FAMPO
FEDERATION OF AFRICAN MEDICAL PHYSICS ORGANIZATIONS

African Journal of Medical Physics
2022; 4(1): 27-40. doi: 10.

Review Paper

Radiotherapy Plan Quality Assessment: Current Tools and Future Directions

Gronberg Mary P.[1, 2,✉], Netherton Tucker J.[1, 2], Beadle Beth M.[3], Cao Wenhua [1], Cardenas Carlos E. [4], Chung Christine[1], Fuller Clifton D.[2, 5], Garcia John [6], Hancock Donald [6], Howell Rebecca M. [1, 2], Jhingran Anuja[5], Lee Anna[5], Lim Tze Yee [1, 2], Moreno Amy C. [2, 5], Olanrewaju Adenike [1], Peterson Christine B. [2, 7], Whitaker Thomas J. [1, 2], Court Laurence E. [1, 2]

1. Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.
2. The University of Texas MD Anderson Cancer Center UTHealth Houston Graduate School of Biomedical Sciences, Houston, Texas, USA.
3. Department of Radiation Oncology, Stanford University, Stanford, California, USA.
4. Department of Radiation Oncology, The University of Alabama at Birmingham, Birmingham, Alabama, USA.
5. Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.
6. Department of Radiation Oncology Physics Support, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.
7. Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

✉ Corresponding author: Mary Gronberg, The University of Texas MD Anderson Cancer Center UTHealth Houston Graduate School of Biomedical Sciences, Houston, Texas, USA; E-mail: mpeters1@mdanderson.org

ISSN 2643-5977

Received: 2022.08.18; Accepted: 2023.03.21; Published: 2023.07.30

Abstract

As artificial intelligence is increasingly adopted in radiation oncology, we have the opportunity to turn our attention toward its use in improving the quality of radiation treatment plans. In this work, we present an overview of patient-specific plan quality assessment tools found in the literature and discuss the pros and cons of each. The findings of an institutional consortium of radiation oncologists, dosimetrists, and physicists who reviewed currently available plan quality assessment tools are presented and include the types of tools they found most beneficial in the clinic, improvements they would like to see in currently available tools, where in the treatment planning process plan quality assessment tools should be implemented, and potential novel applications. Finally, future directions are considered, including the role of plan quality assessment tools in clinical trials, education and training, and peer review.

Keywords: Radiotherapy; quality assessment; treatment planning; artificial intelligence.

Introduction

With advances in data science and data management in radiation oncology [1], we can learn from past experiences to better tailor radiotherapy for each patient. One example of this is knowledge-based planning [2], which refers to the application of data-driven approaches in the development of radiation treatment plans. Machine learning and deep learning approaches have mostly been applied to the automation of the treatment planning process [3-7]; however, these approaches could also play a role in the assessment of plan quality.

In this work, we present an overview of patient-specific radiation treatment plan quality assessment (QA) tools found in the literature and discuss their strengths and weaknesses. Unlike broader reviews of quality assurance in radiotherapy [8-10], this review focuses on available tools for assessing whether the planned dose distribution is optimal in regard to target coverage and normal tissue

sparing. We also discuss which plan QA tools members of the radiation oncology team may find most beneficial, where they should be implemented in the treatment planning process, how they can be improved, and opportunities and promising future directions. We include an added emphasis on the potential for plan QA tools to help address disparities in access to radiotherapy in low- and middle-income countries (LMICs).

Tools for Evaluating Radiation Treatment Plans

The two main contributors to plan quality are contours and dose, which may be assessed together or separately. When assessed separately, target contours are often assessed early in the process, prior to planning. The focus of this review paper is on the evaluation of the dose distribution. During the assessment of the dose distribution, the reviewer scrolls through the patient's computed tomography images and visually inspects the three-dimensional dose distribution in axial, sagittal, and coronal planes, followed by review of the dose-volume histograms (DVHs) of the targets and normal tissues. A DVH is a two-dimensional graphical representation of the dose distribution within a given structure [11]. Dose and dose-volume metrics on a per-structure basis may be reviewed, including minimum, maximum, and mean doses; $D_{V\%}$, the least dose that the hottest V% of the volume receives; or $V_{D\,\mathrm{Gy}}$, the volume receiving a dose of D Gy or more [11-13]. The conformity index [11,14,15] and homogeneity index [11,16] are additional target metrics that indicate how well the prescription isodose curve covers the target and how homogeneous the dose is within the target; an ideal plan reduces cold and hot spots. Lastly, radiobiological metrics such as normal tissue complication probability and tumor control probability can be assessed when clinically relevant [11, 17-19]. In this section, we review the types of tools that can be used to assist with plan QA. A comparison of the tools, including the advantages and disadvantages of each, is displayed in Table 1.

Table 1. A comparison of the types of tools that can be used to assist with radiotherapy plan quality assessment (QA).

Plan QA Tools	Advantages	Weaknesses
Scorecard	Easy to interpret Plan safety check	Constraints may be unreasonable for a given patient Achieving constraints may not indicate that the best plan has been achieved
Patient population comparison	Easy to interpret Shows if a plan is reasonable compared to historical experience	Large database needed Does not show if the plan can be improved Limited by the quality of the plans in the database
Individual patient comparison	Indicates how the plan can be improved The plan is compared to prior plans of patients with similar disease presentation and geometry	Large database needed to find a similar patient Limited by the quality of the plan(s) retrieved from the database
Geometric models for dose-volume histogram (DVH) estimation	The entire DVH curve is predicted Uses mathematical models that relate patient geometry (i.e., distance between an organ at risk (OAR) and target) with achievable normal tissue sparing	Mathematical DVH prediction models are specific to an OAR High-quality database needed for training
Deep learning-based models for three-dimensional dose prediction	The entire dose distribution is predicted Deep learning models can incorporate all treatment planning information (images, prescription, and contours)	Large amounts of data are predicted, which can be challenging to interpret Future work is needed to estimate prediction uncertainty High-quality database is needed for training Deep learning models are less interpretable ("black box")

Scorecards

Scorecards are likely the most commonly used plan evaluation tool. They evaluate how well the current plan meets institution-specific dose-volume objectives and constraints for the given disease site. This information is commonly displayed as a checklist (Figure 1) comparing achieved metrics with the desired objectives and constraints. The main advantage of scorecards is that they are easy to use and interpret. In fact, scorecards are even incorporated in some treatment planning systems. However, scorecards do have several weaknesses. Normal tissue constraints are typically well-established thresholds for severe toxicity [17, 18]. Achieving these constraints indicates a certain level of plan safety, not quality. Moreover, scorecards are not patient-specific, meaning that they do not show if the plan can be improved for the given patient in review. The scorecard showing that a specific normal tissue constraint is being met does not indicate that optimal sparing has been achieved for that structure as further sparing may be achievable. Conversely, the scorecard showing that a normal tissue constraint is not being met does not mean that the plan is sub-optimal, as it may not be possible to meet a constraint given the patient's disease and proximity to normal tissues.

Patient population comparison

There are immense opportunities to answer clinical questions and advance the field of radiation oncology by taking advantage of big data [1, 20, 21]. In one example, the University of Michigan is using its wealth of big data to inform radiation treatment planning through the use of statistical DVH dashboards (Figure 2) that compare a patient's treatment plan with historical experiences [22]. Statistical DVH curves display the planned DVH curve alongside the median historical DVH curve and shaded confidence interval bands. Plan- and structure-level information is displayed using box and whisker plots. A generalized evaluation metric indicates whether a normal tissue constraint was met in comparison with historical achievement. Patient population comparisons thus show if the achieved plan is reasonable; however, they do not show if the plan can be improved for that specific patient. Another limitation of patient population comparisons is the assumption that all prior plans are high-quality.

OAR Dose Constraints

ROI	Constraint	Value	Pass/Fail
Brain	Max < 54 Gy	11.16 Gy	✓
BrainStem	Max < 54 Gy	14.15 Gy	✓
Chiasm	Max < 54 Gy	0.85 Gy	✓
Cochlea_L	Max < 35 Gy	2.01 Gy	✓
Cochlea_R	Max < 35 Gy	3.46 Gy	✓
Lens_L	Max < 7 Gy	0.54 Gy	✓
Lens_R	Max < 7 Gy	0.87 Gy	✓
Mandible	Max < 70 Gy	62.72 Gy	✓
OpticNrv_L	Max< 54 Gy	1.45 Gy	✓
OpticNrv_R	Max < 54 Gy	0.95 Gy	✓
Parotid_L	Mean < 26 Gy	33.47 Gy	✗
Parotid_R	Mean < 26 Gy	20.24 Gy	✓
SpinalCord	Max < 45 Gy	29.53 Gy	✓

Figure 1. Scorecard indicating if the current radiotherapy plan meets institution-specific dose-volume constraints for organs at risk (OARs).

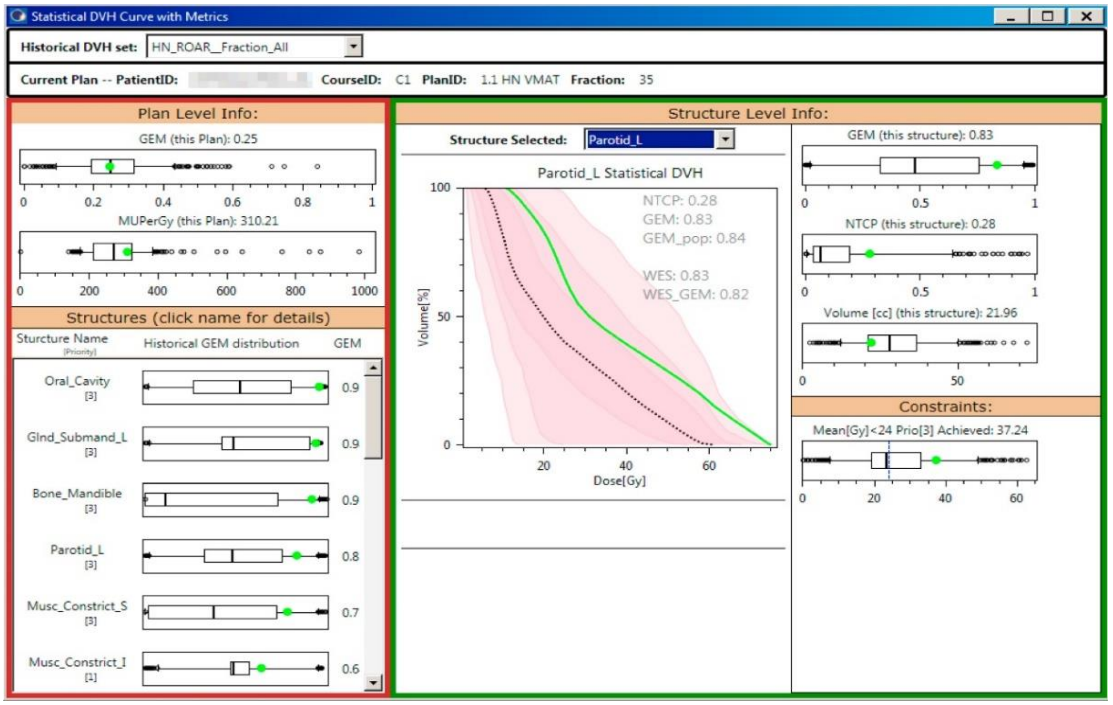

Figure 2. A patient population comparison.

The current radiotherapy plan is compared with all prior plans for the same disease site. Figure reprinted with permission from Advances in Radiation Oncology, Volume 2, Issue 3, Incorporating big data into treatment plan evaluation: Development of statistical DVH metrics and visualization dashboards, Charles S. Mayo et al, Pages 503-514, Copyright Elsevier (2017). Original figure caption: Statistical dose-volume histogram (DVH) dashboard quantifies comparison of statistical metrics for the current plan (green) versus historical experience. Statistical DVH (center) compares the DVH curve to historical experience for the median (dashed line), 50% confidence interval (CI; dark pink), 70% CI (intermediate pink), and 90% CI light pink. Box-and-whisker plots compare plan level (left panel) and structure level (right panel) metrics.

Individual patient comparison

In contrast to statistical comparisons, which compare a patient's plan with all prior plans, individual patient comparisons reference a small number of similar prior plans. They are modeled off clinical practice, where physicians often search for records of patients with similar disease presentation and geometry and review their radiation treatment plans to determine realistic planning goals. Patient geometry descriptors can be used to query a database to find prior patient plans with similar geometry. The overlap volume histogram (OVH) [23], or distance-to-target histogram [24], is a plot of the fractional volume of an organ at risk (OAR) within a certain distance from the planning target volume (PTV). The OVH is a simple way of capturing the geometric relationship between an OAR and the radiation target, depicting the level of normal tissue sparing that is possible.

Ge and Wu [2] published a review of the literature reporting case-based methods, which use prior knowledge from matching cases to inform planning. Although most of the reviewed work reported the use of case-based methods for treatment planning, case-based methods also have great potential for treatment plan QA. Researchers at Johns Hopkins University, who defined the OVH, developed a quality control method to flag plans with suboptimal OAR sparing [23]. The quality control method involves the following steps. (1) Create a database of prior patients with the same disease site and treatment protocol. In the database, store the DVHs of the targets and OARs, as well as the OVHs for each OAR. (2) When a patient's plan is complete, query the database to compare OAR sparing in the current patient's plan with that of the plans in the database. Use the current patient's OAR OVH and planned DVH to query the database using a simple rule-based approach: return prior plans for which the target-to-OAR distance is less but the dose is also less, indicating that further sparing may be achievable for the current patient. (3) Re-plan to reduce the OAR dose to the minimum queried. This quality control method is displayed in Figure 3.

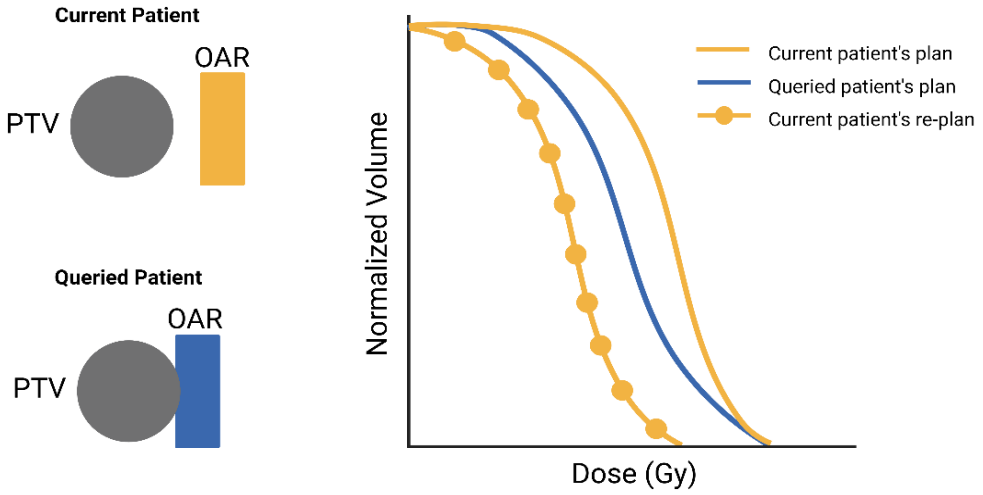

Figure 3. An individual patient comparison.

A prior patient is returned from the database with worse geometry (the blue organ at risk (OAR) is closer to the planning target volume (PTV) than the yellow OAR) but better OAR sparing than the current patient (the blue dose-volume histogram (DVH) is to the left of the yellow DVH). Re-planning the current patient shifts the yellow DVH curve to the left resulting in the dotted yellow DVH curve.

A database query returns a prior patient who had worse geometry than the current patient (the blue OAR is closer to the PTV than the yellow OAR) but better normal tissue sparing (the blue DVH of the queried plan is to the left of the yellow DVH of the current plan). This scenario indicates that the OAR for the current plan can be better spared. Re-planning achieves the yellow dotted DVH.

Wu et al. [23] demonstrated the ability of their quality control method to flag suboptimal parotid gland sparing in a retrospective study of the intensity-modulated radiotherapy (IMRT) treatment plans of 32 patients with head and neck cancer. Each parotid gland was compared with the other 63 parotid glands in the database according to the Radiation Therapy Oncology Group (RTOG) sparing goal of $V_{30\ Gy} < 50\%$, and 17 parotid glands were flagged. All plans in the database were re-planned without compromising target coverage, and re-planning reduced $D_{50\%}$ by an average of 6.6 Gy for flagged parotid glands and 1.9 Gy for unflagged parotid glands. Individual patient comparisons are useful QA tools because they limit comparisons to similar patients. This approach also details where the plan can be improved and by what magnitude. However, limitations of this approach are the need for a large database to have a high likelihood of finding a similar patient, and the assumption that queried plans in the database are high-quality.

Geometric models for DVH estimation

In the early 2010s, researchers began to investigate the relationship between patient geometry and achievable dosimetric sparing. In 2011, Moore et al [25] identified a mathematical correlation between the overlap of the principal OAR and PTV and the achievable mean dose to the OAR. They implemented a script in the treatment planning workflow that displayed predictions of achievable mean doses for principal OARs with PTV overlap. Clinical implementation of the script resulted in improved normal tissue sparing and reduced inter-clinician plan variability. A limitation of this early work is that the model applies only to disease sites for which an OAR overlaps with the PTV (bladder and rectum for prostate cancer; parotid glands, esophagus, and larynx for head and neck cancer). Furthermore, the model predicts only the mean dose and does not predict volumetric dose information.

Since the publication of the overlap model, more sophisticated models have been published in the literature that estimate entire DVH curves. Zhu et al [24] developed a machine learning model that predicts bladder and rectum DVHs for adaptive prostate IMRT plans. Yuan et al [26] developed a similar model that predicts OAR DVHs for prostate and head and neck IMRT plans. Both the Zhu et al and Yuan et al models used principal component analysis applied to DVH and OVH curves for feature selection and support vector regression to establish the correlation between anatomical inputs and DVHs. Zhang et al [27] expanded on the work of Yuan et al and added an anomaly detection method to their prediction model, which identifies patients whose anatomical features are not represented in the training set. Appenzoller et al [28] developed DVH prediction models for prostate and head and neck IMRT plans using a mathematical framework that fits substructure differential DVHs to skew normal probability distributions. Building off the mathematical framework from the Appenzoller et al models, Shiraishi

et al [29] developed DVH prediction models that predict both target and OAR DVHs for intracranial stereotactic radiosurgery.

The DVH prediction models described thus far were trained with a database of prior plans. Ahmed et al [30] developed a unique approach that does not require a training database. This model estimates the best possible DVH for an OAR, which the authors termed feasibility DVH, using mostly first principles. This algorithm assumes a uniform prescription dose to the targets and approximates the minimum dose that out-of-target voxels will receive by approximating the high- and low-gradient dose spreads.

RapidPlan (Varian Medical Systems, Palo Alto, CA) is a commercially available knowledge-based planning system within the Eclipse treatment planning system. RapidPlan allows users to train machine learning DVH prediction models with patients from their clinic or, alternatively, to use pre-loaded DVH estimation models. RapidPlan DVH models predict an achievable DVH range. The lower band of the DVH prediction range is then used to inform the optimization objectives used in inverse planning. To make DVH prediction models more accessible to clinics that do not have in-house models or do not use the Eclipse treatment planning system, researchers at the University of California San Diego developed On-line Real-time Benchmarking Informatics Technology for RadioTherapy (ORBIT-RT) [31], a free web-based DVH estimation platform that closely follows the methodology of Appenzoller et al. Registered users can upload basic plan metadata and anonymized DICOM to https://www.orbit-rt.com/. After performing contour matching, ORBIT-RT provides a DVH estimation for the patient, including a patient-specific DVH prediction error estimate.

Figure 4 demonstrates how DVH prediction models can be used for patient-specific plan QA. The yellow DVH is flagged for re-planning because it is outside and to the right of the predicted DVH band shaded in blue. After replanning, the yellow dotted DVH falls within the prediction band. Tol et al [32] used a head and neck RapidPlan model to evaluate the plan quality of 20 retrospective head and neck plans. Half of the plans were flagged for suboptimal sparing because the mean dose of the predicted DVH was more than 3 Gy, 5 Gy, and 7 Gy lower than the clinical mean doses to the salivary glands, oral cavity, and swallowing muscles, respectively. The 10 manual plans that were flagged were re-planned using RapidPlan, and the predicted sparing gains were generally achieved. Similarly, Cao et al [33] reported on the feasibility of using RapidPlan models for head and neck plan QA, incorporating physician review. In that study, 40 of 45 structures that were flagged by the RapidPlan model were also flagged by physicians, showing the potential of the RapidPlan model in assisting physician plan review.

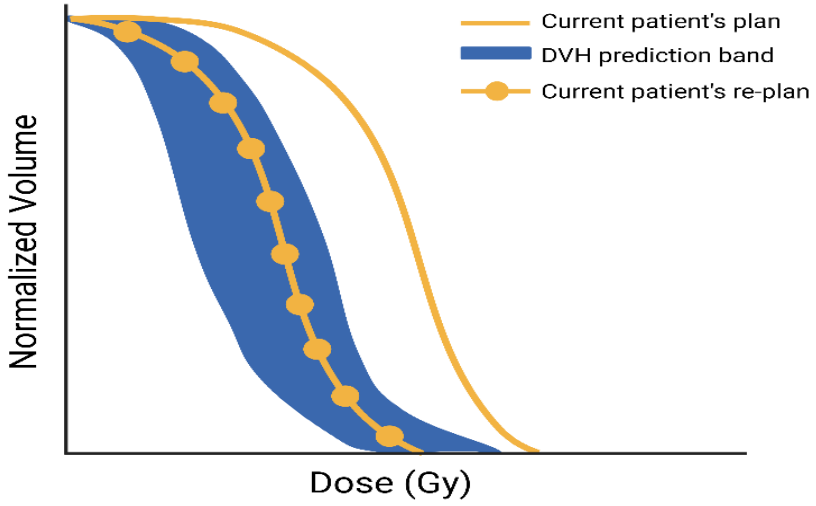

Figure 4. Radiotherapy plan quality assessment (QA) using dose-volume histogram (DVH) prediction.

A mathematical model predicts that the patient's yellow DVH can fall within the shaded blue prediction range. The DVH is flagged because it lies outside and to the right of the predicted DVH band. Following QA, the patient is re-planned to bring the re-planned DVH (yellow dotted DVH) within the predicted region.

An advantage of DVH prediction models compared with patient population comparisons and case-based approaches is that a smaller dataset can be used. However, the quality of the plans in the dataset is arguably more important. RapidPlan models can be trained with a minimum of 20 high-quality plans. Alternatively, one has the option to use clinically

validated RapidPlan models developed at other institutions. ORBIT-RT currently offers 17 validated models, which can be easily downloaded. When using DVH predictions as a QA tool, it is important to consider DVH estimation uncertainty, especially when using models that were trained with patients from a different institution. Covele et al [34] presented an independent error quantification method that can be applied to any DVH-estimation model to quantify its prediction accuracy and uncertainty. To follow this method, an institution uses a cohort of prior high-quality patient plans. The model-predicted DVHs are compared with the clinical DVHs across the cohort to yield a root-mean-square error of the predictions, which serves as a DVH uncertainty band.

Deep learning–based models for three-dimensional dose prediction

A major limitation of DVH prediction models is their inability to provide spatial dose distribution information. In recent years, deep learning models have enabled researchers to predict patient-specific three-dimensional dose distributions for treatment of cancers of the head and neck [35-42], breast [43-45], esophagus [46, 47], lung [48], abdomen [49], cervix [50], prostate [51-53], and rectum [54]. Convolutional neural networks learn from past experiences to predict an optimal solution for a new patient. Deep learning dose prediction algorithms receive the same information as a treatment planner, including computed tomography images, prescriptions, and normal tissue and target contours. These algorithms then extract low- and high-level features from planning inputs and learn patterns related to the dose distribution. Deep learning-based dose prediction algorithms are being developed to automate treatment planning, as described in a review paper by Wang et al [5]. However, these algorithms also have great promise as a tool for patient-specific plan QA. Figure 5 displays how dose prediction can be used for QA. A comparison of the predicted versus planned dose distribution for a patient with head and neck cancer indicates that the spinal cord can be better spared. The deep learning-based dose prediction shows that the spinal cord can be spared by the 30 Gy isodose curve (in lime green) and that the maximum dose to the spinal cord can be reduced by more than 12 Gy, from 41.41 Gy to 28.66 Gy.

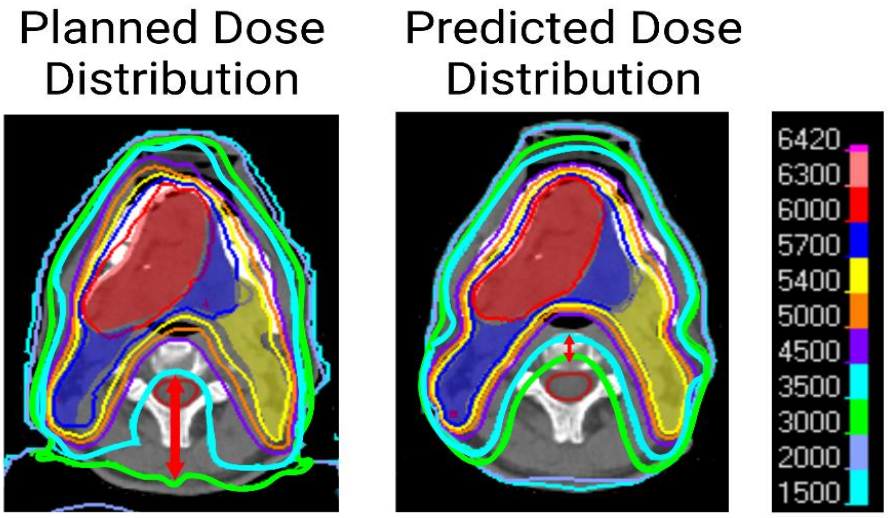

Figure 5. Radiotherapy plan quality assessment using deep learning-based dose prediction.

The deep learning-based dose prediction indicates that the spinal cord can be better spared for this head and neck cancer patient..

Deep learning dose prediction algorithms can consider dosimetric tradeoffs in much the same way as planners and physicians. Because the algorithms learn the physician preferences of the training dataset, there may be challenges in applying these models when physician planning preferences differ. Another challenge with dose prediction models is that they are typically disease site–specific; however, Mashayekhi et al. [55] showed that site-agnostic models can be created using more generalized model inputs. In that study, all OARs were input into one channel in the model rather than having a separate channel for each OAR, as is traditionally done. This approach enables the model to quickly be adapted to a different treatment site.

Another risk of deep learning models for plan QA is their potential misuse in evaluating patient plans that are not representative of the patient plans used to train the model. As models begin to be used for the clinical

judgment of plan quality, model estimation uncertainty needs to be further investigated. Nguyen et al. [56] demonstrated that bootstrap aggregation and Monte Carlo dropout both provide reasonable uncertainty estimations. As dose prediction models continue to be developed, patient-specific prediction uncertainty should continue to be investigated and quantified.

Review of Plan QA Tools by an Institutional Radiation Oncology Consortium

To discuss the current and future roles of plan QA tools, we brought together a consortium of radiation oncology team members at The University of Texas MD Anderson Cancer Center, including two dosimetrists, two radiation oncologists, and two medical physicists. Members were invited to share their feedback on the plan QA tools that they find most beneficial for achieving high-quality treatment plans, as well as how these tools can be improved or better implemented.

Regarding scorecards, the radiation oncologists appreciated their ease of use because their attention is quickly drawn to areas of the plan that do not meet the desired planning goals. However, they highlighted the need for disease subsite–specific scorecards to enable the use of stricter constraints where feasible. In addition to disease subsite–specific scorecards, the radiation oncologists proposed scorecards specific to disease presentation, such as human papillomavirus status or disease laterality.

Consortium members acknowledged the benefits of patient population comparisons and their usefulness in showing the bounds of possibility from a population standpoint. They suggested the incorporation of patient outcome data into patient population comparisons. As planning practices continue to change with the introduction of new plan QA tools, statistical dashboards can help document changes in planning trends.

With respect to individual patient comparisons, a radiation oncologist indicated that she frequently reviews records of prior patients with similar staging when planning complex cases. She found patient-specific comparisons to be very useful for determining a realistic plan. However, she noted that this approach relies on the assumption that the prior plan was of optimal quality, and that this approach may prevent progress.

Much of the focus of the discussion revolved around the desire to incorporate DVH and dose distribution prediction into the treatment planning workflow. In the current workflow at our institution, the physician creates a planning directive for each plan, which includes target definitions and expansions, the prescription, target and normal tissue goals, and any special planning considerations. As in many other institutions, physicians at MD Anderson use a template of planning goals for each disease site that are included in the planning directive. Physicians can modify these general rules for each patient according to what is believed to be achievable given the patient's disease extent and location relative to critical organs. The radiation oncologists and dosimetrists expressed their desire to have dose prediction tools before the start of treatment planning and the creation of the planning directive. This would enable the treating physician and planner to discuss what is achievable for a given patient before planning. The physician could then modify the planning directive, with the expectation that more informed planning directives would ultimately lead to fewer second plans needed.

The promise of predicted dose distributions to serve as educational tools was also discussed. At MD Anderson, which is a large academic institution, physicians, dosimetrists, and physicists have disease site specializations. Peer review is conducted by other radiation oncologists and physicists with the same disease specialization. In contrast, radiation oncology teams at community practices, representing most clinics around the world, treat all disease sites. Most radiation oncology practices could greatly benefit from dose prediction tools that are trained with high-quality plans created by disease site experts, particularly for rare disease sites.

Future Applications of Plan QA Tools

Patient-specific plan QA tools have the potential to improve the quality and consistency of radiation treatment plans around the world. Plan QA tools have an important role to play within clinical trials; these tools could improve trial compliance and increase the power of clinical trials to inform best practices. Plan QA tools can also help to reduce disparities in access to cancer care. It is estimated that by 2030, 75% of cancer deaths will occur in LMICs [57]. Plan QA tools can be used to develop training tools, assist with the safe implementation of new technology, and serve as peer review.

Radiotherapy quality assessment in clinical trials

The quality of radiotherapy is especially important as it pertains to clinical trials because poor-quality radiotherapy brings into question the validity of trial results. Noncompliance with radiotherapy protocols in clinical trials has been linked to poor outcomes [58-70]. Concerns regarding radiotherapy protocol deviations and diversity in the implementation of radiotherapy

quality assurance in clinical trials led to the establishment of the Global Clinical Trials Radiation Therapy Quality Assurance Harmonisation Group in 2010 [71]. Members of this group described important components of clinical trial quality assurance programs, including radiotherapy and imaging credentialing, informatics platforms for data acquisition and management, real-time case review, and educational resources [72].

Gondi et al. [73] made a strong case for the importance of real-time pre-treatment plan review. A two-step quality assurance process including pre-enrollment credentialing and centralized plan review was implemented in RTOG 0933, a phase II trial of hippocampal avoidance during whole-brain radiotherapy for patients with brain metastases. Pre-enrollment credentialing of each physician participating in the trial involved a "dry run" of planning for a sample test patient. Once the physician passed the pre-enrollment credentialing, they then needed to submit their patient's plans for review until three consecutive plans passed review. Despite > 95% of physicians passing pre-enrollment credentialing, 24% of cases submitted for pre-treatment review were deemed unacceptable for deviations in contouring or planning. This finding highlights the need for patient-specific plan QA in clinical trials.

Although human review of every case may not be practical for large trials, predicted DVHs and dose distributions can serve as important automatic plan QA tools. A retrospective study by Moore et al. [74] used DVH predictions to quantify the frequency and clinical severity of suboptimal planning in RTOG 0126, a phase III study of high-dose versus standard-dose radiotherapy for patients with localized prostate cancer. An existing DVH prediction model [28] trained with 20 high-quality prostate IMRT plans administered at the University of California San Diego was used to predict DVHs for 219 IMRT patients from the high-dose arm of the trial. The predicted DVHs of the high-dose arm were compared with the achieved DVHs to identify the top 10% highest quality plans. The DVH prediction model was then re-trained with the identified highest quality plans. DVH predictions were re-computed with the newly trained model, enabling benchmarking of all plans against the 90th percentile of plan quality. Predicted rectal and bladder DVHs were compared with their corresponding clinical DVHs to calculate excess normal tissue complication probabilities. It was found that 42.9% of patients were exposed to > 5% excess risk of grade 2+ late rectal toxicities. A re-planning study with a subset of patients showed that the predicted DVHs are achievable without reducing target coverage. Many of the plans flagged for improvement in this retrospective analysis met RTOG 0126 protocol constraints for the rectum, highlighting the weaknesses of generalized scorecards and the need for patient-specific plan QA in clinical trials. Li et al [75] have since proposed a method for training and validating RapidPlan models for plan quality assessment in clinical trials.

Educational Tools

Another major untapped potential source of plan QA tool improvement is the development of training tools. Insufficient training and a small load of patients with rare disease sites negatively affect patient overall survival rates. Almost 25% of radiation oncology residents in the United States see less than 10 head and neck cancer patients per year [76]. At centers with low patient volumes (median accrual of 4 patients to 21 RTOG trials), patients with locally advanced head and neck cancer treated with three-dimensional conformal radiotherapy had a 91% increased risk of death in RTOG 0129 compared with those treated at centers with high patient volumes (median accrual of 65 patients to 21 RTOG trials) [77]. Retrospective studies on insurance claims data also show improved overall survival for patients with head and neck cancer treated by radiation oncologists who had high patient volumes [78,79]. Findings by Boero et al [79] suggest that radiation oncologist experience is even more important for advanced radiotherapy techniques such as IMRT. Overall, these survival data indicate a need to expand physician training in rare disease sites.

An ongoing research project by the International Atomic Energy Agency (IAEA) and LMIC radiation oncologists is investigating the potential of radiotherapy quality assurance and educational workshops to enhance the quality of radiation treatment planning and improve outcomes for nasopharyngeal carcinoma in LMICs [80]. The first phase of the study was to gather baseline data from participating centers, including a baseline assessment of planning quality. Phase I results indicated that 5 of 14 participating centers had no local radiotherapy QA available. Of 134 plans submitted across 13 centers, 81 plans (60.4%) were assessed as unacceptable, with major deviations in contouring and/or dose to the brainstem, spinal cord, optic nerves, gross tumor volumes, or high-risk PTV. The phase I results highlight the great disparities in plan quality across the globe. Plan prediction tools can help educate these centers about how to identify a good plan.

Advanced training tools for plan QA can also help to address the rate at which we can train clinical staff across the global arena. There is a severe shortage of oncologists across Africa and Asia, with > 1000 incident cancers per clinical oncologist in 25 African countries and 2 Asian countries, and no oncologists in 7 African countries and 1 Asian country [81]. Inadequate staffing can be attributed to a lack of formal training programs. Historically, the

IAEA supported the training of LMIC medical professionals overseas [82, 83]. However, overseas training has contributed to "brain drain," where educated trainees do not return to their home country [84-87]. Volunteer groups [88, 89] and national organizations [90] have traditionally brought in experts from neighboring countries to assist with local staff training. However, many of these efforts were halted as a result of the COVID-19 pandemic. It is increasingly apparent that the establishment of local training programs is crucial to addressing the growing cancer burden [91-93].

As access to radiotherapy expands and centers convert from cobalt-60 to linear accelerator–based treatments and adopt more advanced treatment techniques, such as IMRT/VMAT or hypofractionated treatments, there will be a large learning curve. Plan QA will be important for ensuring the safe delivery of new treatments and can assist staff with determining if a plan is of good quality. Deep learning–based dose prediction can be used to develop interactive online training tools. These tools would enable physicians and planners to upload their treatment plans. Deep learning tools would generate the patient's predicted treatment plan. The assessment tool would then show a comparison of the achieved and predicted plan and provide feedback to the user on how their plan can be improved, including reductions in normal tissue doses or improvements in target coverage.

Peer review

To ensure the quality and safety of radiation plans, many institutions conduct peer review "chart rounds," where each patient's plan is reviewed by members of the treatment team other than the treating physician. Reviews of the literature [94-96] and survey data [21,97,98] suggest that peer review practices vary greatly across institutions, with variation in meeting timing, participation, content, scoring, and documentation. Ford *et al*. [99] investigated the effectiveness of 15 common quality control checks, including chart rounds, to detect near-miss incidents. Each quality control measure was analyzed to determine if it could potentially detect and prevent each near-miss incident. Physician chart rounds were found to have the second-lowest effectiveness at detecting and preventing errors, at 12%. In a blind prospective study by Talcott et al [100], treatment plans with errors identical to those reported to the Radiation Oncology Incident Learning System were randomly inserted into weekly chart rounds to see if they would be detected. Only 55% (11/20) of the clinically significant problematic plans were detected. Artificial intelligence (AI)-driven plan QA tools can help to reduce variability in chart round practices and improve error detection.

Most requested changes resulting from peer review are related to preplanning steps, including target delineation, prescription, or planning directives [97]. It is less common for plans to be flagged for their dose distribution. In the Talcott et al study, all four plans with poor target coverage passed through chart rounds undetected. A common recommendation for complex treatments is that they be reviewed in two stages [94,101-103], with an early peer review of physician-drawn targets [104-106]. Early peer review chart rounds would reduce the need for re-planning and enable a more detailed review of the dosimetric aspects of the plan during late peer review chart rounds. As more institutions adopt a two-step review process, DVH and dose prediction tools can assist with dosimetric plan review in late peer review chart rounds. These late chart rounds often take place immediately before or soon after a patient begins radiotherapy, resulting in resistance to re-planning. AI–based plan QA tools would enable physicians to decide when the benefits of re-planning are large enough to warrant treatment delays.

Despite variability in peer review practices and low plan flagging rates, peer review remains the best way to address the subjectivity of treatment decisions. Unfortunately, many cancer centers in LMICs do not have formal peer review processes in place. There are many ongoing efforts to establish web-based and cloud-based peer review programs in Africa, where oncologists prescribe both chemotherapy and radiotherapy and treat patients who are young, have advanced disease, are pregnant, are malnourished, or have comorbidities such as human immunodeficiency virus or tuberculosis [107,108]. Treating these patients is complex, and physicians would greatly benefit from access to peer review. The experience of Maputo Central Hospital in Mozambique is a prime example of the large impact peer review can have on patient outcomes. Maputo Central Hospital implemented its first multidisciplinary tumor board for patients with breast cancer beginning in March 2016 [109]. Improved coordination of care and peer review resulted in a 53% reduction in mortality among patients with early breast cancer. The improvement in patient outcomes and cost-effectiveness of tumor boards led Maputo Central Hospital to form tumor boards for thoracic conditions, gynecologic cancer, head and neck cancer, and esophageal cancer.

Global partnerships have helped to expand access to peer review in Africa. In 2012, the IAEA piloted a teleconference peer review program, the Africa Radiation Oncology Network (AFRONET), with the goals of improving patient care by providing evidence-based clinical decision-making and strengthening resident

education [107]. During the AFRONET pilot in 2012-2017, 16 cancer centers across 14 African countries regularly participated in monthly peer review sessions with experts from outside the African region, including the United States, Canada, Antigua, India, Indonesia, Bangladesh, Macedonia, New Zealand, and Romania. Feedback from participating centers was very positive, and AFRONET continues to be expanded. However, the uptake of peer review in additional centers has progressed slowly. This slow uptake could be a result of a lack of time and resources at busy clinics, where there are long waitlists of patients needing treatment. These centers would greatly benefit from automated peer review tools that flag patients for whom predicted doses are significantly different from those planned.

Conclusions

Tools for evaluating radiation treatment plans can be categorized into five major groups: scorecards, patient population comparisons, individual patient comparisons, geometric models for DVH estimation, and deep learning-based models for three-dimensional dose distribution prediction. Big data and AI will enable us to move away from general plan QA metrics toward personalized ones. Disease subsite–specific scorecards should be developed from clinical experience and based on realistic objectives and constraints to encourage planners and physicians to further push plans where possible, while also discouraging them from pushing for metrics that are unachievable. Patient population comparisons and statistical analysis will continue to inform planners of whether plans are reasonable compared with prior experience. Patient outcome data should be incorporated into these dashboards to allow us to track planning trends and their impact on patient survival and toxicity. Over the past decade, we have continued to make significant strides toward the development of patient-specific plan QA tools by developing models and approaches that incorporate patients' geometric information. OVHs have led to case-based reasoning approaches and the development of machine learning models that can predict entire DVH curves. In recent years, patient planning information, including computed tomography images and contours, has been incorporated into deep learning models to predict three-dimensional dose distributions. With the ability to predict the optimum achievable plan for an individual patient, we can begin to reduce the subjectivity of treatment planning and strive for improved plan quality and consistency.

The radiation oncology community should continue to investigate novel applications of plan QA tools. DVH and dose distribution prediction tools should be incorporated into clinical trial quality assurance to improve protocol compliance and standardize plan quality across institutions to ensure the validity of trial results. Moreover, educational tools should be developed that incorporate dose prediction to address training deficiencies for rare disease sites and to increase the rate at which we can train radiation oncologists in LMICs. These educational tools can assist with the safe implementation of new technologies and treatment approaches. Finally, peer review tools should be developed to help reduce variability in peer review practices and improve their effectiveness. Automated plan QA tools may be able to serve as peer review in small clinics and LMICs where no local radiotherapy quality assurance is available.

Abbreviations

AFRONET: Africa Radiation Oncology Network; AI: Artificial Intelligence; $D_{V\%}$: the least dose that the hottest V% of the volume receives; DVH: dose-volume histogram; LMICs: low- and middle-income countries; IMRT: intensity-modulated radiotherapy; QA: quality assessment; OAR: organ at risk; ORBIT-RT: On-line Real-time Benchmarking Informatics Technology for RadioTherapy; OVH: overlap volume histogram; PTV: planning target volume; RTOG: Radiation Therapy Oncology Group; $V_{D\,\mathrm{Gy}}$: the volume receiving a dose of D Gy or more.

Acknowledgments

We thank Erica Goodoff, Senior Scientific Editor in the Research Medical Library at The University of Texas MD Anderson Cancer Center, for editing this article. M.P.G. was supported by the American Legion Auxiliary Fellowship in Cancer Research, the Cancer Prevention and Research Institute of Texas (RP200395), the Ellen Taylor Goldin Legacy Scholarship, the Linda M. Wells GSBS Outreach Award, and the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Numbers TL1TR003169 and UL1TR003167. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

All authors contributed to this study. All authors gave their final approval.

Competing Interests

Authors are part of the Radiation Planning Assistant Group which receives funding from the National Cancer

Institute, Cancer Prevention and Research Institute of Texas, Wellcome Trust, and Varian Medical Systems.

References

[1] Mayo, C. S., Kessler, M. L., Eisbruch, A., Weyburne, G., Feng, M., Hayman, J. A., et al. (2016). The big data effort in radiation oncology: Data mining or data farming? *Advances in Radiation Oncology*, 1(4), 260-271. https://doi.org/10.1016/j.adro.2016.10.001.

[2] Ge, Y., & Wu, Q. J. (2019). Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. *Medical Physics*, 46(6), 2760-2775. https://doi.org/10.1002/mp.13526.

[3] Moore, K. L. (2019). Automated radiotherapy treatment planning. *Seminars in Radiation Oncology*, 29(3), 209-218. https://doi.org/10.1016/j.semradonc.2019.02.003.

[4] Netherton, T. J., Cardenas, C. E., Rhee, D. J., Court, L. E., & Beadle, B. M. (2021). The emergence of artificial intelligence within radiation oncology treatment planning. *Oncology*, 99(2), 124-134. https://doi.org/10.1159/000512172.

[5] Wang, M., Zhang, Q., Lam, S., Cai, J., & Yang, R. (2020). A review on application of deep learning algorithms in external beam radiotherapy automated treatment planning. *Frontiers in Oncology*, 10, 580919. https://doi.org/10.3389/fonc.2020.580919.

[6] Cardenas, C. E., Yang, J., Anderson, B. M., Court, L. E., & Brock, K. B. (2019). Advances in auto-segmentation. *Seminars in Radiation Oncology*, 29(3), 185-197. https://doi.org/10.1016/j.semradonc.2019.02.001.

[7] Court, L. E., Kisling, K., McCarroll, R., Zhang, L., Yang, J., Simonds, H., et al. (2018). Radiation Planning Assistant – a streamlined, fully automated radiotherapy treatment planning system. *Journal of Visualized Experiments*, (134), https://doi.org/10.3791/57411.

[8] McNutt, T. R., Moore, K. L., Wu, B., & Wright, J. L. (2019). Use of big data for quality assurance in radiation therapy. *Seminars in Radiation Oncology*, 29(4), 326-332. https://doi.org/10.1016/j.semradonc.2019.05.006.

[9] Kalet, A. M., Luk, S. M. H., & Phillips, M. H. (2020). Radiation therapy quality assurance tasks and tools: The many roles of machine learning. *Medical Physics*, 47(5), e168-e177. https://doi.org/10.1002/mp.13445.

[10] Moore, K. L., Brame, R. S., Low, D. A., & Mutic, S. (2012). Quantitative metrics for assessing plan quality. *Seminars in Radiation Oncology*, 22(1), 62-69. https://doi.org/10.1016/j.semradonc.2011.09.005.

[11] Special considerations regarding absorbed-dose and dose-volume prescribing and reporting in IMRT (2010). *Journal of the ICRU*, 10(1), 27-40. https://doi.org/10.1093/jicru/ndq008.

[12] Dieterich, S., Ford, E., Pavord, D., & Zeng, J. (2016). Treatment planning and quality metrics. *Practical Radiation Oncology Physics* (pp. 189-206). Elsevier. https://doi.org/10.1016/B978-0-323-26209-5.00014-6.

[13] Holmes, T., Das, R., Low, D., Yin, F., Balter, J., Palta, J., et al. (2009). American society of radiation oncology recommendations for documenting intensity-modulated radiation therapy treatments. *International Journal of Radiation Oncology - Biology - Physics*, 74(5), 1311-1318. https://doi.org/10.1016/j.ijrobp.2009.04.037.

[14] Landberg, T., Chavaudra, J., Dobbs, J., Gerard, J.-P., Hanks, G., Horiot, J.-C., et al. (1999). ICRU report 62: Prescribing, recording and reporting photon beam therapy (supplement to ICRU report 50).

[15] Feuvret, L., Noel, G., Mazeron, J. J., & Bey, P. (2006). Conformity index: A review. *International Journal of Radiation Oncology - Biology - Physics*, 64(2), 333-342. https://doi.org/10.1016/j.ijrobp.2005.09.028.

[16] Kataria, T., Sharma, K., Subramani, V., Karrthick, K. P., & Bisht, S. S. (2012). Homogeneity index: An objective tool for assessment of conformal radiation treatments. *Journal of Medical Physics*, 37(4), 207-213. https://doi.org/10.4103/0971-6203.103606.

[17] Marks, L. B., Yorke, E. D., Jackson, A., Ten Haken, R. K., Constine, L. S., Eisbruch, A., et al. (2010). Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology - Biology - Physics*, 76(3 Suppl), S10-19. https://doi.org/10.1016/j.ijrobp.2009.07.1754.

[18] Bentzen, S. M., Constine, L. S., Deasy, J. O., Eisbruch, A., Jackson, A., Marks, L. B., et al. (2010). Quantitative analyses of normal tissue effects in the clinic (QUANTEC): An introduction to the scientific issues. *International Journal of Radiation Oncology - Biology - Physics*, 76(3 Suppl), S3-9. https://doi.org/10.1016/j.ijrobp.2009.09.040.

[19] Allen Li, X., Alber, M., Deasy, J. O., Jackson, A., Ken Jee, K. W., Marks, L. B., et al. (2012). The use and qa of biologically related models for treatment planning: Short report of the TG-166 of the therapy physics committee of the AAPM. *Medical Physics*, 39(3), 1386-1409. https://doi.org/10.1118/1.3685447.

[20] Mayo, C. S., Phillips, M., McNutt, T. R., Palta, J., Dekker, A., Miller, R. C., et al. (2018). Treatment data and technical process challenges for practical big data efforts in radiation oncology. *Medical Physics*, 45(10), e793-e810. https://doi.org/10.1002/mp.13114.

[21] Matuszak, M. M., Fuller, C. D., Yock, T. I., Hess, C. B., McNutt, T., Jolly, S., et al. (2018). Performance/outcomes data and physician process challenges for practical big data efforts in radiation oncology. *Medical Physics*, 45(10), e811-e819. https://doi.org/10.1002/mp.13136.

[22] Mayo, C. S., Yao, J., Eisbruch, A., Balter, J. M., Litzenberg, D. W., Matuszak, M. M., et al. (2017). Incorporating big data into treatment plan evaluation: Development of statistical DVH metrics and visualization dashboards. *Advances in Radiation Oncology*, 2(3), 503-514. https://doi.org/10.1016/j.adro.2017.04.005.

[23] Wu, B., Ricchetti, F., Sanguineti, G., Kazhdan, M., Simari, P., Chuang, M., et al. (2009). Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Medical Physics*, 36(12), 5497-5505. https://doi.org/10.1118/1.3253464.

[24] Zhu, X., Ge, Y., Li, T., Thongphiew, D., Yin, F. F., & Wu, Q. J. (2011). A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Medical Physics*, 38(2), 719-726. https://doi.org/10.1118/1.3539749.

[25] Moore, K. L., Brame, R. S., Low, D. A., & Mutic, S. (2011). Experience-based quality control of clinical intensity-modulated radiotherapy planning. *International Journal of Radiation Oncology - Biology - Physics*, 81(2), 545-551. https://doi.org/10.1016/j.ijrobp.2010.11.030.

[26] Yuan, L., Ge, Y. R., Lee, W. R., Yin, F. F., Kirkpatrick, J. P., & Wu, Q. J. (2012). Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Medical Physics*, 39(11), 6868-6878. https://doi.org/10.1118/1.4757927.

[27] Zhang, J., Sheng, Y., Wolf, J., Kayode, O., Bradley, J., Ge, Y., et al. (2022). Technical note: Determining the applicability of a clinical knowledge-based learning model via prospective outlier detection. *Medical Physics*, 49(4), 2193-2202. https://doi.org/10.1002/mp.15516.

[28] Appenzoller, L. M., Michalski, J. M., Thorstad, W. L., Mutic, S., & Moore, K. L. (2012). Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Medical Physics*, 39(12), 7446-7461. https://doi.org/10.1118/1.4761864.

[29] Shiraishi, S., Tan, J., Olsen, L. A., & Moore, K. L. (2015). Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Medical Physics*, 42(2), 908. https://doi.org/10.1118/1.4906183.

[30] Ahmed, S., Nelms, B., Gintz, D., Caudell, J., Zhang, G., Moros, E. G., et al. (2017). A method for a priori estimation of best feasible DVH for organs-at-risk: Validation for head and neck VMAT planning. *Medical Physics*, 44(10), 5486-5497. https://doi.org/10.1002/mp.12500.

[31] Covele, B. M., Puri, K. S., Kallis, K., Murphy, J. D., & Moore, K. L. (2021). ORBIT-RT: A real-time, open platform for knowledge-based quality control of radiotherapy treatment planning. *JCO Clinical Cancer Informatics*, 5, 134-142. https://doi.org/10.1200/CCI.20.00093.

[32] Tol, J. P., Dahele, M., Delaney, A. R., Slotman, B. J., & Verbakel, W. F. (2015). Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? *Radiation Oncology*, 10, 234. https://doi.org/10.1186/s13014-015-0542-1.

[33] Cao, W., Gronberg, M., Olanrewaju, A., Whitaker, T., Hoffman, K., Cardenas, C., et al. (2022). Knowledge-based planning for the radiation therapy treatment plan quality assurance for patients with head and neck cancer. *Journal of Applied Clinical Medical Physics*, 23(6), e13614. https://doi.org/10.1002/acm2.13614.

[34] Covele, B. M., Carroll, C. J., & Moore, K. L. (2021). A practical method to quantify knowledge-based DVH prediction accuracy and uncertainty with reference cohorts. *Journal of Applied Clinical Medical Physics*, 22(3), 279-284. https://doi.org/10.1002/acm2.13199.

[35] Nguyen, D., Jia, X., Sher, D., Lin, M. H., Iqbal, Z., Liu, H., et al. (2019). 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. *Physics in Medicine and Biology*, 64(6), 065020. https://doi.org/10.1088/1361-6560/ab039b.

[36] Fan, J., Wang, J., Chen, Z., Hu, C., Zhang, Z., & Hu, W. (2019). Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. *Medical Physics*, 46(1), 370-381. https://doi.org/10.1002/mp.13271.

[37] Chen, X., Men, K., Li, Y., Yi, J., & Dai, J. (2019). A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning. *Medical Physics*, 46(1), 56-64. https://doi.org/10.1002/mp.13262.

[38] Liu, S., Zhang, J., Li, T., Yan, H., & Liu, J. (2021). Technical note: A cascade 3D U-net for dose prediction in radiotherapy. *Medical Physics*, 48(9), 5574-5582. https://doi.org/10.1002/mp.15034.

[39] Liu, Z., Fan, J., Li, M., Yan, H., Hu, Z., Huang, P., et al. (2019). A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy. *Medical Physics*, 46(5), 1972-1983. https://doi.org/10.1002/mp.13490.

[40] Yue, M., Xue, X., Wang, Z., Lambo, R. L., Zhao, W., Xie, Y., et al. (2022). Dose prediction via distance-guided deep learning: Initial development for nasopharyngeal carcinoma radiotherapy. *Radiotherapy and Oncology*, 170, 198-204. https://doi.org/10.1016/j.radonc.2022.03.012.

[41] Zimmermann, L., Faustmann, E., Ramsl, C., Georg, D., & Heilemann, G. (2021). Technical note: Dose prediction for radiation therapy using feature-based losses and one cycle learning. *Medical Physics*, 48(9), 5562-5566. https://doi.org/10.1002/mp.14774.

[42] Gronberg, M. P., Gay, S. S., Netherton, T. J., Rhee, D. J., Court, L. E., & Cardenas, C. E. (2021). Technical note: Dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-net architecture. *Medical Physics*, 48(9), 5567-5573. https://doi.org/10.1002/mp.14827.

[43] Ahn, S. H., Kim, E., Kim, C., Cheon, W., Kim, M., Lee, S. B., et al. (2021). Deep learning method for prediction of patient-specific dose distribution in breast cancer. *Radiation Oncology*, 16(1), 154. https://doi.org/10.1186/s13014-021-01864-9.

[44] Bakx, N., Bluemink, H., Hagelaar, E., van der Sangen, M., Theuws, J., & Hurkmans, C. (2021). Development and evaluation of radiotherapy deep learning dose prediction models for breast cancer. *Physics and Imaging in Radiation Oncology*, 17, 65-70. https://doi.org/10.1016/j.phro.2021.01.006.

[45] Hedden, N., & Xu, H. (2021). Radiation therapy dose prediction for left-sided breast cancers using two-dimensional and three-dimensional deep learning models. *Physica Medica*, 83, 101-107. https://doi.org/10.1016/j.ejmp.2021.02.021.

[46] Zhang, J., Liu, S., Li, T., Mao, R., Du, C., & Liu, J. (2019) Voxel-level radiotherapy dose prediction using densely connected network with dilated convolutions. *Artificial Intelligence in Radiation Therapy*, 70-77.

[47] Barragan-Montero, A. M., Thomas, M., Defraene, G., Michiels, S., Haustermans, K., Lee, J. A., et al. (2021). Deep learning dose prediction for IMRT of esophageal cancer: The effect of data quality and quantity on model performance. *Physica Medica*, 83, 52-63. https://doi.org/10.1016/j.ejmp.2021.02.026.

[48] Barragan-Montero, A. M., Nguyen, D., Lu, W., Lin, M. H., Norouzi-Kandalan, R., Geets, X., et al. (2019). Three-dimensional dose prediction for lung IMRT patients with deep neural networks: Robust learning from heterogeneous beam configurations. *Medical Physics*, 46(8), 3679-3691. https://doi.org/10.1002/mp.13597.

[49] Guerreiro, F., Seravalli, E., Janssens, G. O., Maduro, J. H., Knopf, A. C., Langendijk, J. A., et al. (2021). Deep learning prediction of proton and photon dose distributions for paediatric abdominal tumours. *Radiotherapy and Oncology*, 156, 36-42. https://doi.org/10.1016/j.radonc.2020.11.026.

[50] Qilin, Z., Peng, B., Ang, Q., Weijuan, J., Ping, J., Hongqing, Z., et al. (2022). The feasibility study on the generalization of deep learning dose prediction model for volumetric modulated arc therapy of cervical cancer. *Journal of Applied Clinical Medical Physics*, 23(6), e13583. https://doi.org/10.1002/acm2.13583.

[51] Kearney, V., Chan, J. W., Haaf, S., Descovich, M., & Solberg, T. D. (2018). DoseNet: A volumetric dose prediction algorithm using 3d fully-convolutional neural networks. *Physics in Medicine and Biology*, 63(23), 235022. https://doi.org/10.1088/1361-6560/aaef74.

[52] Lempart, M., Benedek, H., Jamtheim Gustafsson, C., Nilsson, M., Eliasson, N., Back, S., et al. (2021). Volumetric modulated arc therapy dose prediction and deliverable treatment plan generation for prostate cancer patients using a densely connected deep learning model. *Physics and Imaging in Radiation Oncology*, 19, 112-119. https://doi.org/10.1016/j.phro.2021.07.008.

[53] Kandalan, R. N., Nguyen, D., Rezaeian, N. H., Barragan-Montero, A. M., Breedveld, S., Namuduri, K., et al. (2020). Dose prediction with deep learning for prostate cancer radiation therapy: Model adaptation to different treatment planning practices. *Radiotherapy and Oncology*, 153, 228-235. https://doi.org/10.1016/j.radonc.2020.10.027.

[54] Song, Y., Hu, J., Liu, Y., Hu, H., Huang, Y., Bai, S., et al. (2020). Dose prediction using a deep neural network for accelerated planning of rectal cancer radiotherapy. *Radiotherapy and Oncology*, 149, 111-116. https://doi.org/10.1016/j.radonc.2020.05.005.

[55] Mashayekhi, M., Tapia, I. R., Balagopal, A., Zhong, X., Barkousaraie, A. S., McBeth, R., et al. (2022). Site-agnostic 3D dose distribution prediction with deep learning neural networks. *Medical Physics*, 49(3), 1391-1406. https://doi.org/10.1002/mp.15461.

[56] Nguyen, D., Sadeghnejad Barkousaraie, A., Bohara, G., Balagopal, A., McBeth, R., Lin, M. H., et al. (2021). A comparison of Monte Carlo dropout and bootstrap aggregation on the performance and uncertainty estimation in radiation therapy dose prediction with deep learning neural networks. *Physics in Medicine and Biology*, 66(5), 054002. https://doi.org/10.1088/1361-6560/abe04f.

[57] Globocan 2018: Counting the toll of cancer (2018). *The Lancet*, 392(10152). https://doi.org/10.1016/s0140-6736(18)32252-9.

[58] Weber, D. C., Tomsej, M., Melidis, C., & Hurkmans, C. W. (2012). QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy. *Radiotherapy and Oncology*, 105(1), 4-8. https://doi.org/10.1016/j.radonc.2012.08.008.

[59] Ohri, N., Shen, X., Dicker, A. P., Doyle, L. A., Harrison, A. S., & Showalter, T. N. (2013). Radiotherapy protocol deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials. *Journal National Cancer Institute*, 105(6), 387-393. https://doi.org/10.1093/jnci/djt001.

[60] Zhong, H., Men, K., Wang, J., van Soest, J., Rosenthal, D., Dekker, A., et al. (2019). The impact of clinical trial quality assurance on outcome in head and neck radiotherapy treatment. *Frontiers in Oncology*, 9, 792. https://doi.org/10.3389/fonc.2019.00792.

[61] Abrams, R. A., Winter, K. A., Regine, W. F., Safran, H., Hoffman, J. P., Lustig, R., et al. (2012). Failure to adhere to protocol specified radiation therapy guidelines was associated with decreased survival in RTOG 9704--a phase iii trial of adjuvant chemotherapy and chemoradiotherapy for patients with resected adenocarcinoma of the pancreas. *International Journal of Radiation Oncology - Biology - Physics*, 82(2), 809-816. https://doi.org/10.1016/j.ijrobp.2010.11.039.

[62] Crane, C. H., Winter, K., Regine, W. F., Safran, H., Rich, T. A., Curran, W., et al. (2009). Phase II study of bevacizumab with concurrent capecitabine and radiation followed by maintenance gemcitabine and bevacizumab for locally advanced pancreatic cancer: Radiation therapy oncology group RTOG 0411. *Journal of Clinical Oncology*, 27(25), 4096-4102. https://doi.org/10.1200/JCO.2009.21.8529.

[63] Duhmke, E., Franklin, J., Pfreundschuh, M., Sehlen, S., Willich, N., Ruhl, U., et al. (2001). Low-dose radiation is sufficient for the noninvolved extended-field treatment in favorable early-stage Hodgkin's disease: Long-term results of a randomized trial of radiotherapy alone. *Journal of Clinical Oncology*, 19(11), 2905-2914. https://doi.org/10.1200/JCO.2001.19.11.2905.

[64] Eisbruch, A., Harris, J., Garden, A. S., Chao, C. K., Straube, W., Harari, P. M., et al. (2010). Multi-institutional trial of accelerated hypofractionated intensity-modulated radiation therapy for early-stage oropharyngeal cancer (RTOG 00-22). *International Journal of Radiation Oncology - Biology - Physics*, 76(5), 1333-1338. https://doi.org/10.1016/j.ijrobp.2009.04.011.

[65] Peters, L. J., O'Sullivan, B., Giralt, J., Fitzgerald, T. J., Trotti, A., Bernier, J., et al. (2010). Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. *Journal of Clinical Oncology*, 28(18), 2996-3001. https://doi.org/10.1200/JCO.2009.27.4498.

[66] Perez, C. A., Stanley, K., Grundy, G., Hanson, W., Rubin, P., Kramer, S., et al. (1982). Impact of irradiation technique and tumor extent in tumor control and survival of patients with unresectable non-oat cell carcinoma of the lung: Report by the radiation therapy oncology group. *Cancer*, 50(6), 1091-1099. https://doi.org/10.1002/1097-0142(19820915)50:6<1091::aid-cncr2820500612>3.0.co;2-0.

[67] White, J. E., Chen, T., McCracken, J., Kennedy, P., Seydel, H. G., Hartman, G., et al. (1982). The influence of radiation therapy quality control on survival, response and sites of relapse in oat cell carcinoma of the lung: Preliminary report of a southwest oncology group study. *Cancer*, 50(6), 1084-1090. https://doi.org/10.1002/1097-0142(19820915)50:6<1084::aid-cncr2820500611>3.0.co;2-w.

[68] Taylor, R. E., Donachie, P. H., Weston, C. L., Robinson, K. J., Lucraft, H., Saran, F., et al. (2009). Impact of radiotherapy parameters on outcome for patients with supratentorial primitive neuro-ectodermal tumours entered into the siop/ukccsg pnet 3 study. *Radiotherapy and Oncology*, 92(1), 83-88. https://doi.org/10.1016/j.radonc.2009.02.017.

[69] Carrie, C., Hoffstetter, S., Gomez, F., Moncho, V., Doz, F., Alapetite, C., et al. (1999). Impact of targeting deviations on outcome in medulloblastoma: Study of the french society of pediatric oncology (sfop). *International Journal of Radiation Oncology - Biology - Physics*, 45(2), 435-439. https://doi.org/10.1016/s0360-3016(99)00200-x.

[70] Donaldson, S. S., Torrey, M., Link, M. P., Glicksman, A., Gilula, L., Laurie, F., et al. (1998). A multidisciplinary study investigating radiotherapy in ewing's sarcoma: End results of POG #8346. Pediatric oncology group. *International Journal of Radiation Oncology - Biology - Physics*, 42(1), 125-135. https://doi.org/10.1016/s0360-3016(98)00191-6.

[71] Melidis, C., Bosch, W. R., Izewska, J., Fidarova, E., Zubizarreta, E., Ishikura, S., et al. (2014). Radiation therapy quality assurance in clinical trials--global harmonisation group. *Radiotherapy and Oncology*, 111(3), 327-329. https://doi.org/10.1016/j.radonc.2014.03.023.

[72] Fitzgerald, T. J., Bishop-Jodoin, M., Bosch, W. R., Curran, W. J., Followill, D. S., Galvin, J. M., et al. (2013). Future vision for the quality assurance of oncology clinical trials. *Frontiers in Oncology*, 3, 31. https://doi.org/10.3389/fonc.2013.00031.

[73] Gondi, V., Cui, Y., Mehta, M. P., Manfredi, D., Xiao, Y., Galvin, J. M., et al. (2015). Real-time pretreatment review limits unacceptable deviations on a cooperative group radiation therapy technique trial: Quality assurance results of RTOG 0933. *International Journal of Radiation Oncology - Biology - Physics*, 91(3), 564-570. https://doi.org/10.1016/j.ijrobp.2014.10.054.

[74] Moore, K. L., Schmidt, R., Moiseenko, V., Olsen, L. A., Tan, J., Xiao, Y., et al. (2015). Quantifying unnecessary normal tissue complication risks due to suboptimal planning: A secondary study of RTOG 0126. *International Journal of Radiation Oncology - Biology - Physics*, 92(2), 228-235. https://doi.org/10.1016/j.ijrobp.2015.01.046.

[75] Li, N., Carmona, R., Sirak, I., Kasaova, L., Followill, D., Michalski, J., et al. (2017). Highly efficient training, refinement, and validation of a knowledge-based planning quality-control system for radiation therapy clinical trials. *International Journal of Radiation Oncology - Biology - Physics*, 97(1), 164-172. https://doi.org/10.1016/j.ijrobp.2016.10.005.

[76] Accreditation Council for Graduate Medical Education. (2020) Radiation oncology case logs. *Requirements and Process Overview*.

[77] Wuthrick, E. J., Zhang, Q., Machtay, M., Rosenthal, D. I., Nguyen-Tan, P. F., Fortin, A., et al. (2015). Institutional clinical trial accrual volume and survival of patients with head and neck cancer. *Journal of Clinical Oncology*, 33(2), 156-164. https://doi.org/10.1200/JCO.2014.56.5218.

[78] Chien, C. R., Lin, H. W., Yang, C. H., Yang, S. N., Wang, Y. C., Kuo, Y. C., et al. (2011). High case volume of radiation oncologists is associated with better survival

of nasopharyngeal carcinoma patients treated with radiotherapy: A multifactorial cohort analysis. *Clinical Otolaryngology*, 36(6), 558-565. https://doi.org/10.1111/j.1749-4486.2011.02405.x.

[79] Boero, I. J., Paravati, A. J., Xu, B., Cohen, E. E., Mell, L. K., Le, Q. T., et al. (2016). Importance of radiation oncologist experience among patients with head-and-neck cancer treated with intensity-modulated radiation therapy. *Journal of Clinical Oncology*, 34(7), 684-690. https://doi.org/10.1200/JCO.2015.63.9898.

[80] Corry, J., Ng, W. T., Moore, A., Choi, H. C. W., Le, Q., Holmes, S., et al. (2021). Can radiation therapy quality assurance improve nasopharyngeal cancer outcomes in low- and middle-income countries: Reporting the first phase of a prospective International Atomic Energy Agency study. *International Journal of Radiation Oncology - Biology - Physics*, 111(5), 1227-1236. https://doi.org/10.1016/j.ijrobp.2021.08.013.

[81] Mathew, A. (2018). Global survey of clinical oncology workforce. *Journal of Global Oncology*, 4, 1-12. doi:10.1200/JGO.17.00188.

[82] Zubizarreta, E. H., Fidarova, E., Healy, B., & Rosenblatt, E. (2015). Need for radiotherapy in low and middle income countries - the silent crisis continues. *Clinical Oncology*, 27(2), 107-114. https://doi.org/10.1016/j.clon.2014.10.006.

[83] Abdel-Wahab, M., Zubizarreta, E., Polo, A., & Meghzifene, A. (2017). Improving quality and access to radiation therapy-an IAEA perspective. *Seminars in Radiation Oncology*, 27(2), 109-117. https://doi.org/10.1016/j.semradonc.2016.11.001.

[84] Patel, V. (2003). Recruiting doctors from poor countries: The great brain robbery? *BMJ*, 327(7420), 926-928. https://doi.org/10.1136/bmj.327.7420.926.

[85] Misau, Y. A., Al-Sadat, N., & Gerei, A. B. (2010). Brain-drain and health care delivery in developing countries. *Journal of Public Health in Africa*, 1(1), e6. https://doi.org/10.4081/jphia.2010.e6.

[86] Stilwell, B., Diallo, K., Zurn, P., Vujicic, M., Adams, O., & Dal Poz, M. (2004). Migration of health-care workers from developing countries: Strategic approaches to its management. *Bulletin of the World Health Organization*, 82(8), 595-600. https://doi.org/S0042-96862004000800009.

[87] Bundred, P. E., & Levitt, C. (2000). Medical migration: Who are the real losers? *The Lancet*, 356(9225), 245-246. https://doi.org/10.1016/S0140-6736(00)02492-2.

[88] Downes, S., & Ralston, A. (2021). Asia-Pacific special interest group: The first ten years. *Journal of Medical Imaging and Radiation Oncology*, 65(4), 439-444. https://doi.org/10.1111/1754-9485.13259.

[89] Hassan, S., Oar, A., Ward, I., Koh, E. S., Shakespeare, T. P., & Yap, M. L. (2021). Equity should know no borders: The role of Australasian radiation oncologists in supporting radiation oncology services in low- and middle-income countries in the Asia-Pacific. *Journal of Medical Imaging and Radiation Oncology*, 65(4), 410-417. https://doi.org/10.1111/1754-9485.13191.

[90] Rosenblatt, E., Acuna, O., & Abdel-Wahab, M. (2015). The challenge of global radiation therapy: An IAEA perspective. *International Journal of Radiation Oncology - Biology - Physics*, 91(4), 687-689. https://doi.org/10.1016/j.ijrobp.2014.12.008.

[91] Chite Asirwa, F., Greist, A., Busakhala, N., Rosen, B., & Loehrer, P. J., Sr. (2016). Medical education and training: Building in-country capacity at all levels. *Journal of Clinical Oncology*, 34(1), 36-42. https://doi.org/10.1200/JCO.2015.63.0152.

[92] Rosenblatt, E., Prajogi, G. B., Barton, M., Fidarova, E., Eriksen, J. G., Haffty, B., et al. (2017). Need for competency-based radiation oncology education in developing countries. *Creative Education*, 08(01), 66-80. https://doi.org/10.4236/ce.2017.81006.

[93] Eriksen, J. G. (2017). Postgraduate education in radiation oncology in low- and middle-income countries. *Clinical Oncology*, 29(2), 129-134. https://doi.org/10.1016/j.clon.2016.11.004.

[94] Lewis, P. J., Court, L. E., Lievens, Y., & Aggarwal, A. (2021). Structure and processes of existing practice in radiotherapy peer review: A systematic review of the literature. *Clinical Oncology*, 33(4), 248-260. https://doi.org/10.1016/j.clon.2020.10.017.

[95] Brunskill, K., Nguyen, T. K., Boldt, R. G., Louie, A. V., Warner, A., Marks, L. B., et al. (2017). Does peer review of radiation plans affect clinical care? A systematic review of the literature. *International Journal of Radiation Oncology - Biology - Physics*, 97(1), 27-34. https://doi.org/10.1016/j.ijrobp.2016.09.015.

[96] Huo, M., Gorayski, P., Poulsen, M., Thompson, K., & Pinkham, M. B. (2017). Evidence-based peer review for radiation therapy - updated review of the literature with a focus on tumour subsite and treatment modality. *Clinical Oncology*, 29(10), 680-688. https://doi.org/10.1016/j.clon.2017.04.038.

[97] Lawrence, Y. R., Whiton, M. A., Symon, Z., Wuthrick, E. J., Doyle, L., Harrison, A. S., Dicker, A. P. (2012). Quality assurance peer review chart rounds in 2011: A survey of academic institutions in the united states. *International Journal of Radiation Oncology - Biology - Physics*, 84(3), 590-595. https://doi.org/10.1016/j.ijrobp.2012.01.029.

[98] Hoopes, D. J., Johnstone, P. A., Chapin, P. S., Kabban, C. M., Lee, W. R., Chen, A. B., et al. (2015). Practice patterns for peer review in radiation oncology. *Practical Radiation Oncology*, 5(1), 32-38. https://doi.org/10.1016/j.prro.2014.04.004.

[99] Ford, E. C., Terezakis, S., Souranis, A., Harris, K., Gay, H., & Mutic, S. (2012). Quality control quantification (QCQ): A tool to measure the value of quality control checks in radiation oncology. *International Journal of Radiation Oncology - Biology - Physics*, 84(3), e263-269. https://doi.org/10.1016/j.ijrobp.2012.04.036.

[100] Talcott, W. J., Lincoln, H., Kelly, J. R., Tressel, L., Wilson, L. D., Decker, R. H., Ford, E., Hartvigson, P.E., Pawlicki, T. (2020). A blinded, prospective study of error detection during physician chart rounds in radiation oncology. *Practical Radiation Oncology*, 10(5), 312-320. https://doi.org/10.1016/j.prro.2020.05.012.

[101] The Royal College of Radiologists. (2017). Radiotherapy target volume definition and peer review RCR guidance. Available at: https://www.rcr.ac.uk/system/files/publication/field_publication_files/bfco172_peer_review_outlining.pdf.

[102] Marks, L. B., Adams, R. D., Pawlicki, T., Blumberg, A. L., Hoopes, D., Brundage, M. D., et al. (2013). Enhancing the role of case-oriented peer review to improve quality and safety in radiation oncology: Executive summary. *Practical Radiation Oncology*, 3(3), 149-156. https://doi.org/10.1016/j.prro.2012.11.010.

[103] Cox, S., Cleves, A., Clementel, E., Miles, E., Staffurth, J., & Gwynne, S. (2019). Impact of deviations in target volume delineation - time for a new RTQA approach? *Radiotherapy and Oncology*, 137, 1-8. https://doi.org/10.1016/j.radonc.2019.04.012.

[104] Cardenas, C. E., Mohamed, A. S. R., Tao, R., Wong, A. J. R., Awan, M. J., Kuruvila, S., et al. (2017). Prospective qualitative and quantitative analysis of real-time peer review quality assurance rounds incorporating direct physical examination for head and neck cancer radiation therapy. *International Journal of Radiation Oncology - Biology - Physics*, 98(3), 532-540. https://doi.org/10.1016/j.ijrobp.2016.11.019.

[105] Riegel, A. C., Vaccarelli, M., Cox, B. W., Chou, H., Cao, Y., & Potters, L. (2019). Impact of multi-institutional prospective peer review on target and organ-at-risk delineation in radiation therapy. *Practical Radiation Oncology*, 9(2), e228-e235. https://doi.org/10.1016/j.prro.2018.10.016.

[106] Rosenthal, D. I., Asper, J. A., Barker, J. L., Jr., Garden, A. S., Chao, K. S., Morrison, W. H., et al. (2006). Importance of patient examination to clinical quality assurance in head and neck radiation oncology. *Head Neck*, 28(11), 967-973. https://doi.org/10.1002/hed.20446.

[107] Rosenblatt, E., Prasad, R., Hopkins, K., Polo, A., Ntokozo, N., Zaghloul, M., et al. (2018). View of Africa Radiation Oncology Network (AFRONET): An IAEA telemedicine pilot project | journal of the international society for telemedicine and ehealth. *Journal of the International Society for Telemedicine and EHealth*, 6(1). https://doi.org/10.29086/JISfTeH.6.e6.

[108] Vanderpuye, V., Hammad, N., Martei, Y., Hopman, W. M., Fundytus, A., Sullivan, R., et al. (2019). Cancer care workforce in Africa: Perspectives from a global survey. *Infectious Agents and Cancer*, 14, 11. https://doi.org/10.1186/s13027-019-0227-8.

[109] Brandao, M., Guisseve, A., Bata, G., Firmino-Machado, J., Alberto, M., Ferro, J., et al. (2021). Survival impact and cost-effectiveness of a multidisciplinary tumor board for breast cancer in Mozambique, Sub-Saharan Africa. *Oncologist*, 26(6), e996-e1008. https://doi.org/10.1002/onco.13643.