Review Article

Machine learning applications in radiation oncology

Matthew Field\(^{a,b}\), Nicholas Hardcastle\(^d,e\), Michael Jameson\(^b,1\), Noel Aherne\(^f,g,\ast\), Lois Holloway\(^a,b,c,e\)

\(^{a}\) South Western Sydney Clinical School, Faculty of Medicine, University of New South Wales, Sydney, NSW, Australia
\(^{b}\) Ingham Institute for Applied Medical Research, Sydney, NSW, Australia
\(^{c}\) Centre for Medical Radiation Physics, University of Wollongong, Wollongong, NSW, Australia
\(^{d}\) Cancer Therapy Centre, Liverpool Hospital, Sydney, NSW, Australia
\(^{e}\) Physical Sciences, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia

Abstract

Machine learning technology has a growing impact on radiation oncology with an increasing presence in research and industry. The prevalence of diverse data including 3D imaging and the 3D radiation dose delivery presents potential for future automation and scope for treatment improvements for cancer patients. Harnessing this potential requires standardization of tools and data, and focused collaboration between fields of expertise. The rapid advancement of radiation oncology treatment technologies presents opportunities for machine learning integration with investments targeted towards data quality, data extraction, software, and engagement with clinical expertise. In this review, we provide an overview of machine learning concepts before reviewing advances in applying machine learning to radiation oncology and integrating these techniques into the radiation oncology workflows. Several key areas are outlined in the radiation oncology workflow where machine learning has been applied and where it can have a significant impact in terms of efficiency, consistency in treatment and overall treatment outcomes. This review highlights that machine learning has key early applications in radiation oncology due to the repetitive nature of many tasks that also currently have human review. Standardized data management of routinely collected imaging and radiation dose data are also highlighted as enabling engagement in research utilizing machine learning and the ability integrate these technologies into clinical workflow to benefit patients. Physicists need to be part of the conversation to facilitate this technical integration.

1. Introduction

Recently the emergence of machine learning techniques in radiation oncology has led to increased interest in the future of automation in the field [1]. There are a number of components in the oncology patient pathway which may be augmented with machine learning techniques in order to improve efficiency, consistency and most importantly patient outcomes [2,3]. For an incoming patient, the initial decision following appropriate cancer staging is whether to treat with a treatment modality such as surgery, chemotherapy, radiation therapy or immunotherapy, either alone or in combination [4]. These decisions are typically supported by clinical trial results and from outcome and/or toxicity modelling [5]. For radiation therapy, subsequent decisions become progressively detailed as to target volumes, in relation to tumour location, nodal regions and organs at risk (OARs) with the aid of segmentation tools. This is followed by treatment management, through optimizing treatment planning and compensating for motion during treatment. The decisions are ultimately linked by review of treatment details once a complete treatment strategy is derived and can lead to adaptive changes throughout the clinical process. Appropriate data management may facilitate continuous feedback and learning based on each new patient that enters the clinic [6,7].

A survey of radiation oncology medical physicists found that most are either using or preparing to use machine learning in the clinic (69%)

---

\(^{\ast}\) Corresponding author at: Department of Radiation Oncology, Mid-North Coast Cancer Institute, 345 Pacific Highway, Coffs Harbour, NSW 2450, Australia. E-mail address: noel.aherne@health.nsw.gov.au (N. Aherne).

https://doi.org/10.1016/j.phro.2021.05.007
Received 6 November 2020; Received in revised form 19 May 2021; Accepted 22 May 2021
Available online 24 June 2021
2405-6316/© 2021 Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
2. Methods and materials

The fundamentals of machine learning are first reviewed for background, this is followed by the review of applications as shown in Fig. 1 which depicts the workflow of planning and treatment processes in radiation therapy. The highlighted areas that provide focus for this review are based on the perceived impact that machine learning can have on patient outcomes in these areas. We further discuss potential future challenges and directions relevant to medical physicists in radiation oncology and medical imaging. Search strategy and terms were targeted closely with key areas identified in the surveys as current or future applications of machine learning [8].

3. Machine learning

The goal of machine learning is to build systems that automatically improve through experience, where this experience is encoded in statistical models derived from past examples of input–output data. These systems seek to automate the generation of outputs given future inputs by modelling the underlying statistical patterns that manifest in the input–output observations with intention of developing decision rules requiring minimal to no human intervention [16,17]. There are substantial benefits in deploying such automation in medicine, including increasing the speed and efficiency of manually laborious tasks [18], standardizing output where subjective or human errors are a significant factor [19], or potentially improving accuracy where outcomes are definitive. There have been increased investments in machine learning technology by commercial and research sectors in recent years due to the highly practical solutions that can be obtained [17,20].

In the context of physics, imaging and radiation oncology, uncertainty exists over appropriate treatment selection, precisely where to apply treatment and how to adapt and manage treatments [21]. Where there is a lack of clarity for clinicians making decisions in these areas, machine learning models may provide a personalised risk analysis of potential scenarios based on prior data. From this perspective, the role of machine learning is largely an assistive one as decision support. Another consideration is more definitive tasks, such as those currently requiring manual interventions in an adaptive therapy setting or involve real-time tracking, which could be partially or fully automated with the aid of machine learning models, either reducing clinician workload or improving treatment techniques.

Fundamentally, the problem needs to be defined in terms of a set of output or outcome values and a set of input or predictor values. If output values are available, the problem is considered supervised learning and the model is optimised or trained to match the input–output relation. This takes the form of a regression problem for continuous (real-valued) outputs and classification if the output stratifies into distinct groups or labels. If the set of output values are constant the problem is a one-class classification or anomaly detection, whereby the current data set corresponds to ‘normal’ samples and any future significant deviation from this data is considered an outlier. If, however, there are no output values available, then the problem is by contrast referred to as unsupervised learning. Without outcomes to guide the model the focus is on describing a latent structure to the data. Success is then gauged by how useful the predicted structure of outputs is in future learning tasks or if a previously unknown structure to the data is indicated, the model can lead to further exploratory analyses or inform hypothesis generation. In practice, significant manual effort or extensive measurements are usually required to obtain output values. However, to mitigate this limitation, a semi-supervised approach can leverage patterns present in unlabelled input data towards a supervised prediction model. This is achieved by treating the unsupervised learning as an initialization of the model parameters before fine-tuning with a smaller set of complete input–output pairs. Recent image recognition models have employed this approach with one or zero-shot learning for scarce categories [22,23].

An important aspect of machine learning is the selection of models among many potential candidates and this is conducted through a further optimization. This entails exploring the space of potential model types or structures as defined by hyperparameters which control model complexity. Model performances are compared by testing the model on data unseen by the model. This is necessary to assess whether the model has been overfit, that is, the model has specialised on the training data.

Fig. 1. Radiation Oncology treatment process with red boxes indicating steps where machine learning has been applied referenced to relevant sections of the article. The green boxes refer to clinical review before and after treatment, the red include the treatment planning process and blue treatment. Section 4; predicting outcomes following radiation therapy, section 5; decision support, section 6; segmentation, section 7; treatment planning, section 8; image guidance and motion management. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
and is incapable of generalising to unseen data in the same problem domain. If minimal data is available for independent testing, a cross-validation procedure can be employed whereby data is partitioned into alternating observed and unobserved groups for sequential evaluation. Alternatively, if the likelihood function for the model is tractable, readily sampled or can be approximated then a Bayesian model selection is an effective approach due to convergence properties [24]. Many models involve the assumption of independent and identically distributed (IID) data and consequently training and validation data sets are often randomly shuffled ensuring this statistical property in model construction. However, this may not always be an appropriate assumption in an applied setting. The data distribution may not be stationary, it may change over time, in oncology due to treatment technique and population demographic changes, or the distribution in another geographical location may not match that of the model development cohort [25,26]. In recent years, to account for this situation attention or memory mechanisms can be embedded into the model development [27]. Fig. 2 provides a graphical description of the process of building and validating machine learning models.

From the health epidemiology perspective, there can be concerns over whether the construction of these automated systems have appropriately considered the detailed circumstances of the patient and whether the models suffer from unforeseen biases in data collection [28,29]. Many health-related statistical models are designed for highly selected cohorts of patients to minimise the impact of competing hazards that can confound predictions. Therefore, an appropriate experimental design considering all of these factors is key to the development of an effective machine learning model.

Recent trends in machine learning have been driven by application challenges such as learning across distributed computer systems (termed federated learning), which is either for computational efficiency or privacy preservation [30,31]. Significant progress is also evident in the field of deep learning [32], where multilayered neural network architectures are able to couple sparsely correlated features in large data sets which is particularly useful for imaging applications [33]. Longer term trends are in formulating probabilistic model structures where the complexity of model is dynamically bounded by the data [34,35], which minimises biases in model selection.

4. Predicting outcomes following radiation therapy

Predicting cancer outcomes accurately for individuals given the heterogeneity in patient, disease and treatment characteristics is of significant benefit to clinicians and patients making treatment decisions, however achieving this is challenging. Recent modelling of clinical practice datasets with machine learning techniques have demonstrated the ability to develop models which can be updated relatively quickly, keeping up with changes in treatment technology and ensuring that the models remain relevant [36]. Many different machine learning approaches have been considered for modelling cancer outcome including support vector machines (SVM), Bayesian networks, artificial neural networks, decision trees and ensemble methods [37–40].

Predicting normal tissue toxicity is arguably as important as tumour control in establishing the most appropriate radiation therapy treatment for patients, given the importance of quality of life for cancer patients. Achieving this is challenging due to the wide number of influencing factors and the variation in radiation dose distributions [41]. Simple dose metrics, such as the volume of healthy tissue receiving 20 Gy (V20) for lung, are commonly used clinically without including other contributing factors. Marks et al [41], in an overview on normal toxicity models in the QANTEC series of papers describe ‘machine learning’ algorithms as ‘sophisticated statistical methods which can support development of a robust multimetric approach for normal tissue toxicity modelling’, although at the time of that publication there were limited models that had been developed, validated and published using machine learning in this area. In a more recent publication from 2015 providing an overview of machine learning approaches for radiation oncology clinicians, Kang et al commented that ‘the overall methodology (of machine learning) has progressively matured, and the field is ready for larger-scale further investigation’ [14]. For normal tissue toxicity modelling support vector machines have been used to predict esophagitis and pneumonitis due to lung cancer radiation therapy; urinary, bladder and rectum complications arising from prostate radiation therapy, and oral mucositis and xerostomia occurring from head and neck radiation therapy [42–44].

Modelling has long benefited from imaging information [45]. The addition of anatomical imaging such as CT and MRI has added accurate volume data and the inclusion of functional imaging from PET or

![Fig. 2. Illustration of the process of building machine learning models flowing initialisation and data on the left to optimised models and applications on the right. A data set with known outcomes is partitioned to iteratively train models guided by performance on validation data or likelihoods until an optimal model is reached based on a pre-defined performance objective. The iterative process of selecting a model structures (defined by hyperparameters) is a search that can be directed by various methods such as a pre-defined sequence or via Bayesian optimisation. Performance is finally assessed on independent test data or external verification data before deploying the model in practice.](image-url)
quantitative MRI allows for values of hypoxia, proliferation and receptor expression [46]. Key predictive values used for radiation oncology outcome models may include patient characteristics (e.g. age, weight), disease and clinical data (e.g. blood test results, stage of disease), image data (e.g. radiomics features), genomics data and treatment approach (e.g. prescribed radiation dose and fractionation). Published models to date have rarely incorporated all of these types of data within a single model, presumably due to two factors; availability of data and the size of the datasets necessary to avoid over-fitting with large numbers of parameters. Interestingly, radiation dose is not always included within radiation oncology cancer outcome models, instead cancer outcome models have been developed for patients who have received the same radiation dose and tumour control probability models which incorporate dose commonly remain separate. In contrast, for normal tissue toxicity models radiation dose is commonly the key predictor e.g. the use of lung V20 within the clinic. This apparent contradiction reflects the need for variation within parameters when modelling, with commonly large variation in normal tissue dose between patients of between 0 Gy and the prescription dose. Variation in dose to the target volume (within the uncertainty of recorded and delivered radiation doses) is small given the International Commission on Radiation Units (ICRU) recommendations of this range being between 95% and 107% of the prescribed dose. Also, most departments adopt very similar treatment guidelines with changes in prescription only being considered within a clinical trial. As uncertainties in radiation dose are reduced and larger datasets are available there will be significant interest in the impact small changes has on cancer outcome models.

Incorporation of radiomics or other imaging information has demonstrated potential in modelling outcome for both cancer outcomes [47–50] as well as normal tissue toxicity [51]. Techniques applied with these data have included logistic regression as well as comparison of multiple machine learning methods [50,52]. Inclusion of both radiomics and clinical factors has been assessed although confirming the benefit of including both factors is still an area of research [48,49]. Genetic factors have also been found to have predictive power in both cancer and normal tissue toxicity [53–56] outcome, and are being considered in machine learning developed models [57].

Given the potential impact of outcome models, validation should be carefully considered and adhere to accepted reporting standards. The TRIPOD statement provides clear guidelines on appropriate validation for multivariable prediction models [58]. The reporting of models with clear statements on the type of validation, with reference to TRIPOD, is increasing [59,60].

The clinical adoption of more complex models, e.g. including more variables than lung V20, will need methods for interpreting the predictions to overcome the perception of models as a ‘black box’. Interpretability has been an issue for complex machine learning models especially with the use of deep learning networks for image classification [61,62], however there are techniques to visualise the focal points of models allowing users to review the factors involved in decision making [63]. Increased data availability and variation is expected to improve the development of outcome models and help to assess both changes in practice and outcome. An important factor in developing clinically useful models is the collection of sufficient follow up information such that outcomes involving progression-free survival, distant-metastases-free survival and overall survival can be assessed where appropriate. Features shown to have the most influence on predicting outcome may change as increased feedback is delivered to clinicians on routine patients and data collection practices improve. This is feasible due to both distributed learning approaches [30,31,64], publicly available data [65] and a trend to recording, storing and collecting larger datasets.

5. Decision support

Decision-making in radiation oncology depends on specialists routinely considering myriad pieces of clinic-pathological information prior to arriving at a therapeutic decision. Ultimately, the foundation for this is their professional experience, with the benefit of seeing thousands of clinical cases as a reference library. This process often involves a complex balance of priorities including feasibility in terms of treatment delivery, patient preferences for balancing the expectations of treatment success and quality of life. However, definitive and broadly applicable decision rules are not always achievable. Simple staging and prognostication systems are limited in the amount of information they can reasonably be expected to capture. Clinical decision support systems (DSS) are designed to provide evidence-based personalised information on the risks of selecting between medical interventions in these situations. Initial expert derived DSSs were built on a clinical knowledge base, termed expert systems, but could be adapted with the aid of electronic records. Statistical models displayed in the form of nomograms trained with electronic health records have been used extensively by clinicians. Despite the widespread adoption of established and validated medical DSSs, in Otto et al. [66] a nomogram was used by 55% of surveyed clinicians, only a small proportion of models have undergone prospective comparison to independent clinician decisions [3], or impact analysis to demonstrate effectiveness in the clinic [67]. A broad systematic review in [68] found that less than 1% of articles on DSS models included results from randomised controlled trials in a real clinical setting, where a definitive result can be demonstrated, but generally DSS usage indicates a positive impact on outcomes. A synthesis of systematic reviews also indicated that the use of DSSs positively impacted patient outcomes [69]. In radiation oncology, DSSs are already in use for prostate cancer in predicting staging, survival, recurrence and metastatic progression [70]. An online breast cancer prognosis of survival and treatment selection DSS on the inclusion of adjuvant systemic treatment [71] is actively updated and independently validated [72]. Engelhardt et al. [73] present a systematic review of DSS in metastatic colorectal cancer and show that the current level of evidence is limited by the extensive validation and reporting of the model calibration details. Decision support has been extended to assess the cost effectiveness of treatments, such as the likely benefit of proton therapy over photon therapy [74]. Furthermore, DSSs are used in reducing the likelihood of unnecessary imaging procedures [75] in emergency departments which reduces costs while minimizing harm to patients.

Beyond merely informing the clinician, some DSSs are designed to include the patient preferences in a shared decision-making (SDM) process. Risks in relation to quality of life versus life expectancy can be personal and an SDM tool is designed to ensure the patient is equally informed on their prospects and choices. Challenges ensue in the design of DSS to eliminate unintended bias and ensuring lay interpretability of the model implications [7]. Assessment of the success of SDM is also dependent on patient satisfaction [76], and surveys indicate that SDM reduces decision regret and anxiety with treatment [77]. Ultimately, personalised treatment involves providing a complete tailored strategy based on the patient and disease characteristics as informed by the latest evidence and priorities of the patient, and conforming to accepted and progressing standards for the construction and delivery of these decision aids is crucial to the patient-clinician confidence in the process [78].

6. Segmentation

Image segmentation, whereby the voxels of tumours and normal tissues are identified on medical images is one of the first steps in the radiation therapy treatment planning process. The Dice Similarity Coefficient (DSC) is a statistical tool first introduced to assess the similarity between samples [1]. In medical image segmentation the DSC represents a measure of overlap between two segmentations and has become the de facto metric used in delineation studies. A DSC above 0.7 is generally considered to represent good agreement, although the use of DSC alone is not recommended as it does not contain any location information [79,80]. Further, DSC can be subject to variation with volume, with
small changes in small volume structures resulting in larger DSC impact. For certain anatomical sites manual segmentation can take hours to complete and is plagued with multiple sources of uncertainty, making automated approaches desirable. Currently, the most common automated techniques in radiation therapy planning systems are thresholding, deformable shape models (DSMs) and atlas based segmentation. Thresholding is useful where there is a clear gradient between a structure and the surrounding tissues (i.e. lung or brain), but otherwise, results can be poor. DSMs are representations of anatomical structures that are constrained by a set of parameters relating to the structure and the image modality. These parameters reduce the dimensionality of the shape space for computational efficiency at the sacrifice of shape reliability. Atlas based segmentation uses a priori knowledge in the form of a reference image(s) that has been previously segmented. The segmented reference image is registered with the target image and the same image transform is applied to the image segments. The resulting image segmentation accuracy is contingent on the registration accuracy, reference image segmentation and image quality. Deep learning, a subfield of machine learning has shown some promise of providing an automated, fast, accurate and precise image segmentation in radiation oncology.

Convolutional neural networks (CNNs), have become prevalent for deep learning image processing applications as the inputs to such networks can be two dimensional (2D) and 3D images where the adjacency of pixels or voxels are encoded in the network. These algorithms have been employed for many image based applications including classification, object localisation and semantic segmentation. In biomedical imaging, CNNs have been applied to classification in the diagnosis of skin cancer, where they out performed experienced dermatologists. Anas et al used a deep convolution neural network approach to delineate the prostate for low dose rate brachytherapy on trans-rectal ultrasound (TRUS) images. Using this approach, they achieved a mean Dice similarity coefficient (DSC) of 0.94, mean surface distance error of 1.05 mm and mean Hausdorff distance error of 3.0 mm. The time to segment a single TRUS image (415x490 pixels) was 0.15 s. DeepMind Technologies, a subsidiary of Alphabet Inc and partners (Mountainview, California, United States of America) have published a protocol for head and neck tumour and normal tissue segmentation using deep learning. Ihrigimov and Xing published the first OAR segmentation paper using CNNs in radiation therapy. The organs they considered included the spinal cord, mandible, parotids, submandibular glands, larynx, pharynx, eye globes and the optic chiasm. The DSC ranged from 0.37 (chiasm) to 0.90 (mandible). Segmentation took approximately 4 min in total. Since then, several other groups have published on deep learning segmentation on head and neck cancer data. A further study by Lustberg et al used a CNN implemented in a research version of a commercial radiation therapy software for thoracic segmentation of OAR. Results required manual correction but resulted in significant reductions in segmentation time. A group at Peking Medical College have published two deep learning segmentation papers where not only OAR but clinical target volumes (CTVs) have been segmented for rectal and nasopharyngeal cancer. Deep learning has also been applied to multi-parametric MRI for localisation and segmentation of rectal cancer with a DSC of 0.68 and 0.70 compared to manual segmentations.

Men et al trained a very deep residual network (DD-ResNet) on a dataset of 800 patients who received breast conserving therapy from 2013 to 2016. They employed cross-validation and reported mean DSCs of 0.91 for both right and left breasts CTVs. CNNs have also been applied to adaptive radiation therapy for lung cancer. Wang et al used a longitudinal imaging dataset comprised of 9 lung cancer patients with 6–7 weekly T2-weighted MRI scans. The network utilised previous weekly scans to delineate current MRIs. The network was trained within 2 h and resulted in DSC of 0.81±0.10. Overall deep learning approaches to image segmentation, are at least as accurate as previous methods, faster to segment clinical images but generally require larger training datasets. One of the advantages of CNNs over atlas based segmentation is that segmentation computation time does not scale linearly with the amount of training data, however training of CNNs is a time consuming process.

7. Treatment planning

Planning and delivery of contemporary radiation oncology treatments is technically challenging. Given the geometric complexity of tumours and their relationship to the surrounding anatomy, computational techniques to determine optimal treatment geometric parameters are required to achieve the desired 3D dose distribution. Beam angle selection has been determined using genetic, evolutionary and cross-entropy algorithms and to determine the optimal beam weights for conformal and intensity modulated treatments. Given the increased use of arc-based intensity modulation these applications may be limited, and may have a role in optimisation of 4π treatment geometry. Further applications of machine learning for treatment plan automation have been in the inverse planning process, where an algorithm was trained to mimic human decision making to adjust treatment plan optimisation parameters during optimisation.

There has been a recent emergence in use of prior information on the effect of patient anatomy on achievable radiation dose distributions to inform current patient planning parameters. This process, so-called ‘knowledge-based planning’, has emerged as a powerful tool to determine the achievable dose distribution. Initial focus including commercial implementation has focused on prediction of dose-volume histograms (DVHs) based on geometric relationships between targets and OARs. Subsequently, DVHs have been used to generate optimisation objectives, reducing the dependence on the user and leading to more consistent and efficient treatment plan optimisation. More recently, prediction of full 3D dose distributions has been achieved through deep learning approaches. In these methods, the input data for training includes 3D masks of target and OAR structures and/or CT image data with spatially associated 3D dose grids from manually created treatment plans. U-Net, DenseNet and Generative Adversarial Networks (GANs) have been implemented to predict the 3D dose distribution given an anatomical input. Patient-specific optimisation objectives have then been derived from these dose distributions to create a deliverable treatment plan using conventional inverse planning. Along with improvements in efficiency and quality of radiation therapy treatment planning, prediction of dose distributions for a given anatomy has the potential to provide rapid assessment of feasibility and quality of radiation therapy for a given patient, in the context of multi-disciplinary decision making.

Machine learning techniques also have a role in evaluation of radiation therapy treatment plan quality, given the complex interplay between doses delivered to tumours and OARs and treatment side effects. Early work evaluated the use of a neural network to score radiation therapy treatment plans based on a learning set of physician-scored plans. Knowledge-based planning has also been used for evaluation of treatment plan quality, as compared with that which has been achieved previously. This has an application for quality assurance in clinical trials, where plan quality is assessed according to whether it has met the constraints. Assessment based on whether the treatment plan is optimal for that particular patient may improve the overall quality of radiation therapy clinical trials. This process is not without challenges; given the relatively short turnaround times for plan submission prior to start of treatment, feedback to submitting institutions should be based on estimates of clinical impact of any deviations in plan quality. Machine learning has also further been applied to prediction of treatment plan delivery fidelity, with the aim to automate measurement of whether a complex treatment plan is deliverable on a linear accelerator with sufficient agreement with the treatment plan.
Deep learning algorithms can generate ‘synthetic’ 3D image data from other image types such as CT scans from MRI or cone-beam CT scans, or predict a 3D dose distribution from an image as demonstrated in Fig. 3 from [117]. Radiation dose calculations rely on the relationship between CT number and electron density or physical material, therefore, the derivation of CT scans from MRI scans with deep learning models is attractive, as it negates the need to acquire a CT scan for treatment planning purposes. With the advent of MRI-linac devices, in which an MRI is merged with a linear accelerator, the MRI is also the only available imaging modality. Fuzzy c-means clustering has been used to classify tissues head and neck and abdominal MRIs, which are then mapped based on attenuation properties into a synthetic CT representing most probably Hounsfield units [136,137]. CNNs and GANs have proven capable of generating high accuracy synthetic CTs from brain, head and neck and liver MRIs for the purposes of photon and proton treatment planning, as shown in Fig. 5, and to provide a reference CT for CT based IGRT [138-142]. Further, synthetic CTs have been generated from MRI using a cycleGAN for the purposes of providing a ‘bridge’ between MR and CT for the purposes of image registration [143].

8. Image guidance and motion management

Image guidance aims to use imaging prior to and during radiation therapy to ensure the target is accurately localised relative to the retreatment beam [144]. Motion management in radiation therapy typically refers to intervention to limit motion and monitoring of tumour or surrogates during radiation delivery to ensure the planned dose is delivered to targets that move due to physiological processes such as respiration [145]. Image guidance and image based motion management require localisation of the target on acquired images. Accurate and fast measurement of motion facilitates interventions such as gating the radiation beam or tracking the moving tumour with the radiation beam. Such interventions increase the probability of hitting the tumour, and may facilitate margin and OAR dose reduction [132,146]. Machine learning plays two key roles in this process – detection of the tumour or surrogate before and during treatment from imaging or otherwise, as well as prediction of motion based on previous measurements.

Detection of the tumour in images acquired during radiation therapy suffers from a lack of contrast between the tumour and surrounding anatomy, as well as the speed with which information is required for adaptation to the tumour position. Kilovoltage x-ray imaging beams are ubiquitous in contemporary linear accelerators and can provide planar imaging prior to treatment, as well as continuous fluoroscopy during radiation therapy treatment. Implanted fiducial markers, visualised on planar kilovoltage imaging, can act as a surrogate for tumour location in the prostate, liver and pancreas. Mylonas et al. used CNNs to improve detection of these fiducials during treatment; the advantages over typical template matching approaches being the ability to detect arbitrarily shaped markers with limited prior information [147].

Further progress has been made in so-called ‘markerless tracking’, to which neural networks, SVMs and Bayesian statistics have been applied with sufficient accuracy and efficiency for pre-treatment setup imaging and real-time tumour tracking for lung, pancreas and prostate cancer [133,148–152]. An example of lung tumour tracking is illustrated in Fig. 4 [150]. Given the increasing multi-modality treatment options for upper abdominal cancers such as pancreas and liver cancer, it is not uncommon for patients to present for radiation therapy with existing ‘markers’ such as surgical clips and radiopaque gel from previous surgery or interventional radiology procedures. The application to machine learning to detection of anatomical and already existing foreign objects has strong potential to provide accurate tumour tracking for a large proportion of patients receiving radiation therapy.

Prediction of respiratory motion is highly advantageous on the sub-second timescale for real-time tumour tracking. Accurate prediction of future positions reduces tracking uncertainty in physical devices such as multileaf collimators and patient positioning systems which arise from system latency between measurement of tumour position and positional adaptation [153]. Machine learning techniques are particularly well suited to this application due to the similarity of future breath characteristics with previously recorded breaths. A large body of work has shown neural network, SVM, manifold learning and kernel density estimation can efficiently predict respiratory motion based on previously measured motion traces [153–159]. Other applications of machine learning involve prediction of motion extent based on tumour size and location in the lungs, automatic diaphragm motion trajectory assessment and incorporation of lung tumour motion into patient setup and prediction of tumour baseline shifts in the short term (approximately 5 s) [160–163].

Cone-beam CTs, the dominant soft-tissue volumetric imaging modality on linear accelerators used for image guidance, suffer from poor image quality due to a range of issues related to increased scatter contribution, beam hardening and motion artefacts [164]. Improvement of CBCT image quality facilitates improved IGRT accuracy through improved soft-tissue contrast, offline and online adaptive radiation therapy treatment based on CBCT imaging, and derivation of quantitative image metrics from CBCTs. Machine learning has been employed in both the projection and image domain to resolve image quality issues. Estimation and correction of photon scatter contribution to CBCT

Fig. 3. The framework of the proposed system, involving images with structures (Slmg), coarse dose maps (CDM), and fine dose maps (FDM). Reprinted from “A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning”, by Chen X, et al., 2019, Medical Physics, 46(1), p56–64.
projections was performed using CNNs using Monte Carlo estimation of primary and scatter contribution as training data [165,166]. Virtual ‘scatter free’ projections are then used for CBCT reconstruction, resulting in improved Hounsfield unit consistency and soft tissue contrast. This approach has a significant advantage over Monte Carlo based correction methods, thereby allowing correction at time of acquisition and reconstruction. Finally, machine learning has also found applications in generation of ventilation images from CT scans [167], deformable image registration [168] and its subsequent quality assurance [169,170].

9. Discussion

The applications of machine learning in radiation oncology are varied in scope and complexity. We have provided a review of applications for this disruptive technology and future challenges. A vision for a fully integrated data management system with continuous feedback between patient outcomes and model input parameters will lead us to improvements in clinical decision making involving the patient, accurate prediction of treatment outcomes and quality of life, efficient and consistent treatment planning and highly targeted delivery. Significant challenges remain, namely highly complex, patient specific disease characteristics and the interplay with systemic and local therapies, this is often coupled with disparate methods of data recording, imaging acquisition methods, all of which impact on accuracy and applicability of any models used. Moreover, limitations in accurate and useful treatment outcome reporting and the prevalence of data sharing or distributed learning currently hinder attempts to mine existing data for model building.

There are several ways in which these challenges can be addressed. The maintenance of a standardised clinical DICOM picture archiving and communication system (PACS) which handles DICOM RT objects would allow for future machine learning uses within the clinic and facilitate external research partnerships. This can be further enhanced by radiation oncology vendors providing batch data anonymisation and data exporting functionality. Physics and imaging specialists can standardise community-wide imaging acquisition protocols such that radiomic or deep learning models can have wider applicability [171,172]. Physics and imaging specialists should focus efforts to standardise ontology in emerging areas in which machine learning may play a role [173,174]. Research staff in radiation oncology can engage with community-wide efforts to participate in distributed or federated learning consortia targeted at building generalizable and useful models across clinics [31,64]. The release of research software as open source will encourage a research community of open collaboration. There should be safe pathways for machine learning models to be used in the clinic to start benefiting patient outcomes [11]. Lastly, radiation oncology departments can allow physics and imaging staff to have training opportunities in computer science and machine learning increasing the capacity to exploit these emerging technologies to benefit their patients [175].

Fig. 4. The markerless tumor tracking trajectories of cases with (a) the lowest and (b) the highest 3D tracking errors. The tracked tumor positions are highlighted in the blue contours. LR: left–right; SI: superior-inferior; AP: anterior-posterior. Reprinted from “A Bayesian approach for three-dimensional markerless tumor tracking using kV imaging during lung radiotherapy”, by Shieh CC, et al., 2017, Phys Med Biol, 2017;62:3065–80. © Institute of Physics and Engineering In Medicine. Reproduced by permission of IOP Publishing. All rights reserved. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: N Hardcastle receives funding through a Varian Medical Systems Collaborative Research Grant for Kidney SABR. This grant includes components of machine learning as applied for treatment planning.

References

[1] Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. Br J Radiol 2019;92(1100):20190001. https://doi.org/10.1259/bjr.20190001.

[2] Lustberg T, van Soest J, Gooding M, Peressutti D, Aljabar P, van der Stoep J, et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. Radiother Oncol 2018;126(2):312-7. https://doi.org/10.1016/j.radonc.2017.11.012.

[3] Omerji C, Nabavant G, Dekker A, Boersma I, Borger J, Reynen 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(1):37-43. https://doi.org/10.1016/j.radonc.2014.04.012.

[4] Markham MJ, Wachtner K, Agarwal N, Bertagnolli MM, Chang SM, Dale W, et al. Clinical Cancer Advances 2020: Annual Report on Progress Against Cancer From the American Society of Clinical Oncology. J Clin Oncol 2020;38(10):1081. https://doi.org/10.1200/JCO.19.03141.

[5] Lambin P, Zindler J, Vanneste B, van de Voorde I, Jacobs M, Eekers D, et al. Modern clinical research: How rapid learning health care and cohort multiple randomized clinical trials complement traditional evidence based medicine. Acta Oncol 2015;54(9):1289-300. https://doi.org/10.3109/0284186X.2015.1062136.

[6] Abnerethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, Neti C, et al. Rapid-learning system for cancer care. J Clin Oncol 2010;28(27):2668-74. https://doi.org/10.1200/JCO.2010.28.5476.

[7] Lambin P, Zindler J, Vanneste BGL, De Voorde LV, Eekers D, Compter I, et al. Decision support systems for personalized and participative radiation oncology. Adv Drug Deliv Rev 2017;109:131-53. https://doi.org/10.1016/j.addr.2016.01.006.

[8] Brouwer CL, Dinkla AM, Vandewinckele L, Crijns W, Claessens M, Verellen D, et al. Machine learning applications in radiation oncology: Current use and needs to support clinical implementation. Phys Imaging Radiat Oncol. 2020;16:144-8. https://doi.org/10.1016/j.phro.2020.11.002.

[9] Batumalai V, Jameson MG, King O, Walker R, Slater C, Dundas K, et al. Cautiously optimistic: a survey of radiation oncology professionals’ perceptions of automation in radiotherapy planning. Tech Innov Patient Support Radiat Oncol. 2020;16:58-64. https://doi.org/10.1016/j.tipsro.2020.10.003.

[10] Sahiner B, Pesesh A, Hadijiki LM, Wang X, Drukker K, Cha KH, et al. Deep learning in medical imaging and radiation therapy. Med Phys 2019;46(1):e1–e6. https://doi.org/10.1002/mp.13260.

[11] Vandewinckele L, Claessens M, Dinkla A, Brouwer C, Crijns W, Verellen D, et al. Overview of artificial intelligence-based applications in radiotherapy: recommendations for implementation and quality assurance. Radiother Oncol 2020;153:55–66. https://doi.org/10.1016/j.radonc.2020.09.008.

[12] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749-62. https://doi.org/10.1038/nrclinonc.2017.141.

[13] Vial A, Stirling D, Field M, Rizs M, Carolan M, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. Transl Cancer Res 2018;7(3):803-16.

[14] Kang J, Schwartz R, Flickinger J, Beriwal S. Machine learning approaches for predicting radiation therapy outcomes: a clinician’s perspective. Int J Radiat Oncol Biol Phys 2015;93(5):1127-35. https://doi.org/10.1016/j.ijrobp.2015.07.2286.

[15] Tseng H-H, Wei L, Cui S, Luo Y, Ten Haken R, El Naqa I. Machine learning and imaging informatics in oncology. Oncology 2020;98(Suppl. 6):344-62. https://doi.org/10.1159/000493575.

[16] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[17] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science 2015;349(6245):255-60. https://doi.org/10.1126/science.aaa8415.

[18] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[19] Gross P, Honnorat N, Varol E, Wallner M, Trappanese DM, Sharp TE, et al. Nuquantus: Machine learning software for the characterization and quantification of cell nuclei in complex immunofluorescent tissue images. Sci Rep 2016;6(1). https://doi.org/10.1038/srep23431.

[20] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[21] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[22] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[23] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[24] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[25] Bishop C. Pattern recognition and machine learning. Springer; 2006.

[26] Bishop C. Pattern recognition and machine learning. Springer; 2006.
[165] Hansen DC, Landry G, Kamp F, Li M, Belka C, Parodi K, et al. ScatterNet: A convolutional neural network for cone-beam CT intensity correction. Med Phys 2018;45(11):4916–26. https://doi.org/10.1002/mp.2018.45.issue-11.10.1002/mp.13175.

[166] Nomura Y, Xu Q, Shirato H, Shimizu S, Xing L. Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network. Med Phys 2019;46(7):3142–55. https://doi.org/10.1002/mp.2019.46.issue-7.10.1002/mp.13583.

[167] Liu Z, Miao J, Huang P, Wang W, Wang X, Zhai Y, et al. A deep learning method for producing ventilation images from 4DCT: First comparison with technegas SPECT ventilation. Med Phys 2020;47(3):1249–57. https://doi.org/10.1002/mp.2019.47.issue-3.10.1002/mp.14004.

[168] Fu Y, Lei Y, Wang T, Higgins K, Bradley JD, Curran WJ, et al. LungRegNet: an unsupervised deformable image registration method for 4D-CT lung. Med Phys 2020;47(4):1763–74. https://doi.org/10.1002/mp.2019.47.issue-4.10.1002/mp.14665.

[169] Galib SM, Lee HK, Guy CL, Riblett MJ, Hugo GD. A fast and scalable method for quality assurance of deformable image registration on lung CT scans using convolutional neural networks. Med Phys 2020;47(1):99–109. https://doi.org/10.1002/mp.2019.47.issue-1.10.1002/mp.13896.

[170] Neylon J, Min Y, Low DA, Santinanan A. A neural network approach for fast, automated quantification of DIR performance. Med Phys 2017;44(8):4126–38. https://doi.org/10.1002/mp.2017.44.issue-8.10.1002/mp.13171.

[171] Rai R, Holloway LC, Brink C, Field M, Christiansen RL, Sun Y, et al. Multicenter evaluation of MRI-based radiomic features: a phantom study. Med Phys 2020;47(7):3654–63. https://doi.org/10.1002/mp.2019.47.issue-7.10.1002/mp.14173.

[172] Zhouannik I, Buxink J, Traverso A, Shi Z, Kalendrais P, Wee L, et al. Learning from scanners: bias reduction and feature correction in radiomics. Clin Transl Radiat Oncol 2019;19:33–8. https://doi.org/10.1016/j.ctro.2019.07.003.

[173] Traverso A, van Soest J, Wee L, Dekker A. The radiation oncology ontology (ROO): publishing linked data in radiation oncology using semantic web and ontology techniques. Med Phys 2018;45(10):e854–62. https://doi.org/10.1002/mp.2018.45.issue-10.1002/mp.12879.

[174] Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295(2):328–38. https://doi.org/10.1148/radiol.2020191145.

[175] Clark CH, Gagliardi G, Heijmen B, Malicki J, Thorwarth D, Verellen D, et al. Adapting training for medical physicists to match future trends in radiation oncology. Phys Imaging Radiat Oncol 2019;11:71–5. https://doi.org/10.1016/j.phro.2019.09.003.