

Received:  
24 December 2018

Revised:  
28 March 2019

Accepted:  
15 May 2019

© 2019 The Authors. Published by the British Institute of Radiology under the terms of the Creative Commons Attribution 4.0 Unported License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Cite this article as:

Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. *Br J Radiol* 2019; **92**: 20190001.

## REVIEW ARTICLE

# Applications and limitations of machine learning in radiation oncology

<sup>1,2</sup>DANIEL JARRETT, <sup>1</sup>ELEANOR STRIDE, <sup>3</sup>KATHERINE VALLIS and <sup>2</sup>MARK J. GOODING

<sup>1</sup>Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, UK

<sup>2</sup>Mirada Medical Ltd, Oxford, UK

<sup>3</sup>Department of Oncology, Oxford Institute for Radiation Oncology, University of Oxford, UK

Address correspondence to: Dr Mark J. Gooding  
E-mail: [mark.gooding@mirada-medical.com](mailto:mark.gooding@mirada-medical.com)

## ABSTRACT

Machine learning approaches to problem-solving are growing rapidly within healthcare, and radiation oncology is no exception. With the burgeoning interest in machine learning comes the significant risk of misaligned expectations as to what it can and cannot accomplish. This paper evaluates the role of machine learning and the problems it solves within the context of current clinical challenges in radiation oncology. The role of learning algorithms within the workflow for external beam radiation therapy are surveyed, considering simulation imaging, multimodal fusion, image segmentation, treatment planning, quality assurance, and treatment delivery and adaptation. For each aspect, the clinical challenges faced, the learning algorithms proposed, and the successes and limitations of various approaches are analyzed. It is observed that machine learning has largely thrived on reproducibly mimicking conventional human-driven solutions with more efficiency and consistency. On the other hand, since algorithms are generally trained using expert opinion as ground truth, machine learning is of limited utility where problems or ground truths are not well-defined, or if suitable measures of correctness are not available. As a result, machines may excel at replicating, automating and standardizing human behaviour on manual chores, meanwhile the conceptual clinical challenges relating to definition, evaluation, and judgement remain in the realm of human intelligence and insight.

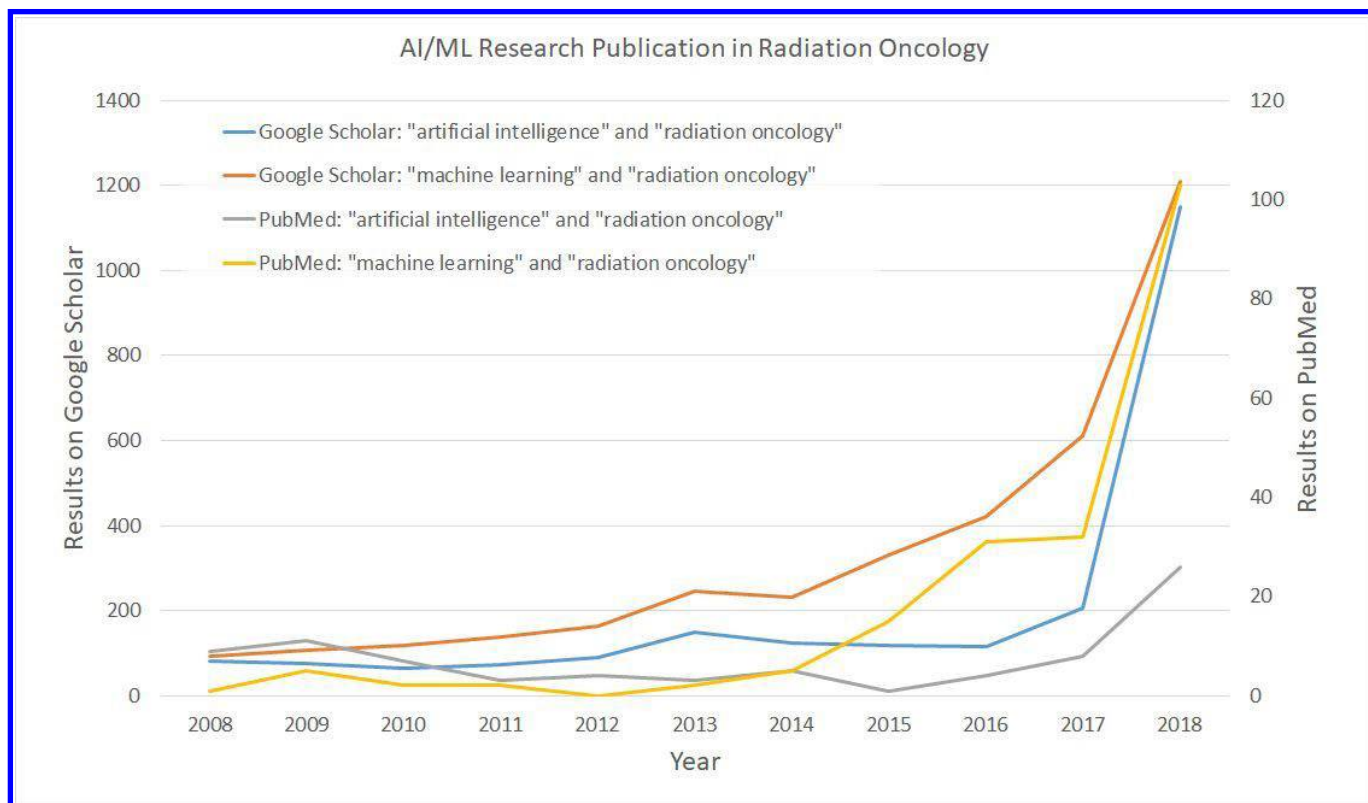
## INTRODUCTION

There is currently considerable enthusiasm for *Artificial Intelligence* (AI) in healthcare, including in radiation oncology. AI is an umbrella term covering all approaches to imitating human intelligence through the use of machines. However, the predominant technical approach currently generating interest in AI for healthcare is best categorized as *machine learning* (ML): the development of data-driven algorithms that learn to mimic human behaviour on the basis of prior example or experience. As a consequence, this paper focuses on the applications of ML in radiation oncology. The burgeoning interest in healthcare is evidenced by the rapid increase in clinical publications in this area, illustrated in [Figure 1](#). Perhaps inevitably, one consequence of such enthusiasm in the evolving field is the risk of overblown expectations. Several reviews have surveyed the use of AI or ML in radiation oncology,<sup>1–5</sup> or more broadly in healthcare.<sup>6,7</sup> In order to calibrate our expectations, this paper first considers what the current clinical challenges are in radiation oncology and then assesses how, or whether ML,

is addressing these challenges. With this clinical orientation, we do not describe algorithmic innovations in detail, and instead refer the interested reader to the more technically focused survey paper of Meyer et al<sup>4</sup> or the book by El Naqa et al.<sup>8</sup>

The clinical workflow for external beam radiation therapy (EBRT) is considered to provide focus on how ML is addressing clinical challenges in radiation oncology, although many of the challenges faced will also apply in other areas of radiation oncology such as brachytherapy. A generalized representation of this workflow is shown in [Figure 2](#); the precise details may vary between institutions. The workflow can be split into three conceptual domains: (1) diagnosis and decision support, (2) treatment planning, and (3) treatment delivery. The process of diagnosis and treatment decision-making can be considered part of the broader oncology workflow, and therefore this review only briefly touches on this aspect; instead, attention is focused on the more concrete applications of ML to treatment planning and delivery.

Figure 1. Number of search results by year for publications relating to “Radiation Oncology” and “Artificial Intelligence” or “Machine Learning”. Results from Google Scholar may represent a wider cross-section of publications than from PubMed. AI, Artificial Intelligence; ML, machine learning.



## MACHINE LEARNING IN PLANNING

### Simulation imaging

Simulation is typically based on either CT imaging, MRI, or both—in which case image fusion is used to combine the information. CT imaging has been the predominant approach due to its cost-effectiveness, speed of acquisition, high spatial resolution, and availability of machines.

### CT simulation

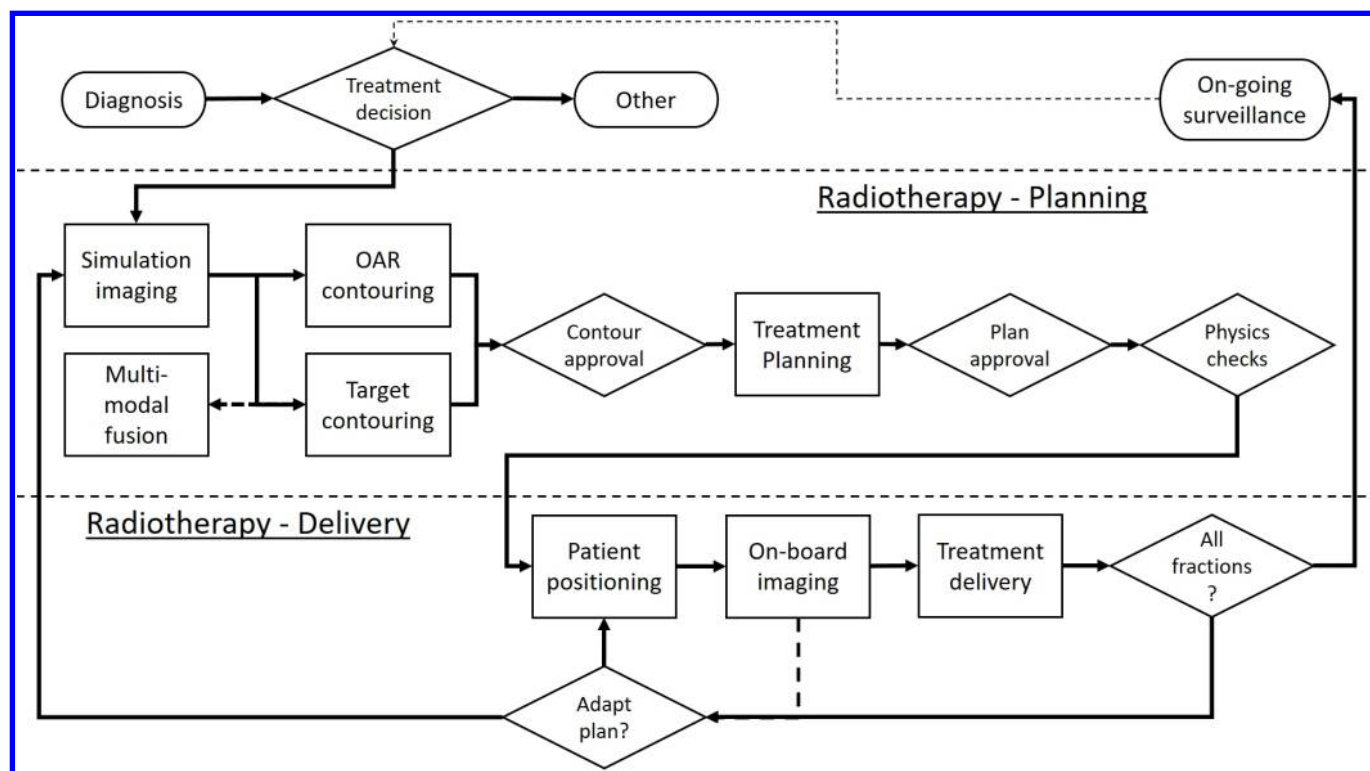
CT simulation imaging has been standard practice in external beam radiation therapy since the 1980s.<sup>9</sup> It might be assumed therefore that many of the challenges have been overcome. However, a number of areas for clinical improvement remain. For example; a CT image only provides a single snapshot of the tumour, introducing location uncertainty right at the beginning of the planning process<sup>10</sup>; increasing the number and types of images used for planning, such as four-dimensional CT, and treatment monitoring may introduce additional risks to the patient with respect to radiation dose.<sup>11</sup> Furthermore, uncertainty also stems from gauging the extent of the tumour<sup>10</sup>; this is discussed in more detail in the section on contouring.

Broadly speaking, there is a trade-off between dose reduction and image quality enhancement, and the achievable balance depends on the specific reconstruction algorithm. Improvements in image reconstruction aim to push the frontier of this trade-off, so that similar quality images can be acquired with a lower dose, or better quality images for the same dose.

Research on ML for reconstruction follows on from improvements in iterative reconstruction,<sup>12</sup> and many incorporate the use of neural networks (NNs). The majority of work has been focused on low-dose CT imaging for diagnosis, where dose reduction is the goal, rather than radiotherapy planning or delivery, where improving image quality is of greater importance, with the 2016 AAPM Low-dose CT grand challenge being a driver for several publications.<sup>13–15</sup> ML techniques can be broadly classified by how the learning algorithm fits into the reconstruction pipeline. *Image-to-image* ML approaches aim to denoise the CT image *after* reconstruction has already taken place.<sup>13,15,16</sup> Other approaches employ ML *during* the reconstruction process, to learn the prior information to guide iterative reconstruction from the raw imaging data,<sup>14,17,18</sup> a two-dimensional array containing raw X-ray projections through the patient known as a sinogram. In addition, some studies specifically focus on the challenge of reducing metal artefacts in CT reconstruction; similarly, these can be categorized by whether artefact reduction occurs post-reconstruction,<sup>19</sup> or directly by improving the sinogram.<sup>20</sup>

With their focus on diagnostic imaging, all of these studies utilize the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) to assess visual image quality based on comparison with some putative “ground truth” image such as a higher dose CT image with a large number of views. On these counts, technical innovations often produce improvements over standard filtered back-projection or iterative

Figure 2. Schematic overview of the external beam radiation therapy workflow. Conceptually, we split this into (1) diagnosis and decision support, (2) treatment planning, and (3) treatment delivery. OAR, organ at risk.



reconstruction algorithms. However, these are seldom evaluated within the context of the radiotherapy workflow. For instance, while Gjestebj et al<sup>19</sup> motivate their work with the need for high image quality in proton therapy and show a 26% increase in peak signal-to-noise ratio for metal artefact reduction compared to conventional approaches, no impact on the derived electron density, or the corresponding radiation dosimetry, of these methods have been considered.

### MRI simulation

While CT imaging is ubiquitous due to its low cost and high speed, an inherent limitation of CT imaging is the uncertainty of tumour location due to poor soft-tissue contrast. In some cases, this makes MRI more attractive for simulation. While MRI is increasingly used for planning and treatment monitoring, progress has focused on how methods developed for CT can be adapted to work with MRI. However, the use of MRI for simulation introduces a new obstacle: the electron density information is unavailable for calculation of the treatment plan—a challenge shared with the task of attenuation correction in PET/MR hybrid-imaging.<sup>21</sup> Conventional algorithmic approaches have been developed to synthesize CT images from MRI,<sup>22</sup> and more recently ML approaches have been considered. These learn from training sets of MR and CT images, and broadly fall into two categories: direct *encoder-decoder* networks that map an MRI to pseudo CT images,<sup>23,24</sup> and *generative adversarial networks* (GANs) that pit image “generators” against “discriminators” to ensure that synthesised images are as realistic as possible.<sup>21,25,26</sup> While the majority of studies evaluate methods based on the similarity of pseudo CT images to known CT images, Maspero

et al<sup>25</sup> demonstrated that MR-based GAN approaches can achieve dosimetric deviations of as little as 0.5% under realistic conditions—a level of accuracy comparable to tissue classification-based approaches.<sup>27</sup>

### Multimodal fusion

Image registration is predominantly used in radiotherapy to align supporting imaging to the planning CT, such that additional information in the former can be used for target contouring.<sup>28</sup> Typically, finding the best transformation between source and target images is posed as an optimization problem on a case-by-case basis. Specifically, they aim to maximize some measure of *similarity* between the images, subject to pre-specified physical constraints. However, conventional methods are limited clinically on two fronts. First, there is a clinical desire to accurately and automatically quantify spatial error throughout the image to inform treatment margins,<sup>29,30</sup> which is particularly challenging in homogenous regions.<sup>31</sup> Second, a significant conceptual challenge is the fact that image content may change between scans (*i.e.* before/after surgery or tumour growth), violating the assumption of a simple voxel-wise correspondence. While a few example of ML have been applied in the context of radiotherapy to date,<sup>32,33</sup> further studies have investigated this approach for other medical image registration tasks.

ML approaches to the registration problem have primarily focused on solving narrower technical issues related to the optimization problem: defining more sophisticated measures

of similarity between images, as well as accelerating the computational procedure. Generally, it is difficult for handcrafted statistical similarity metrics to be simultaneously sophisticated enough to capture complex correlations between modalities, while remaining general enough to be robust to different tissues and noise levels without significant fine-tuning. One solution is to *learn* an appropriate similarity metric from training data. Building on earlier ML approaches,<sup>34,35</sup> NNs have been trained as binary classifiers to learn patchwise correspondences between images.<sup>36,37</sup> In technical studies, these learned measures of similarity were found to outperform conventional metrics such as mutual information and cross-correlation—both in terms of the correctness of the metric itself (in terms of prediction errors),<sup>37</sup> as well as its effectiveness when used within the image registration workflow (in terms of overlap measures).<sup>36</sup> Similarly, encoder-decoder networks that learn hierarchical features as a pre-processing step have been shown to be advantageous both in computational efficiency and robustness to noise.<sup>33</sup> Alternatively, learning algorithms can also be employed to first construct shared latent representations of anatomical structures across different modalities,<sup>38</sup> or even to perform cross-modal image synthesis,<sup>26</sup> effectively reducing the multimodal problem to a more tractable mono-modal one, with promising improvements in performance.<sup>32</sup> Such promising technical advances still require further validation in a clinical setting.

ML strategies have also focused on accelerating the optimisation process. A recent approach attempts to predict deformation parameters *directly* from appearances of image patches, with a 35x speed up being reported for deformable registration of MRI brain imaging.<sup>39</sup> In fact, learning the transformation function directly obviates the need for iterative optimization—the procedure simply becomes a much more efficient matter of applying a function. Furthermore, the registration task has been cast in terms of *reinforcement learning* (RL), where an *artificial agent* is trained to choose sequences of actions that improve image alignment.<sup>40</sup> It has also been formulated within a GAN framework, where the generator estimates transformation parameters while the discriminator evaluates the quality of those predictions.<sup>41</sup>

Currently, for the purpose of clinical use, ML methods have been trained and evaluated on the basis of historical registrations within clinical data sets. “Ground truth” registrations are therefore estimates of the underlying transformations, and depend on the specific protocols and algorithms used to produce them in the first place. An alternative approach to training would be to use digital phantoms, as has been proposed for commissioning of clinical systems.<sup>42</sup> Such an approach enables the true registration to be known but would still require the registration used to be clinically meaningful to train a registration method to generate clinically plausible registrations. Thus, while learning algorithms can be trained to imitate registration patterns from examples, they are limited by the clinical correctness and realism of those examples. Importantly, the open clinical challenges related to quantifying spatial errors or accounting for anatomical changes have not been addressed to date. Despite these limitations, there

are considerable benefits to be had both in time- and memory-efficiency gains by using ML for image registration.

### Contouring

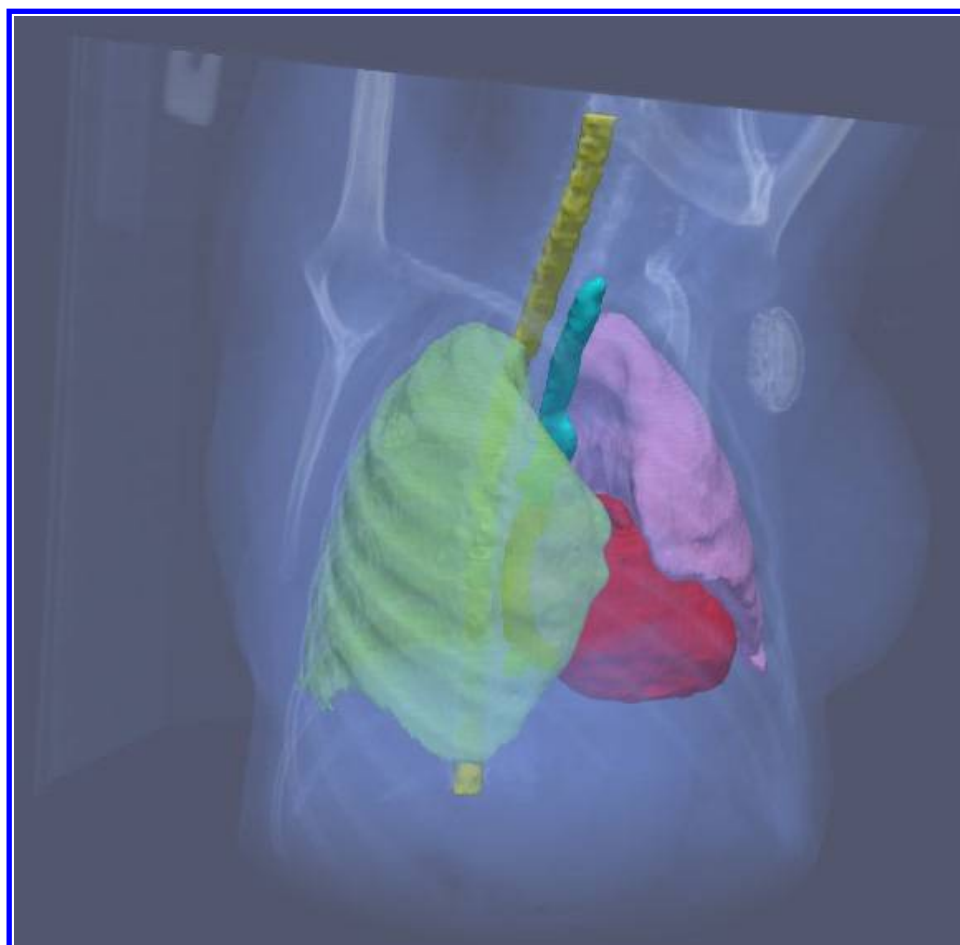
In standard contouring workflows, the segmentation of tumour regions and normal tissue is manually performed by clinical staff, normally on a slice-by-slice basis. As a result, the procedure is lengthy and subject to a high degree of interobserver variability—constituting one of the largest sources of uncertainty in treatment planning.<sup>43,44 45</sup> The earliest attempts at automatic segmentation—such as edge- and region-based methods—relied purely on the informational content of each image in question. Later methods began to incorporate prior knowledge, such as relative anatomical locations or expected size variations of organs, into the process. The approach most commonly found in clinical use is *Atlas-based* segmentation.<sup>46,47</sup> First, registration techniques are used to match target images to one or more selected reference images. Then, ground-truth segmentations on the references are mapped onto the target. However, atlas-based methods are highly sensitive to the atlas-selection strategy,<sup>48</sup> as well as the robustness of the—often time-consuming—registration itself.

The ML approach is to learn the structure labelling of each image voxel directly, more flexibly incorporating prior knowledge in the form of parameterised models. Successful techniques include the use of statistical and decision-learning classifiers,<sup>49,50</sup> and more recently deep learning. For segmentation of organs at risk (OARs), *convolutional neural networks* (CNNs) have shown competitive performance in the context of thoracic cancer,<sup>51</sup> head-and-neck cancer,<sup>52</sup> prostate cancer,<sup>53</sup> as well as more challenging organs such as the oesophagus.<sup>54</sup> Although to date most research has focus on quantitative comparison against “ground truth”, Lustberg et al show an average time saving of 61% compared to existing clinical practice, and 22% compared to the use of atlas-based contouring.<sup>51</sup> Figure 3 shows an example of unedited segmentation of OARs generated by a CNN. In addition, models have been shown to generalize successfully across imaging modalities without specific customisation or large volumes of data.<sup>55</sup> The task of tumour volume segmentation is generally more challenging due to the varied shape, size, appearance, and localisation of tumours, as well as the lack of clear boundaries, rendering the process more reliant on oncologist knowledge and experience. Nevertheless, various deep learning architectures have shown considerable progress, for instance in brain cancer,<sup>56</sup> breast cancer,<sup>57</sup> oropharyngeal cancer,<sup>58</sup> and rectal cancer.<sup>59</sup>

Contouring is inherently subjective,<sup>44</sup> and highly variable as to what a “correct” result should look like. Learning algorithms are simply trained to maximize measures of similarity between their output and the examples given to them. So, while they are increasingly adept at mimicking human-drawn contours, they are constrained—as is any data-driven method—by the nature and quality of training examples. Machines can be no more “correct” than the human input taken as clinical ground truth<sup>60</sup>; until more concrete consensus definitions for object boundaries can be specified, machine “accuracy” is only meaningful within



Figure 3. Example of unedited segmentation of OARs. The use of automatic OAR segmentation based on deep learning methods has demonstrated time savings in the clinical workflow. OAR, organ at risk.



the context of individuals and institutional protocols. Ultimately, the fundamental challenge in generalizability lies in the standardization of existing clinical practices, without which it is difficult for any generic algorithm to perform equally well for different clinicians and centres. Nevertheless, the use of ML may assist in driving this standardization, enabling different institutions to start from a common reference and encouraging conformity to a standard.

Deep learning models have largely demonstrated superior quantitative performance relative to state-of-the-art algorithms for both target and OAR contouring,<sup>52,56</sup> significantly decreasing the manual editing required for contours to be acceptable for use.<sup>51</sup> Therefore, deep learning contouring approaches offers considerable potential to enable the automation and standardisation of contouring in clinical practice.

#### Treatment planning

Radiation treatment planning is considered a problem of optimization—*i.e.* determining the parameters for positioning, fractionation, distribution, and other machine settings to best manage a patient's cancer. However, the notion of *optimality* is often far from concrete in practice: in the presence of competing objectives for target coverage and organ sparing,

the planning process requires multiple *subjective* trade-offs on the part of the oncologist and planner. As a result, conventional human-driven iterative workflows are often time-consuming and labour-intensive, with a high degree of variability.

Several planning strategies have been developed to address these issues. *Multicriteria optimization* generates a broad range of plans on the Pareto surface enabling the user to explore the impact of the trade-offs to be made in real-time.<sup>61,62</sup> *Autoplanning* approaches mimic the process of iterative evaluations and adjustments made by experienced operators. They automatically identify regions of interest, appropriately introducing additional objectives and priorities within the traditional optimization procedure.<sup>63,64</sup> *Knowledge-based planning* methods develop fixed associations between geometric and dosimetric parameters from a selection of previous plans. Rather than starting a new for each patient, this method leverages prior experience in predicting feasible dose-volume histograms<sup>65</sup> or voxel dose distributions,<sup>66</sup> navigating dosimetric trade-offs,<sup>67</sup> or initializing optimization parameters to serve as individualized starting points for fine-tuning.<sup>68</sup> Plans generated in this way have generally demonstrated similar coverage of target volume and superior OAR sparing compared to manually developed plans, as well as reduced planning time and interoperator variability in plan quality.<sup>69,70</sup>

Deep learning methods have recently been investigated for use in planning automation, predominantly focused on predicting a single solution as an input to knowledge-based or auto-planning or as a guide for the clinician/dosimetrist to predict achievable doses,<sup>71</sup> rather than developing improved MCO methods. As a more dynamic and flexible incarnation of knowledge-based planning, NNs can be trained on prior plans to predict voxelwise dose values on the basis of contours and anatomy.<sup>72,73</sup> In lieu of handcrafted features, CNNs are effective in automatic feature extraction, allowing models to learn contour-to-dose mappings directly.<sup>71,74,75</sup> In the spirit of iterative autoplanning, GAN architectures have also been proposed to mimic the roles of dosimetrists (generators) and oncologists (discriminators).<sup>76</sup> So far, deep learning methods have managed to produce highly accurate predictions of dose distributions,<sup>72</sup> resulting in plans of equivalent or superior quality relative to historical plans—both in terms of meeting clinical satisfaction criteria, where GAN-based KBP outperformed the original clinical plans in meeting acceptance criteria in 75% of cases compared 73%,<sup>76</sup> and in qualitative evaluation by clinicians where KBP using NN was preferred to the clinical plan in 81% of cases.<sup>73</sup>

While learning-based models are adept at incorporating human knowledge in planning and quality control, they are—once again—only as good as their training data. Importantly, there is a difference between generating an *optimal* plan, vs one that is simply *preferred* relative to its manual counterpart. Existing strategies have generally aimed—with much success—at the latter, broadly producing clinically acceptable and preferable results in head-to-head clinical evaluations. Although, subjective decisions requiring human judgement remain in determining the “best” compromise when not all constraints can be met or understanding how to vary the plan appropriately from “standard” plans for medically complicated cases, ML methods offer the potential to support this decision nevertheless, by accelerating the planning process and proposing high quality plans automatically.

### Approval & QA

Key components for plan evaluation and quality assurance include the delineation of target volumes and OARs, field arrangement, collimation, target coverage, dose homogeneity, and normal tissue sparing. Due to the increasing sophistication of radiation techniques, the potential scope of assessment has become correspondingly complex. Conventional workflows involve manual reviews with checklists and treatment charts—a process that becomes challenging with limited staff availability. Despite rigorous human review procedures, errors can and do slip through.<sup>77</sup> Learning algorithms have the potential to automate elements of the process, model complexities without exhaustive rules and definitions, as well as enabling individualized assessments for more intelligent allocation of planning resources.

In terms of contour quality, ML algorithms can be trained to recognize and label anatomical structures within segmented images, using *confidence measures* to detect potential errors along the way.<sup>78</sup> To validate physical parameters—such as beam configurations, monitor units, energies, and

fractions—anomaly detection methods with statistical classifiers and clustering algorithms offer the ability to flag potential outliers for human intervention.<sup>79,80</sup> As for dose distributions, knowledge-based approaches to predicting achievable doses have been effective for benchmarking the quality of planned dose distributions.<sup>81–83</sup> In the same way as knowledge-based planning, libraries of historical plans are used to identify specific outliers for replanning, enabling a more targeted approach to improving suboptimal plans. Furthermore, ML algorithms can potentially be used to enhance the quality of delivery. For instance, it has been demonstrated that by predicting differentials between planned and delivered treatments variables, discrepancies can be incorporated into dose calculations, increasing  $\gamma$  passing rates during QA delivery.<sup>84</sup> In addition, accurate a priori predictions of  $\gamma$  passing rates can be made on the basis of fluence maps and other plan details through probabilistic and deep learning techniques.<sup>85,86</sup> However caveats apply: an algorithm cannot transcend its data and objectives. For instance, while ML may predict  $\gamma$  passing rates with great accuracy, it offers no help if passing rates do not actually correlate meaningfully with clinical safety.<sup>87</sup>

Importantly, great care must be taken to ensure that planning and validation algorithms are sufficiently independent—for instance, in the case of knowledge-based planning methods being used for quality assurance as well as autoplanning. After all, no matter how sophisticated an algorithm is, it cannot detect its own errors. Nevertheless, compared to generic checklist-based approaches, ML has the clear potential to augment the scope and efficiency of QA procedures via benchmarks and criteria tailored to specific patients, physicians, and institutions.

### Machine learning in treatment delivery

During the course of radiation therapy, treatment may need to be adjusted to ensure appropriate delivery of the plan. Adjustments may stem from both online factors such as the pre-treatment positioning of the patient, as well as longer term factors related to anatomical changes and treatment response. The technical challenges of aligning onboard imaging to the planning CT are similar to those discussed in the registration section—with the additional limitation that only restricted rigid adjustment can be performed, while the process of replanning requires additional imaging, contouring and treatment planning. An additional wrinkle, however, lies in determining which patients will most benefit from replanning—a decision that greatly influences the allocation of imaging and planning resources.

In terms of dose calculation for daily treatment adaptation, recent work has demonstrated promising gains in both efficiency and output quality by using NNs for directly mapping cone beam CT to planning CT images,<sup>88,89</sup> where the use of GANs to correct CBCT values reduce mean intensity absolute errors compared to an original CT from 158HU to 57HU.<sup>89</sup> During treatment delivery, the precision of irradiation may be compromised by respiratory motion. Learning techniques are particularly suited to capturing the heterogeneous variations in breathing patterns without explicit biomechanical models.

Basic NNs have been shown to be effective in inferring and predicting tumour location from measurements of respiratory motion, thereby enabling adaptive beam realignment to occur in real time with minimal latency.<sup>90</sup> In addition, learning frameworks combining individual predictors have been shown to significantly improve performance beyond the best existing methods.<sup>91</sup>

If the decision for plan adaptation is made during the course of treatment, knowledge-based methods can be adapted for automatic replanning, by initializing the planning procedure with the existing plan for the same patient.<sup>92</sup> Recently, *reinforcement learning* (RL) approaches to automated dose-fractionation adaptation have been developed. In this framework, an artificial agent navigates a model of the radiotherapy environment, selecting sequences of planning decisions to maximise measures of reward—such as tumour control probability and normal tissue complication probability.<sup>93</sup> More generally, RL algorithms can be flexibly trained to first *learn* representations of the radiotherapy state space, and then to *optimize* measures of treatment outcome on the basis of diagnostic, dosimetric, biological, and genetic features.<sup>94</sup> By incorporating prior knowledge and clinical protocols via appropriate reward functions, as well as augmenting historical plans with adversarially generated synthetic data, this technique has been shown to achieve results comparable to those chosen by clinicians.<sup>95</sup>

ML has also been developed to identify candidate patients for replanning intervention. Based on anatomical and dosimetric variations (such as those caused by tumour shrinkage, organ movement, or changes in setup), classifiers and clustering algorithms have been developed to automatically predict patients who would most benefit from updated plans during fractionated radiotherapy treatment.<sup>96,97</sup> However, since learning proceeds from data on historical patients, plans, and their adaptations, the limitation is—once again—that the algorithm is simply learning to mimic past prescriptions and protocols, instead of determining the truly ideal time for replanning intervention on the basis of outcomes.

Despite the caveats, the additional automation and efficiency that ML could bring to the radiotherapy workflow in registration, contouring and planning means that the cost associated with replanning is reduced. Thus, ML may allow the decision to replanning to be made with a greater focus on the clinical benefit, than on the cost of doing so.

## DISCUSSION

For an ML algorithm to be effective, three ingredients are required: (1) a well-defined problem, (2) a well-defined ground truth for which there is sufficient data, and (3) a quantitative measure with which the algorithm is trained and evaluated. In the context of these requirements,

**Table 1** summarizes the clinical challenges in radiotherapy, as well as the areas that ML research has focused on. The three requirements pose challenges to developing ML solutions at each step of the radiotherapy workflow and give insight into why

some problems are readily addressed by ML, while others remain open problems.

First, is the problem well-defined? In other words, is there a *correct* answer in theory? For many clinical problems the answer is yes, but this is not always the case. In treatment planning, for instance, there is no clearly defined criterion for correctness. While a “good” plan can be conceived of as one on the Pareto optimal surface, there are no established guidelines regarding the choice of a single “best” point on that surface—especially since the “best” clinical outcome is analogously poorly defined.

Second, if there is a correct answer, do we have practical *access* to this ground truth? Once the problem is clearly defined, the next requirement is for training data. While the radiotherapy workflow can provide a lot of data, not all of it has a well-defined ground truth. For example, in image registration there is ambiguity regarding the correct displacement field within homogeneous image regions.

Finally, can we define a concrete measure to *optimize*? ML techniques are often trained by optimizing one or more performance metrics. This may pose a problem, since an appropriate measure is not always definable. For example, an ideal registration approach, whether through AI or traditional algorithms, would minimize the real clinical/anatomical error. However, this error is largely unknown (although it can be approximated for specific locations if they are marked-up). As a result, optimization is performed on surrogate measures, which may not accurately reflect clinical/anatomical correctness.

**Table 1** summarizes current research in applications of ML to the radiotherapy pathway. As a consequence of the challenges faced in defining the problem, the ground truth, or the measure. ML research in radiotherapy currently does not address some of the greater clinical challenges, since the challenges are not necessarily always in the technical details of data analysis, but in defining either the question, the type of answer, or the way to measure how well we are doing. What is the correct registration for dose summation where tissues have changed? What is the clinical extent of the tumour on this CT scan? Which plan will deliver the best clinical outcome for this patient? How should we go about measuring this? These are problems that cannot be addressed by ML, but rather require the real intelligence and insight of humans.

Nevertheless, ML has seen considerable success in providing fast and parallelizable technical implementations to automate conventional workflows. While ML is not able to provide better definitions of the problem, application of ML in radiotherapy can bring efficiency and consistency to any solution to a problem. Increasing efficiency offers the opportunity to free time to consider the open clinical challenges, while increasing consistency has the potential to allow better assessment of the impact of intentional changes to treatment practice. By mimicking the current state-of-the-art reliably, ML could

Table 1. Summary of current ML research focus in the radiotherapy pathway

Clinical application	Clinical Need	Current ML focus	Well-defined problem?	Well-defined ground truth?	Quantitative measure of correctness?
CT simulation	Image reconstruction quality / dose reduction	Image reconstruction quality / dose reduction	Yes	No	No
MRI simulation	Pseudo CT creation	Pseudo CT creation	Yes	No	Yes
Image fusion	<ul style="list-style-type: none"> <li>Estimate spatial uncertainty</li> <li>Accommodation of anatomical changes</li> </ul>	<ul style="list-style-type: none"> <li>Registration efficiency</li> <li>Appropriate similarity metric</li> </ul>	No - Depends on use-case	No	No
Contouring	<ul style="list-style-type: none"> <li>OAR/Target Contouring efficiency</li> <li>OAR/Target consistency</li> <li>Target contouring accuracy</li> </ul>	<ul style="list-style-type: none"> <li>OAR/Target Contouring efficiency</li> <li>OAR/Target consistency</li> </ul>	Yes	No - Subjective clinical contours used	Yes
Treatment planning	<ul style="list-style-type: none"> <li>Planning efficiency</li> <li>Plan consistency</li> <li>Determining the plan to deliver the best clinical outcome</li> </ul>	<ul style="list-style-type: none"> <li>Planning efficiency</li> <li>Plan consistency</li> </ul>	No - Depends on clinical satisfaction criteria	No - Subjective treatment plans used	No
QA	<ul style="list-style-type: none"> <li>Efficiency and automation</li> <li>Identification of clinically meaningful errors</li> </ul>	<ul style="list-style-type: none"> <li>Efficiency and automation</li> </ul>	n/a	n/a	n/a
Delivery	<ul style="list-style-type: none"> <li>Dose accuracy in the presence of motion (see Image fusion, Contouring, and planning)</li> <li>Determining who will most benefit from replanning</li> </ul>	<ul style="list-style-type: none"> <li>Dose accuracy in the presence of motion (see Image fusion, Contouring, and Treatment planning)</li> </ul>	No	No	No

ML, machine learning; OAR, organ at risk; QA, quality assurance.

Training machine learning requires a well-defined problem, with a well-defined ground truth, and a simple measure with which to assess effectiveness. The application to QA is not considered in detail, as the status depends on what is being assured and to what degree.



facilitate both transfer of best-practice between clinics and greater process automation.

## ACKNOWLEDGMENTS

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 766,276. This paper reflects only

the authors' views and the European Commission Research Executive Agency is not responsible for any use that may be made of the information it contains.

## CONFLICTS OF INTEREST

Daniel Jarrett and Mark Gooding are employees of Mirada Medical Ltd.

## REFERENCES

1. Bibault J-E, Giraud P, Burgun A. Big data and machine learning in radiation oncology: state of the art and future prospects. *Cancer Lett* 2016; **382**: 110–7p.. doi: <https://doi.org/10.1016/j.canlet.2016.05.033>
2. Feng M, Valdes G, Dixit N, Solberg TD. Machine learning in radiation oncology: opportunities, requirements, and needs. *Frontiers in Oncology* 2018; **8**. doi: <https://doi.org/10.3389/fonc.2018.00110>
3. Lustberg T, van Soest J, Jochems A, Deist T, van Wijk Y, Walsh S, et al. Big data in radiation therapy: challenges and opportunities. *Br J Radiol* 2017; **90**: 20160689. doi: <https://doi.org/10.1259/bjr.20160689>
4. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med* 2018; **98**: 126–46p.. doi: <https://doi.org/10.1016/j.combiomed.2018.05.018>
5. Thompson RF, Valdes G, Fuller CD, Carpenter CM, Morin O, Aneja S, et al. Artificial intelligence in radiation oncology: a specialty-wide disruptive transformation? *Radiotherapy and Oncology* 2018; **129**: 421–6. doi: <https://doi.org/10.1016/j.radonc.2018.05.030>
6. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Informatics* 2017;: 505–15p..
7. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017; **42**: 60–88p.. doi: <https://doi.org/10.1016/j.media.2017.07.005>
8. Naqa E I, Li R, Murphy MJ. *Machine Learning in Radiation Oncology*: Springer; 2015.
9. Aird EGA, Conway J. CT simulation for radiotherapy treatment planning. *Br J Radiol* 2002; **75**(937–949): 937–49. doi: <https://doi.org/10.1259/bjr.75.900.750937>
10. van Herk M. Errors and margins in radiotherapy. *Semin Radiat Oncol* 2004; **14**: 52–64p.. doi: <https://doi.org/10.1053/j.semradonc.2003.10.003>
11. Murphy MJ, Balter J, Balter S, BenComo JJA, Das IJ, Jiang SB, et al. The management of imaging dose during image-guided radiotherapy. *Medical Physics* 2007; **34**: 4041–63p..
12. Willemink MJ, Noël PB. The evolution of image reconstruction for CT-from filtered back projection to artificial intelligence. *European Radiology* 2018;.
13. Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med Phys* 2017; **44**: e360–75. doi: <https://doi.org/10.1002/mp.12344>
14. Wu D, Kim K, El Fakhri G, Li Q. Iterative low-dose CT reconstruction with priors trained by artificial neural network. *IEEE Trans Med Imaging* 2017; **36**: 2479–86p.. doi: <https://doi.org/10.1109/TMI.2017.2753138>
15. Han YS, Yoo J, JC Y. Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis. 2016; **06391**arXiv preprint arXiv:1611.
16. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, et al. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imaging* 2017; **36**: 2524–35p.. doi: <https://doi.org/10.1109/TMI.2017.2715284>
17. Chen Y, Liu J, Xie L, Hu Y, Shu H, Luo L, et al. Discriminative Prior - Prior Image Constrained Compressed Sensing Reconstruction for Low-Dose CT Imaging. *Sci Rep* 2017; **7**: 13868p.. doi: <https://doi.org/10.1038/s41598-017-13520-y>
18. Lee H, Lee J, Kim H, Cho B, Cho S. Deep-neural-network based sinogram synthesis for sparse-view CT image reconstruction. 2018; **00694**arXiv preprint arXiv:1803.
19. Gjestebly L, Shan H, Yang Q, Xi Y, Claus B, Jin Y, et al. Deep learning methods to guide CT image reconstruction and reduce metal artifacts. In *Medical Imaging 2017: Physics of Medical Imaging*; 2017: International Society for Optics and Photonics. p **101322W**.
20. Zhang Y, Yu H. Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE Trans Med Imaging* 2018; **37**: 1370–81p.. doi: <https://doi.org/10.1109/TMI.2018.2823083>
21. Nie D, Trullo R, Petitjean C, Ruan S, Shen Nie D D, Trullo R, et al. Medical image synthesis with context-aware generative adversarial networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* 2017;: 417–25p..
22. Guerreiro F, Burgos N, Dunlop A, Wong K, Petkar I, Nutting C, et al. Evaluation of a multi-atlas CT synthesis approach for MRI-only radiotherapy treatment planning. *Phys Med* 2017; **35**: 7–17p.. doi: <https://doi.org/10.1016/j.ejmp.2017.02.017>
23. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys* 2017; **44**: 1408–19p.. doi: <https://doi.org/10.1002/mp.12155>
24. Jang H, Liu F, Zhao G, Bradshaw T, McMillan AB. Technical note: deep learning based MRAC using rapid ultrashort echo time imaging. *Medical Physics* 2018; **45**: 3697–704May.. doi: <https://doi.org/10.1002/mp.12964>
25. Maspero M, Savenije MHF, Dinkla AM, Seevinck PR, Intven MPW, Jurgenliemk-Schulz IM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Phys Med Biol* 2018; **63**: 185001. doi: <https://doi.org/10.1088/1361-6560/aada6d>
26. Wolterink JM, Dinkla AM, Savenije MH, Seevinck PR, van den Berg CA, Išgum I. Deep Mr to CT synthesis using upaired data. In *International Workshop on Simulation and Synthesis in Medical Imaging* 2017;: 14–23p..
27. Tyagi N, Fontenla S, Zhang J, Cloutier M, Kadbi M, Mechalakos J, et al. Dosimetric and workflow evaluation of first commercial synthetic CT software for clinical use in pelvis. *Phys Med Biol* 2017; **62**: 2961–75p.. doi: <https://doi.org/10.1088/1361-6560/aa5452>
28. Kessler ML. Image registration and data fusion in radiation therapy. *Br J Radiol* 2006;

- 79 Spec No 1**(Special Issue 1): S99–S108p.. doi: <https://doi.org/10.1259/bjr/70617164>
29. Nix MG, Prestwich RJD, Speight R, utomated SR. Automated, reference-free local error assessment of multimodal deformable image registration for radiotherapy in the head and neck. *Radiother Oncol* 2017; **125**: 478–84p.. doi: <https://doi.org/10.1016/j.radonc.2017.10.004>
  30. Kierkels RGJ, den Otter LA, Korevaar EW, Langendijk JA, van der Schaaf A, Knopf AC, et al. An automated, quantitative, and case-specific evaluation of deformable image registration in computed tomography images. *Phys Med Biol* 2018; **63**((4)): 045026p.. doi: <https://doi.org/10.1088/1361-6560/aa9dc2>
  31. Veiga C, Lourenço AM, Mouinuddin S, van Herk M, Modat M, Ourselin S, et al. Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm. *Med Phys* 2015; **42**: 760–9p.. doi: <https://doi.org/10.1118/1.4905050>
  32. Cao X, Gao Y, Yang J, Wu G, Shen D. Learning-Based multimodal image registration for prostate cancer radiation therapy. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 2016;: 1–9p..
  33. Kearney V, Haaf S, Sudhyadhom A, Valdes G, Solberg TD. An unsupervised convolutional neural network-based algorithm for deformable image registration. *Phys Med Biol* 2018; **63**: 185017. doi: <https://doi.org/10.1088/1361-6560/aada66>
  34. Lee D, Hofmann M, Steinke F, Altun Y, Cahill N, Scholkopf B. Learning similarity measure for multi-modal 3D image registration. In *Conference on Computer Vision and Pattern Recognition* 2009;: 186–93p..
  35. Michel F, Bronstein M, Bronstein A, Paragios N. Boosted metric learning for 3D multi-modal deformable registration. In *IEEE International Symposium on Biomedical Imaging* 2011;: 1209–14p..
  36. Simonovsky M, Gutiérrez-Becker B, Mateus D, Navab N, Komodakis N. a deep metric for multimodal registration. In *International Conference On Medical Image Computing & Computer Assisted Intervention* 2016;.
  37. Cheng X, Zhang L, Zheng Y. Deep similarity learning for multimodal medical images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 2018; **6**: 248–52p.. doi: <https://doi.org/10.1080/21681163.2015.1135299>
  38. Oktay O, Schuh A, Rajchl M, Keraudren K, Gomez A, Heinrich MP, et al. Structured decision forests for multi-modal ultrasound image registration. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 2015;: 363–71p..
  39. Yang X, Kwitt R, Styner M, Niethammer M. Quicksilver: Fast predictive image registration - A deep learning approach. *Neuroimage* 2017; **158**: 378–96p.. doi: <https://doi.org/10.1016/j.neuroimage.2017.07.008>
  40. Liao R, Miao S, de Tournemire P, Grbic S, Kamen A, Mansi T, et al. An Artificial agent for robust image registration. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* 2017;: 4168–75p..
  41. Yan P, Xu S, Rastinehad A, Wood B. Adversarial image registration with application for MR and TRUS image fusion. *Proceedings of International Workshop on Machine Learning in Medical Imaging* 2018;: 197;: --204p..
  42. 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. *Medical Physics* 2017; **44**: : e43–76p.. doi: <https://doi.org/10.1002/mp.12256>
  43. Olsen DR, Thwaites DI. Now you see it... imaging in radiotherapy treatment planning and delivery. *Radiotherapy and Oncology* 2007; **85**: : 173–5p.. doi: <https://doi.org/10.1016/j.radonc.2007.11.001>
  44. Roques TW. Patient selection and radiotherapy volume definition — can we improve the weakest links in the treatment chain? *Clinical Oncology* 2014; **26**: : 353–5p.. doi: <https://doi.org/10.1016/j.clon.2014.02.013>
  45. Njeh CF. Tumor delineation: the weakest link in the search for accuracy in radiotherapy. *Journal of Medical Physics* 2008; **33**: : 136p.. doi: <https://doi.org/10.4103/0971-6203.44472>
  46. Sharp G, Fritscher KD, Pekar V, Peroni M, Shusharina N, Veeraraghavan H, et al. Vision 20/20: perspectives on automated image segmentation for radiotherapy. *Medical Physics* 2014; **41**: 050902. doi: <https://doi.org/10.1118/1.4871620>
  47. Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: a survey. *Medical Image Analysis* 2015; **24**: : 205–19p.. doi: <https://doi.org/10.1016/j.media.2015.06.012>
  48. Peressutti D, Schipaanboord B, van Soest J, Lustberg T, van Elmpt W, Kadir T, et al. TU-AB-202-10: how effective are current atlas selection methods for atlas-based Auto-Contouring in radiotherapy planning? *Medical Physics* 2016; **43**: : 3738–9p.. doi: <https://doi.org/10.1118/1.4957432>
  49. Dolz J, Laprie A, Ken S, Leroy HA, Reynolds N, Massotier L, et al. Supervised machine learning based classification scheme to segment the brainstem on MRI in Multicenter brain tumor treatment context. *Hyper Articles en Ligne* 2015;.
  50. Pinto A, Pereira S, Correia H, Oliveira J, Rasteiro D, Silva C. Brain Tumour Segmentation based on Extremely Randomized Forest with High-Level Features. In: EMBC, ed. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; 2015.
  51. 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. *Radiotherapy and Oncology* 2018; **126**: : 312–7p.. doi: <https://doi.org/10.1016/j.radonc.2017.11.012>
  52. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med Phys* 2017; **44**: 547–57p.. doi: <https://doi.org/10.1002/mp.12045>
  53. Guo Y, Gao Y, Shen D. Deformable Mr prostate segmentation via deep feature learning and sparse patch matching. *IEEE Trans Med Imaging* 2016; **35**: 1077–89p.. doi: <https://doi.org/10.1109/TMI.2015.2508280>
  54. Fechter T, Adebahr S, Baltas D, Ben Ayed I, Desrosiers C, Dolz J. Esophagus segmentation in CT via 3D fully convolutional neural network and random walk. *Med Phys* 2017; **44**: 6341–52p.. doi: <https://doi.org/10.1002/mp.12593>
  55. Bobo MF, Bao S, Huo Y, Yao Y, Virostko J, Plassard AJ et al. *Fully Convolutional Neural Networks Improve Abdominal Organ Segmentation*. **10574**. Medical Imaging: Image Processing; 2018.
  56. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 2017; **36**: 61–78p.. doi: <https://doi.org/10.1016/j.media.2016.10.004>
  57. Men K, Zhang T, Chen X, Chen B, Tang Y, Wang S, et al. Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning. *Phys Med* 2018; **50**: 13–19p.. doi: <https://doi.org/10.1016/j.ejmp.2018.05.006>
  58. 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. *International Journal of Radiation*

- oncology, biology. *Physics* 2018; **101**: 468–78p..
59. 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**: 6377–89p.. doi: <https://doi.org/10.1002/mp.12602>
  60. Gooding MJ, Smith AJ, Tariq M, Aljabar P, Peressutti D, van der Stoep J, Stoep J, et al. Comparative evaluation of autocontouring in clinical practice: a practical method using the Turing test. *Med Phys* 2018; **45**: 5105–15p.. doi: <https://doi.org/10.1002/mp.13200>
  61. Craft DL, Halabi TF, Shih HA, Bortfeld TR. Approximating convex Pareto surfaces in multiobjective radiotherapy planning. *Med. Phys.* 2006; **33**: : 3399–407p.. doi: <https://doi.org/10.1118/1.2335486>
  62. Monz M, Küfer KH, Bortfeld TR, Thieke C. Pareto navigation: algorithmic Foundation of interactive multi-criteria IMRT planning. *Phys Med Biol* 2008; **53**: 985–98p.. doi: <https://doi.org/10.1088/0031-9155/53/4/011>
  63. Hazell I, Bzdusek K, Kumar P, Hansen CR, Bertelsen A, Eriksen JG, et al. Automatic planning of head and neck treatment plans. *J Appl Clin Med Phys* 2016; **17**: 272–82. doi: <https://doi.org/10.1120/jacmp.v17i1.5901>
  64. Gallio E, Giglioli FR, Girardi A, Guarneri A, Ricardi U, Ropolo R, et al. Evaluation of a commercial automatic treatment planning system for liver stereotactic body radiation therapy treatments. *Physica Medica* 2018; **46**: : 153–9p.. doi: <https://doi.org/10.1016/j.ejmp.2018.01.016>
  65. Shiraishi S, Tan J, Olsen LA, Moore KL. Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Med Phys* 2015; **42**: 908–17p.. doi: <https://doi.org/10.1118/1.4906183>
  66. McIntosh C, Welch M, McNiven A, Jaffray DA, Purdie TG. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. *Phys. Med. Biol.* 2017; **62**: 5926–44. doi: <https://doi.org/10.1088/1361-6560/aa71f8>
  67. Valdes G, Simone CB, Chen J, Lin A, Yom SS, Pattison AJ, et al. Clinical decision support of radiotherapy treatment planning: a data-driven machine learning strategy for patient-specific dosimetric decision making. *Radiother Oncol* 2017; **125**: 392–7p.. doi: <https://doi.org/10.1016/j.radonc.2017.10.014>
  68. Chanyavanich V, Das SK, Lee WR, Lo JY. Knowledge-based IMRT treatment planning for prostate cancer. *Med Phys* 2011; **38**: 2515–22. doi: <https://doi.org/10.1118/1.3574874>
  69. Kusters JMAM, Bzdusek K, Kumar P, van Kollenburg PGM, Kunze-Busch MC, Wendling M, et al. Automated IMRT planning in pinnacle. *Strahlenther Onkol* 2017; **193**: 1031–8. doi: <https://doi.org/10.1007/s00066-017-1187-9>
  70. Wang J, Hu W, Yang Z, Chen X, Wu Z, Yu X, et al. Is it possible for knowledge-based planning to improve intensity modulated radiation therapy plan quality for planners with different planning experiences in left-sided breast cancer patients? *Radiation Oncology* 2017; **12**. doi: <https://doi.org/10.1186/s13014-017-0822-z>
  71. 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**: 1076p.. doi: <https://doi.org/10.1038/s41598-018-37741-x>
  72. Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. *Med Phys* 2016; **43**: 378–87p.. doi: <https://doi.org/10.1118/1.4938583>
  73. Ziemer BP, Shiraishi S, Hattangadi-Gluth JA, Sanghvi P, Moore KL, automated F. Fully automated, comprehensive knowledge-based planning for stereotactic radiosurgery: preclinical validation through blinded physician review. *Pract Radiat Oncol* 2017; **7**: e569–78. doi: <https://doi.org/10.1016/j.prro.2017.04.011>
  74. Mardani M, Dong P, Xing L. Deep-Learning based prediction of achievable dose for Personalizing inverse treatment planning. *International Journal of Radiation oncology, biology. Physics* 2016; **96**.
  75. Nguyen D, Long T, Jia X, Lu W, Gu X, Iqbal Z, et al. Dose prediction with U-net: a feasibility study for predicting dose distributions from contours using deep learning on prostate IMRT patients. 2017; **09233**arXiv preprint arXiv:1709..
  76. Mahmood R, Babier A, McNiven A, Diamant A, Chan TCY. Automated treatment planning in radiation therapy using generative Adversarial networks. *Proceedings of Machine Learning Research: Machine Learning for Healthcare* 2018; **85**.
  77. PHE Centre for Radiation. Chemical and environmental hazards. Radiotherapy errors and near misses data report (December 2013 to November 2015) Report No.: : : PHE-CRCE-029.. London: Public Health England, Patient Safety in Radiotherapy Steering Group; 2016.
  78. McIntosh C, Svistoun I, Purdie TG. Groupwise conditional random forests for automatic shape classification and contour quality assessment in radiotherapy planning. *IEEE Trans Med Imaging* 2013; **32**: 1043–57p.. doi: <https://doi.org/10.1109/TMI.2013.2251421>
  79. Naqa E I. An anomaly detector for radiotherapy quality assurance using machine learning. *Medical Physics* 2011; **38**(3458–3458).
  80. Li RJiang S. BT. P ValidationI. E Naqa ILi R et al. eds.2015 . pp.Cham Heidelberg New York Dordrecht LondonSpringer
  81. Tol JP, Dahele M, Delaney AR, Slotman BJ, Verbakel W. Can knowledge-based DVH predictions be used for automated. *individualized quality assurance of radiotherapy treatment plans? Radiation Oncology* 2015; **10**.
  82. Stanhope C, Wu QJ, Yuan L, Liu J, Hood R, Yin F-F, et al. Utilizing knowledge from prior plans in the evaluation of quality assurance. *Phys Med Biol* 2015; **60**: 4873–91p.. doi: <https://doi.org/10.1088/0031-9155/60/12/4873>
  83. Gits HC, Dow J, Feng M, Matuszak MM. Using a knowledge-based planning model for quality assurance of liver stereotactic body radiation therapy plans. *International Journal of Radiation oncology, biology. Physics* 2016; **96**(2S).
  84. Carlson JNK, Park JM, Park S-Y, Park JJ, Choi Y, Ye S-J. A machine learning approach to the accurate prediction of multi-leaf collimator positional errors. *Phys Med Biol* 2016; **61**: 2514–31p.. doi: <https://doi.org/10.1088/0031-9155/61/6/2514>
  85. Interian Y, Rideout V, Kearney VP, Gennatas E, Morin O, Cheung J, et al. Deep nets vs expert designed features in medical physics: an IMRT QA case study. *Medical Physics* 2018; **45**: : 2672–80p.. doi: <https://doi.org/10.1002/mp.12890>
  86. Valdes G, Chan MF, Lim SB, Scheuermann R, Deasy JO, Solberg TD. IMRT QA using machine learning: a multi-institutional validation. *J Appl Clin Med Phys* 2017; **18**: 279–84. doi: <https://doi.org/10.1002/acm2.12161>
  87. Nelms BE, Zhen H, Tomé WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Medical Physics* 2011; **38**: : 1037–44p.. doi: <https://doi.org/10.1118/1.3544657>
  88. Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Kotoku Jun'ichi, Kotoku J, et al. Cone beam computed tomography image quality improvement using a deep Convolutional neural network. *Cureus* 2018; **10**: e2548: e2548): : p.. doi: <https://doi.org/10.7759/cureus.2548>
  89. Kurz C, Hansen DC, Savenije MHE, Landry G, Maspero M, Kamp F, et al. OA127] Cone-

- beam CT intensity correction for adaptive radiotherapy of the prostate using deep learning. *Physica Medica* 2018; **52**: : 48p.. doi: <https://doi.org/10.1016/j.ejmp.2018.06.199>
90. Murphy MJ, Naqa E I, Li R, Murphy MJ.. (eds) Artificial Neural Networks to Emulate and Compensate Breathing Motion During Radiation Therapy. In: *Machine Learning in Radiation Oncology*. Cham Heidelberg New York Dordrecht London: Springer; 2015. pp. 203 -23.
  91. Tatinati S, Nazarpour K, Tech Ang W, Veluvolu KC. Ensemble framework based real-time respiratory motion prediction for adaptive radiotherapy applications. *Medical Engineering & Physics* 2016; **38**: : 749–57p.. doi: <https://doi.org/10.1016/j.medengphy.2016.04.021>
  92. Zarepisheh M, Long T, Li N, Tian Z, Romeijn HE, Jia X, et al. A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning. *Medical Physics* 2014; **41**: 061711. doi: <https://doi.org/10.1118/1.4875700>
  93. El Naqa I, Feng M, Bazzi L, Dow J, Cuneo KC, Matuszak MM, et al. Reinforcement learning strategies for decision making in knowledge-based adaptive radiation therapy: application in liver cancer. *International Journal of Radiation Oncology\*Biophysics* 2016; **96**: S45p.. doi: <https://doi.org/10.1016/j.ijrobp.2016.06.119>
  94. Tseng H-H, Luo Y, Ten Haken RK, El Naqa I, Haken RKT, Naqa E I. The role of machine learning in knowledge-based Response-Adapted radiotherapy. *Frontiers in Oncology* 2018; **8**. doi: <https://doi.org/10.3389/fonc.2018.00266>
  95. Tseng H-H, Luo Y, Cui S, Chien J-T, Ten Haken RK, Naqa IE. Deep reinforcement learning for automated radiation adaptation in lung cancer. *Medical Physics* 2017; **44**: : 6690–705p.. doi: <https://doi.org/10.1002/mp.12625>
  96. Guidi G, Maffei N, Vecchi C, Gottardi G, Ciarmatori A, Mistretta GM, et al. Expert system classifier for adaptive radiation therapy in prostate cancer. *Australas Phys Eng Sci Med* 2017; **40**: : 337–48p.. doi: <https://doi.org/10.1007/s13246-017-0535-5>
  97. 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. *Physica Medica* 2016; **32**: : 1659–66p.. doi: <https://doi.org/10.1016/j.ejmp.2016.10.005>