Sun X, Light K, Marks L, et al. An artificial neural network for predicting the incidence of radiation pneumonitis. *Med Phys* 2005; **32**: 318–25. doi: <https://doi.org/10.1118/1.1835611>
 106. El Naqa I, Bradley JD, Lindsay PE, Hope AJ, Deasy JO. Predicting radiotherapy outcomes using statistical learning techniques. *Phys Med Biol* 2009; **54**: S9–30. doi: <https://doi.org/10.1088/0031-9155/54/18/S02>
 107. Naqa IE, Deasy JO, Mu Y, Huang E, Hope AJ, Lindsay PE, et al. Datamining approaches for modeling tumor control probability. *Acta Oncol* 2010; **49**: 1363–73. doi: <https://doi.org/10.3109/02841861003649224>
 108. Chen S, Zhou S, Yin F-F, Marks LB, Das SK. Investigation of the support vector machine algorithm to predict lung radiation-induced pneumonitis. *Med Phys* 2007; **34**: 3808–14. doi: <https://doi.org/10.1118/1.2776669>
 109. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. Mr imaging Radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. *Radiology* 2016; **281**: 382–91. doi: <https://doi.org/10.1148/radiol.2016152110>
 110. El Naqa I, Bradley J, Blanco AI, Lindsay PE, Vicic M, Hope A, et al. Multivariable modeling of radiotherapy outcomes, including dose-volume and clinical factors. *Int J Radiat Oncol Biol Phys* 2006; **64**: 1275–86. doi: <https://doi.org/10.1016/j.ijrobp.2005.11.022>
 111. Oberije C, Nalbantov G, Dekker A, Boersma L, Borger J, Reymen B, et al. A prospective study comparing the predictions of doctors versus models for treatment outcome of lung cancer patients: a step toward individualized care and shared decision making. *Radiother Oncol* 2014; **112**: 37–43. doi: <https://doi.org/10.1016/j.radonc.2014.04.012>
 112. Bradley J, Deasy JO, Bentzen S, El-Naqa I. Dosimetric correlates for acute esophagitis in patients treated with radiotherapy for lung carcinoma. *Int J Radiat Oncol Biol Phys*

- 2004; **58**: 1106–13. doi: <https://doi.org/10.1016/j.ijrobp.2003.09.080>
113. Hope AJ, Lindsay PE, El Naqa I, Alaly JR, Vicic M, Bradley JD, et al. Modeling radiation pneumonitis risk with clinical, dosimetric, and spatial parameters. *Int J Radiat Oncol Biol Phys* 2006; **65**: 112–24. doi: <https://doi.org/10.1016/j.ijrobp.2005.11.046>
 114. Wang X, Peng Y, Lu L, et al. ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on Weakly-Supervised classification and localization of common thorax diseases. *arXiv.org* 2017;.
 115. Zhu W, Vang YS, Huang Y, et al. DeepEM: deep 3D ConvNets with em for weakly supervised pulmonary nodule detection. *arXiv.org* 2018;.
 116. Chen W, Giger ML, Bick U. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Acad Radiol* 2006; **13**: 63–72. doi: <https://doi.org/10.1016/j.acra.2005.08.035>
 117. Liu X, Langer DL, Haider MA, Van der Kwast TH, Evans AJ, Wernick MN, et al. Unsupervised segmentation of the prostate using Mr images based on level set with a shape prior. *Conf Proc IEEE Eng Med Biol Soc* 2009; **2009**: 3613–6. doi: <https://doi.org/10.1109/IEMBS.2009.5333519>
 118. Jamieson AR, Giger ML, Drukker K, Li H, Yuan Y, Bhooshan N, et al. Exploring nonlinear feature space dimension reduction and data representation in breast Caxd with Laplacian eigenmaps and t-SNE. *Med Phys* 2010; **37**: 339–51. doi: <https://doi.org/10.1118/1.3267037>
 119. El Naqa I, Irrer J, Ritter TA, DeMarco J, Al-Hallaq H, Booth J, et al. Machine learning for automated quality assurance in radiotherapy: a proof of principle using EpiD data description. *Med Phys* 2019; **46**: 1914–21. doi: <https://doi.org/10.1002/mp.13433>
 120. de Vos BD, Berendsen FF, Viergever MA, Sokooti H, Staring M, Išgum I, et al. A deep learning framework for unsupervised affine and deformable image registration. *Med Image Anal* 2019; **52**: 128–43. doi: <https://doi.org/10.1016/j.media.2018.11.010>
 121. Tseng H-H, Luo Y, Cui S, Chien J-T, Ten Haken RK, El Naqa I, et al. Deep reinforcement learning for automated radiation adaptation in lung cancer. *Med Phys* 2017; **44**: 6690–705. doi: <https://doi.org/10.1002/mp.12625>
 122. Tseng H-H, Luo Y, Ten Haken RK, El Naqa I. The role of machine learning in knowledge-based Response-Adapted radiotherapy. *Front Oncol* 2018; **8**: 266. doi: <https://doi.org/10.3389/fonc.2018.00266>
 123. Shen C, Gonzalez Y, Klages P, Qin N, Jung H, Chen L, et al. Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. *Phys Med Biol* 2019; **64**: 115013. doi: <https://doi.org/10.1088/1361-6560/ab18bf>
 124. Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis; symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science* 1959; **130**: 9–21. doi: <https://doi.org/10.1126/science.130.3366.9>
 125. Shortliffe EH, Buchanan BG. A model of inexact Reasoning in medicine. *Mathematical Biosciences* 1975; **23**(3-4): 351–79. doi: [https://doi.org/10.1016/0025-5564\(75\)90047-4](https://doi.org/10.1016/0025-5564(75)90047-4)
 126. Yu VL, Fagan LM, Wraith SM, Clancey WJ, Scott AC, Hannigan J, et al. Antimicrobial selection by a computer. A blinded evaluation by infectious diseases experts. *JAMA* 1979; **242**: 1279–82.
 127. Price WN, Gerke S, Cohen IG. *Potential liability for physicians using artificial intelligence*. *JAMA* 2019; **322**: 1765. doi: <https://doi.org/10.1001/jama.2019.15064>
 128. Formation. E.G.o.L.a.N.T.N.T. *Liability for Artificial Intelligence and Other Emerging Digital Technologies*; 2019.
 129. Schwartz WB. Medicine and the computer. *New England Journal of Medicine* 1970; **283**: 1257–64. doi: <https://doi.org/10.1056/NEJM197012032832305>
 130. Schwartz WB, Patil RS, Szolovits P. Artificial intelligence in medicine. *New England Journal of Medicine* 1987; **316**: 685–8. doi: <https://doi.org/10.1056/NEJM198703123161109>
 131. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. *N Engl J Med* 2017; **376**: 2507–9. doi: <https://doi.org/10.1056/NEJMp1702071>
 132. Tacchella A, et al. Collaboration between a human group and artificial intelligence can improve prediction of multiple sclerosis course A proof-of-principle study. *F1000Research* 2017; **6**: 2172.
 133. Patel BN, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J, et al. Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *NPJ Digit Med* 2019; **2**: 111. doi: <https://doi.org/10.1038/s41746-019-0189-7>
 134. Wang Jet al. *Learning Credible Models*, in Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. *ACM: London, United Kingdom* 2018;: 2417–26.
 135. Topol EJ. High-Performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44–56. doi: <https://doi.org/10.1038/s41591-018-0300-7>
 136. <https://www.prnewswire.com/news-releases/quantitative-insights-gains-industrys-first-fda-clearance-for-machine-learning-driven-cancer-diagnosis-300495405.html>
 137. Giger ML. Machine learning in medical imaging. *J Am Coll Radiol* 2018; **15**(3 Pt B): 512–20. doi: <https://doi.org/10.1016/j.jacr.2017.12.028>
 138. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer* 2018; **18**: 500–10. doi: <https://doi.org/10.1038/s41568-018-0016-5>
 139. Chan S, Siegel EL. Will machine learning end the viability of radiology as a thriving medical specialty? *Br J Radiol* 2019; **92**: 20180416. doi: <https://doi.org/10.1259/bjr.20180416>
 140. Duong MT, Rauschecker AM, Rudie JD, Chen P-H, Cook TS, Bryan RN, et al. Artificial intelligence for precision education in radiology. *Br J Radiol* 2019; **92**: 20190389. doi: <https://doi.org/10.1259/bjr.20190389>
 141. Langlotz CP, Allen B, Erickson BJ, Kalpathy-Cramer J, Bigelow K, Cook TS, et al. *A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy workshop*. *Radiology* 2019; **291**: 781–91.. doi: <https://doi.org/10.1148/radiol.2019190613>
 142. Oliveira FPM, Tavares JMRS. Medical image registration: a review. *Comput Methods Biomech Biomed Engin* 2014; **17**: 73–93. doi: <https://doi.org/10.1080/10255842.2012.670855>
 143. Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy Committee task group No. 132. *Med Phys* 2017; **44**: e43–76. doi: <https://doi.org/10.1002/mp.12256>
 144. Haskins G, Kruger U, Yan P. Deep learning in medical image registration: a survey. 2019;.
 145. Bi WLet al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin* 2019; **69**: 127–57.
 146. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med* 2018; **98**: 126–46. doi: <https://doi.org/10.1016/j.compbiomed.2018.05.018>

147. Feng M, Valdes G, Dixit N, Solberg TD. Machine learning in radiation oncology: opportunities, requirements, and needs. *Front Oncol* 2018; **8**: 110. doi: <https://doi.org/10.3389/fonc.2018.00110>
100. El Naqa I, Ruan D, Valdes G, Dekker A, McNutt T, Ge Y, et al. Machine learning and modeling: data, validation, communication challenges. *Med Phys* 2018; **45**: e834–40. doi: <https://doi.org/10.1002/mp.12811>
149. Rhee DJ, Cardenas CE, Elhalawani H, McCarroll R, Zhang L, Yang J, et al. Automatic detection of contouring errors using convolutional neural networks. *Med Phys* 2019; **46**: 5086–97. doi: <https://doi.org/10.1002/mp.13814>
150. Guidi G, Maffei N, Vecchi C, Ciarmatori A, Mistretta GM, Gottardi G, et al. A support vector machine tool for adaptive tomotherapy treatments: prediction of head and neck patients criticalities. *Phys Med* 2015; **31**: 442–51. doi: <https://doi.org/10.1016/j.ejmp.2015.04.009>
151. Sonke J-J, Aznar M, Rasch C. Adaptive radiotherapy for anatomical changes. *Semin Radiat Oncol* 2019; **29**: 245–57. doi: <https://doi.org/10.1016/j.semradonc.2019.02.007>
152. E Naqa, I, R Li, M. J Murphy. eds. *Machine Learning in Radiation Oncology: Theory and Application*. Switzerland: Springer International Publishing; 2015 .
153. Kong F-M, Ten Haken RK, Schipper M, Frey KA, Hayman J, Gross M, et al. Effect of Midtreatment PET/CT-Adapted radiation therapy with concurrent chemotherapy in patients with locally advanced non-small-cell lung cancer: a phase 2 clinical trial. *JAMA Oncol* 2017; **3**: 1358–65. doi: <https://doi.org/10.1001/jamaoncol.2017.0982>
154. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging Archive (TCIA): maintaining and operating a public information Repository. *J Digit Imaging* 2013; **26**: 1045–57. doi: <https://doi.org/10.1007/s10278-013-9622-7>
155. Watson DSet al. Clinical applications of machine learning algorithms: beyond the black box. *BMJ* 2019; **1**: 886.
156. Philbrick KA, Yoshida K, Inoue D, Akkus Z, Kline TL, Weston AD, et al. *What does deep learning see? Insights from a classifier trained to predict contrast enhancement phase from CT images*. *American Journal of Roentgenology* 2018; **211**: 1184–93. doi: <https://doi.org/10.2214/AJR.18.20331>
157. Seah JCY, Tang JSN, Kitchen A, Gaillard F, Dixon AF. Chest radiographs in congestive heart failure: visualizing neural network learning. *Radiology* 2019; **290**: 514–22. doi: <https://doi.org/10.1148/radiol.2018180887>
158. Luna JM, Gennatas ED, Ungar LH, Eaton E, Diffenderfer ES, Jensen ST, et al. Building more accurate decision trees with the additive tree. *Proc Natl Acad Sci U S A* 2019; **116**: 19887–93. doi: <https://doi.org/10.1073/pnas.1816748116>
159. Nazmul Haque K, Latif S, Rana R. Disentangled representation learning with information maximizing Autoencoder. 2019;.
160. Maier AK, Syben C, Stimpel B, Würfl T, Hoffmann M, Schebesch F, et al. Learning with known operators reduces maximum training error bounds. *Nat Mach Intell* 2019; **1**: 373–80. doi: <https://doi.org/10.1038/s42256-019-0077-5>
161. Eaton-Rosen Zet al. *Towards Safe Deep Learning: Accurately Quantifying Biomarker Uncertainty in Neural Network Predictions*. Cham: Springer International Publishing; 2018.
162. Sundararajan M., A, Taly, Yan Q. Axiomatic Attribution for deep networks. 2017;arXiv e-prints.
163. Kikinis R, Pieper SD, Vosburgh KG. *3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support, in Intraoperative Imaging and Image-Guided Therapy*. Springer New York: New York; 2014. pp. 277–89.
164. NVIDIA. Clara. 2019. Available from: <https://developer.nvidia.com/clara>.
165. Abadi Met al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. 2016;arXiv e-prints.
166. Paszke Aet al. Automatic differentiation in PyTorch. in *NIPS* 2017;.
167. Deng Jet al. ImageNet: a large-scale hierarchical image database. in *CVPR09* 2009;.
168. Raji ID, Buolamwini J. Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. ACM: Honolulu, HI, USA; 2019. pp. 429–35.
169. Caruana Ret al. Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM: Sydney, NSW, Australia; 2015. pp. 1721–30.
170. Biggio B, Roli F. Wild patterns: ten years after the rise of adversarial machine learning. *Pattern Recognit* 2018; **84**: 317–31p. doi: <https://doi.org/10.1016/j.patcog.2018.07.023>
171. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS, et al. Adversarial attacks on medical machine learning. *Science* 2019; **363**: 1287–9. doi: <https://doi.org/10.1126/science.aaw4399>
172. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115–8. doi: <https://doi.org/10.1038/nature21056>
173. Rajpurkar Pet al. CheXNet: Radiologist-Level pneumonia detection on chest x-rays with deep learning. 2017;arXiv e-prints,.
174. Luo Y, Tseng H-H, Cui S, Wei L, Ten Haken RK, El Naqa I, et al. Balancing accuracy and interpretability of machine learning approaches for radiation treatment outcomes modeling. *BJR|Open* 2019; **1**: 20190021. doi: <https://doi.org/10.1259/bjro.20190021>