

Beyond auto-segmentation: the case for planning and dosimetry AI in head and neck radiation oncology

Abdulla Elkhadrawy ^{1,2} Sean J Choi ^{1,3} Edward Christopher Dee ⁴ Milit S Patel,⁵ Yingzhi Wu,¹ Nadeem Riaz,¹ Harini Veeraraghavan,⁶ Donghoon Lee,⁶ Sharif Elguindi,⁶ Nancy Y Lee¹

To cite: Elkhadrawy A, Choi SJ, Dee EC, *et al.* Beyond auto-segmentation: the case for planning and dosimetry AI in head and neck radiation oncology. *BMJ Oncology* 2026;5:e001115. doi:10.1136/bmjonc-2026-001115

Received 20 February 2026
Accepted 15 June 2026



© Author(s) (or their employer(s)) 2026. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

¹Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, New York, USA

²Weill Cornell Medical College, New York, New York, USA

³Case Western Reserve University School of Medicine, Cleveland, Ohio, USA

⁴Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, New York, USA

⁵Memorial Sloan Kettering Cancer Center, New York, New York, USA

⁶Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, New York, USA

Correspondence to
Professor Nancy Y Lee;
leen1@mskcc.org and
Edward Christopher Dee;
deee1@mskcc.org

ABSTRACT

Artificial intelligence (AI) has been rapidly integrated into radiation oncology workflows, with the most visible successes occurring in auto-segmentation of target volumes and organs at risk. While these advances have delivered meaningful efficiency gains and improved standardisation, their impact on clinical outcomes in head and neck cancer remains indirect. In head and neck radiotherapy, long-term toxicity and quality of life are influenced not only by contouring accuracy but more directly by nuanced dosimetric trade-offs during treatment planning. This narrative review synthesises the current landscape of AI applications in head and neck radiotherapy planning and highlights the limitations of a segmentation-dominated paradigm. We review the evolution of planning AI from knowledge-based statistical models to deep learning-based dose prediction and emerging reasoning-driven frameworks. We also aim to discuss how these approaches may function as complementary components of a broader planning intelligence ecosystem. In addition to potential patient-level benefits, we examine the system-level implications of planning and dosimetry AI. Finally, we outline key considerations for the safe evaluation and deployment of planning AI. Together, this review positions planning and dosimetry AI as an enabling infrastructure for predictive, adaptive and equitable head and neck radiotherapy.

INTRODUCTION

Artificial intelligence (AI) has been rapidly integrated across multiple stages of the radiation oncology workflow driven by advances in deep learning and increasing amounts of large, annotated imaging data sets. In clinical practice, AI-enabled tools support workflow automation across radiotherapy, including automated segmentation of target volumes and organs at risk (OARs), treatment planning optimisation, and quality assurance processes. This reflects a broader transition towards data-driven radiotherapy, in which complex imaging and clinical data are leveraged to guide contouring, plan generation and evaluation.¹ To date, much of the progress in the space has focused on auto-segmentation

of OARs and more recently target volumes.^{2–8} This emphasis is practical as contouring is lengthy and subject to interobserver variability. Various studies have shown that auto-segmentation can improve efficiency and standardisation, justifying its integration into commercial treatment planning systems with clinician oversight.⁹

Head and neck radiotherapy represents a unique and complex clinical domain. Treatment planning requires steep dose gradients across many critical structures within a confined space. Small variations in planning decisions can have downstream effects on treatment-related toxicity. In addition to significant advancements in auto-segmentation, further improvements in workflow may impact clinical outcomes. In head and neck radiotherapy, interplanner variability persists even when contours are standardised. Additionally, treatment plan quality remains highly dependent on planner expertise and optimisation strategy. These limitations suggest that although auto-segmentation is a critical prerequisite, advances in AI for treatment planning and dosimetry optimisation may offer improvements in both clinical outcomes and system-level performance.

Progress of auto-segmentation: from bottleneck to enabling infrastructure

The rapid progress of auto-segmentation represents a tangible success of AI in radiation oncology. Over the past decade, advances in deep learning architectures, publicly annotated imaging data sets and multi-institutional validation efforts have transformed automated delineation from a research concept to a clinical tool. Commercial deep-learning segmentation tools as well as home-grown systems have been incorporated into departmental workflows after multidisciplinary

evaluation, producing contours that often require only minor adjustment before use in planning.¹⁰ The level of clinical confidence is further reflected in their use within prospective, multicentre clinical trials, where automated contouring has been leveraged not merely as a convenience tool but as a mechanism for standardised practice.

In the TROG V.18.01 NINJA trial, an automated contour quality assurance tool for prostate clinical target volume was used to reduce interobserver variability and manual review burden.¹¹ In a multivendor benchmarking study, Doolan *et al* demonstrated that five commercial auto-segmentation systems achieved good geometric agreement with expert contours (median volumetric Dice similarity coefficients ranging from 0.82 to 0.88) while providing substantial time savings, particularly in head and neck cases where contouring time was reduced by approximately 74–93 min per patient.¹² The HaN-Seg challenge demonstrated robust multi-institutional performance of head and neck OAR segmentation across CT and MRI, supporting generalisability across data sets and imaging platforms.⁴ Complementing this, recent real-world implementation studies such as one conducted by Young *et al* have shown that AI-assisted contouring can achieve favourable dosimetric acceptability and clinician uptake in routine workflows.¹³ Together, these studies extend the role of auto-segmentation beyond workflow acceleration towards clinically scalable deployment.

Consistent with this trial-level validation, widespread real-world adoption has also been reported. For example, survey data from German-speaking countries indicate that a majority of radiation oncologists and medical physicists report using AI-based auto-contouring in routine clinical practice.¹⁴ Approximately two-thirds of respondents reported current use and 90% reported the tools as helpful.¹⁴ In parallel, the scope of automated contouring has expanded substantially within head and neck radiotherapy, evolving from limited OAR delineation to contouring of complex anatomical structures, nodal levels and target volumes.¹⁵

Several factors explain the speed and scale of this progress in auto-segmentation. Contouring is a universal prerequisite of head and neck radiotherapy workflows, as reflected in international consensus guidelines that standardised target and OAR delineation across disease subsites and treatment techniques.¹⁶ From a methodological standpoint, segmentation is well suited for early AI development because segmentation is a bounded task with clearly defined inputs and outputs. This structure enables objective performance evaluation using standardised geometric metrics, such as Dice similarity and surface distance, and has facilitated benchmarking, comparison across studies, and rapid iteration and validation supported by large data sets.^{17 18} Clinically, the contouring process is both time-intensive and repetitive, creating an immediate and easily quantifiable opportunity for efficiency gains. Multiple clinical implementation studies have demonstrated substantial reductions in contouring time, allowing institutions to directly measure

workflow and productivity benefits.^{19 20} Importantly, auto-segmentation outputs remain subject to clinician review prior to treatment delivery, thus allowing for lower regulatory and safety barriers to adoption.²¹ Together, these factors explain why auto-segmentation has emerged as the most mature and widely adopted application of AI in radiation oncology, while also highlighting why its success does not directly translate to more complex planning and dosimetric decision-making tasks.

Radiomics and AI-enabled prognostication in head and neck cancer

Beyond geometric automation, parallel advances in AI have focused on extracting prognostic information from imaging data through radiomics. In head and neck cancer, radiomic analyses have been explored as predictors of treatment response, locoregional control and radiation-induced toxicity, with applications like the prediction of mucositis in patients receiving volumetric modulated radiotherapy.²² These approaches aim to capture imaging features that reflect tumour biology or normal tissue susceptibility.^{23 24} These efforts extend AI from workflow optimisation towards individualised risk stratification. More recently, AI-based prognostication in head and neck oncology has expanded beyond handcrafted radiomic features to include deep learning approaches applied to digital pathology and multimodal data to predict outcomes such as recurrence and survival. For example, Tian *et al* developed a multimodal deep learning model integrating CT, whole-slide histopathology and clinical features from 1087 patients with head and neck squamous cell carcinoma to predict prognosis and postoperative radiotherapy response.²⁵ Emerging studies have demonstrated that integrating imaging, histopathology, genomic and clinical features can improve survival prediction in oral squamous cell carcinoma.²⁶ Although these approaches are not yet used clinically on the same scale as prostate cancer prognostic assays such as the guideline-endorsed ArteraAI Prostate Test (Artera, Los Altos, CA, USA), they highlight the growing role of AI in predicting biological behaviour and radiosensitivity.²⁷

Despite this promise, radiomics-driven prognostication in head and neck radiotherapy remains limited in clinical impact. Many models are retrospective, sensitive to imaging acquisition parameters and difficult to generalise across institutions. For example, radiomics features extracted from pretreatment imaging have been shown to predict the risk of severe radiation-induced xerostomia following intensity-modulated radiotherapy (IMRT).²⁸ These findings remain largely correlative and have not been prospectively translated into treatment decision algorithms. Broader reviews have similarly highlighted methodological limitations related to variability in imaging protocols, limited external validation and reproducibility challenges in head and neck radiomic studies.²⁹ As a result, radiomic predictions are often generated

in isolation, identifying elevated risk without directly informing how treatment planning should be modified.

These limitations underscore a critical gap between prognostic insight and actionable intervention. Recent work illustrates how this gap can be closed in practice. Van den Bosch *et al* developed comprehensive normal tissue complication probability (NTCP) profiles that map dose across 14 OARs to 22 specific toxicity endpoints, quantifying each patient's individual risk landscape prior to planning.³⁰ Van der Laan *et al* then showed these NTCP profiles can be directly used as optimisation objectives, and their quality-of-life-weighted volumetric modulated arc therapy planning approach reduced predicted dysphagia risk by 6%–7% by redistributing dose away from high-risk structures, without compromising target coverage.³¹ This sequence, from patient-specific risk quantification to dose redistribution guided by that risk, demonstrates the mechanism through which prognostic information can modify treatment planning rather than remain an isolated prediction. Prognostic signals derived from imaging, pathology or clinical features are most impactful when coupled to systems capable of operationalising them within the treatment planning process. In this context, planning and dosimetry AI provide a natural interface for integrating prognostic information, enabling risk-aware optimisation strategies that adapt dose distributions to patient-specific anatomy and biological vulnerability.³² Such integration offers a pathway for moving AI-based prognostication beyond risk prediction towards clinically meaningful personalisation of head and neck radiotherapy.

Clinical leverage of planning and dosimetry AI in head and neck radiotherapy

In head and neck radiotherapy, clinical outcomes and long-term quality of life are driven by geometric accuracy and by nuanced dosimetric trade-offs across many critical structures. Even when target volumes and OARs are perfectly delineated, substantial interplanner variability often persists.³³ This variability is due to differences in optimisation priorities, constraint weighting, and compromise strategies that meaningfully affect dose–volume distributions and downstream toxicity. Head and neck cancer is particularly sensitive to subtle dosimetric differences. Small changes in mean dose to the parotid glands, pharyngeal constrictor muscles, laryngeal structures, or oral cavity can affect long-term functional outcomes.³⁴ Planning and dosimetry AI can address this challenge by systematically navigating complex and multiobjective spaces to develop consistent and high-quality plans that balance tumour control with organ preservation.³⁵

In addition, planning AI may also facilitate real-time evaluations of the dosimetric ‘costs’ associated with specific planning decisions, such as more generous clinical target volumes or larger planning target expansions. Specifically, planning AI can model how incremental changes in target volume definition translate to measurable increases in dose to adjacent OARs, enabling clinicians to

weigh the oncologic benefit of broader coverage against the expected toxicity cost before committing to a final plan.³⁶ Commercial tools such as Pinnacle Auto-Planning (Philips Healthcare, Fitchburg, WI, USA), Varian RapidPlan and Ethos Intelligent Optimisation Engine (IOE) (Varian Medical Systems, Palo Alto, CA, USA) and RayStation multicriteria optimisation (RaySearch Laboratories, Stockholm, Sweden) have demonstrated the feasibility of automated or semiautomated planning in complex head and neck cases. In a multi-institutional comparison of five automated treatment planning solutions, Kravensbuehl *et al* showed that multiple commercial auto-planning systems achieved clinically acceptable target coverage and OAR sparing while substantially reducing effective planning time, supporting the maturity of these tools for clinical deployment.³⁷ Beyond knowledge-based and rule-based automation, newer generative AI approaches such as GPT-RadPlan suggest a further evolution towards reasoning-enabled planning systems capable of interfacing with commercial treatment planning environments to automate optimisation decisions.³⁸

In adaptive radiotherapy settings for head and neck cancer, AI-guided reoptimisation has been shown to consistently reduce dose to multiple OARs, including the parotid glands, submandibular glands, oral cavity and larynx, while maintaining comparable target coverage.³⁹ Similarly, studies employing machine-learning-driven optimisation engines within commercial adaptive platforms such as Varian Ethos demonstrate that clinically acceptable plans balance target coverage and OAR sparing without iterative manual replanning.⁴⁰ Critically, not all patients with head and neck cancers benefit equally from adaptive workflows, making patient selection essential for scalable implementation. Mastella *et al* systematically reviewed how AI can automate the CT-based adaptive radiotherapy chain—from synthetic CT generation and OAR segmentation to predictive models that identify candidates most likely to benefit from adaptation, with reported accuracy exceeding 80%.⁴¹ Aristophanous *et al* further supported this with clinical data from a conventional C-arm linear accelerator-based offline adaptive radiotherapy programme, where their custom-developed Automated Watchdog for Adaptive Radiotherapy Environment system monitored volumetric changes using cone-beam CT and deformable image registration.⁴² The system identified patient-specific anatomical triggers for replanning, including a parotid volume reduction of 7% or greater and a nodal gross tumour volume decrease of 29% or greater, both of which were associated with clinically meaningful dosimetric improvements.⁴² These findings underscore that planning AI may support not only adaptive reoptimisation itself but also prospective identification of patients most likely to benefit from adaptive radiotherapy.

Beyond improvements in plan quality and consistency, planning AI can substantially shorten planning timelines, enabling more streamlined and scalable radiotherapy workflows and facilitating earlier treatment initiation with

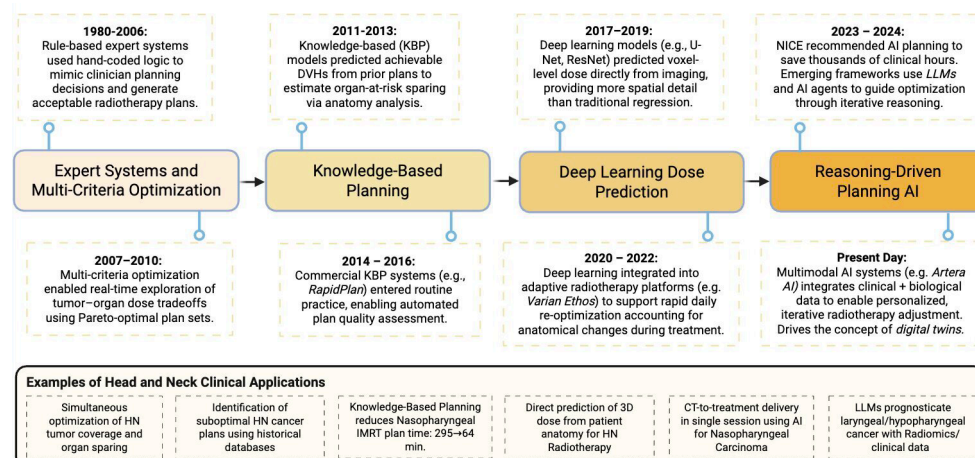


Figure 1 Evolution of radiotherapy planning. AI, artificial intelligence; DVHs, dose-volume histograms; HN, head-and-neck; IMRT, intensity-modulated radiotherapy; KBP, knowledge-based planning; LLMs, large language models; NICE, National Institute for Health and Care Excellence. Created in BioRender. Choi, S. (<https://BioRender.com/tj34of4>) is licensed under CC BY 4.0.

potential clinical benefit for patients with head and neck cancer. This implementation underscores the potential of planning AI to support efficient and reproducible adaptive workflows. Collectively, these studies suggest that planning and dosimetry AI can improve consistency, reduce planning time and achieve clinically acceptable OAR sparing in selected cohorts. However, most published evaluations emphasise dosimetric endpoints, workflow efficiency, or blinded plan preference rather than prospective reductions in toxicity, improvements in patient-reported quality of life, or survival outcomes.^{39 40 43-45} Future prospective multicentre studies and real-world implementation analyses will be necessary to determine whether these technical gains translate into clinically meaningful patient benefit.

Technical roadmap: pattern learning to reasoning-based planning AI

In head and neck radiotherapy, where outcomes are driven by planning trade-offs rather than contouring accuracy alone, recent technical advances in planning AI outline a clear pathway beyond auto-segmentation. As illustrated in [figure 1](#) (Evolution of Radiotherapy Planning), the development of planning AI has progressed from static, database-driven prediction models towards more adaptive and reasoning-oriented systems.⁴⁶⁻⁴⁹ Early efforts in improving outcomes in radiotherapy focused on knowledge-based planning, in which statistical models trained on databases of high-quality previously treated patient plans are used to estimate achievable dose-volume relationships and guide optimisation.⁵⁰

A key limitation of knowledge-based planning is that model performance is inherently dependent on the quality, consistency, and representativeness of the training plan library. Poorly curated or heterogeneous data sets may propagate suboptimal planning strategies rather than encode expert-level performance. This dependency has been demonstrated directly in head and

neck radiotherapy, where iterative retraining of RapidPlan models using improved model-generated outputs produced stronger regression performance and reduced mean dose to parallel OARs such as the parotid glands, suggesting that model quality can improve through deliberate maintenance strategies rather than remain fixed at deployment.⁵¹ Importantly, this limitation may therefore be manageable rather than prohibitive, provided institutions invest in high-quality plan libraries, periodic retraining and ongoing model refinement.⁵² These considerations help explain why later approaches sought methods capable of moving beyond static database-derived priors.

These approaches were subsequently extended by neural network-based methods capable of learning more complex spatial dose patterns and optimisation outcomes directly from data.⁵³ While such models address the limitations of handcrafted features and improve consistency in dose prediction, they primarily operate as static predictors rather than decision-making agents.³⁶

However, deep learning-based dose prediction introduces its own limitations, particularly related to generalisability across institutions and patient populations. A recent systematic review and meta-analysis found substantial heterogeneity across studies ($I^2 > 99\%$), with performance varying by radiotherapy technique, network architecture, and disease subsite, highlighting the instability of results across settings.⁵⁴ Similarly, external validation studies have demonstrated meaningful performance degradation when models are transferred across institutions, although transfer learning using small external data sets has shown partial restoration of model performance, suggesting a practical pathway for cross-institutional adaptation rather than a fundamental barrier to deployment.⁵⁵ More recent architectures incorporating physics-informed priors have also demonstrated improved generalisability across tumour sites, including

head and neck, without institution-specific retraining.⁵⁶ These findings suggest that while generalisability remains a recognised limitation, it is increasingly being addressed through technical and model adaptation strategies.

These limitations have also been recognised at the field level. The ESTRO-AAPM joint guideline noted that widespread clinical adoption of AI planning models remains constrained by concerns regarding availability, applicability, quality, generalisability, interpretability, and safety, with much of the validation literature remaining retrospective and single-institutional.⁴³ At the same time, emerging multicentre evidence suggests this landscape may be evolving. In a three-institution evaluation, over 80% of AI-generated plans met clinical criteria and 60% were preferred over manual plans in blinded review, supporting the view that AI planning may already provide a high-quality foundation for clinician refinement, even if not yet a replacement for expert oversight.⁴⁵

An underexplored but clinically important concept in treatment planning is the degree of precision needed for contours. Prior work suggests that in some clinical contexts, similar dose distributions can be achieved across a range of target volume definitions, highlighting the importance of robust planning strategies. Ploquin *et al* compared intensity-modulated and three-dimensional conformal plans for oropharyngeal cancer under simulated setup errors and found that even the more geometrically complex IMRT plans maintained acceptable OAR doses despite positional perturbations, demonstrating that dosimetric stability is driven more by the underlying optimisation strategy than by millimetre-level contouring precision.⁵⁷ Understanding and exploiting these tolerance limits represents an important opportunity for planning AI, shifting emphasis away from contour precision towards dose robustness, trade-off navigation and sensitivity analysis.

Established multicriteria optimisation and Pareto navigation frameworks provide one example of how such uncertainty-aware planning can be operationalised in practice.^{46 47} For example, human-like intelligent automatic treatment planning systems incorporate a Virtual Treatment Planner, an AI agent designed to mimic expert dosimetrists by iteratively evaluating plan quality.⁵⁸ Similarly, GPT-RadPlan evaluates dose distributions and dose-volume histograms, generates feedback on unmet objectives and iteratively modifies optimisation parameters to improve plan quality.³⁸ As shown in figure 1, this concept shift reframes treatment planning as an iterative reasoning process rather than a single-shot prediction task. In this context, emerging approaches that incorporate large language models (LLMs) represent a further step in this evolution of planning AI. These systems are able to evaluate quality metrics, assess dosimetric trade-offs and guide optimisation adjustments in a manner analogous to expert human planners.⁵⁹

At present, however, reasoning-driven planning systems remain largely exploratory. Early multimodal planning agents have demonstrated promising feasibility but

validation has generally been limited to small test cohorts, narrow disease settings and single-institution environments.³⁸ Similarly, the NRG Oncology working group has emphasised that integrating AI into clinical trials presents ongoing methodological, regulatory, and implementation challenges requiring further development before widespread adoption.⁴⁴ As such, reasoning-driven systems should presently be viewed as investigational decision-support concepts rather than clinically established planning technologies, with prospective evaluation, external validation, and regulatory clarity remaining active areas of development.

Deep learning-based dose prediction and reasoning-driven planning assistance should be viewed as complementary rather than competing paradigms. Neural networks provide the quantitative foundation for modelling achievable dose trade-offs and spatial dose distributions, whereas LLM-based systems operate at a higher level of abstraction, supporting navigation of these trade-offs through higher-level reasoning.⁶⁰ A near-term and clinically realistic application of reasoning-based planning AI lies in the use of AI agents as information synthesisers rather than autonomous decision makers.⁶¹ By integrating plan quality metrics, dose-toxicity relationships, prior cases, and institutional preferences, such systems could support clinicians in making more informed, transparent, and consistent planning decisions.^{59 62} Together, these developments outline a plausible and technically grounded pathway towards advanced planning AI capable of addressing the optimisation decisions that most directly influence outcomes in head and neck radiotherapy.

Ultimately, these developments outline not a linear succession in which each approach replaces its predecessor but a layered evolution in which each generation addresses recognised limitations of the prior one while introducing new challenges of its own. This framing may more accurately position reasoning-driven planning AI not as a mature endpoint but as an emerging extension of a broader planning intelligence ecosystem.

Designing safe and trustworthy planning AI

As planning and dosimetry AI systems move closer to supporting clinical decision-making, their success will depend not only on technical performance but also on how effectively they are designed to operate safely, transparently and in partnership with human expertise. As outlined in figure 2 (Framework for Safe and Trustworthy Planning and Dosimetry AI), safety in planning AI requires an integrated framework that combines failure awareness, clinician oversight, uncertainty modelling, and structured recovery pathways. Unlike auto-segmentation errors, which are often visually apparent and readily corrected, failures in planning and dosimetric reasoning may be subtle yet clinically consequential.⁶³ This distinction is particularly important in head and neck radiotherapy, where small deviations in dose distribution can meaningfully affect toxicity profiles and long-term quality of life.

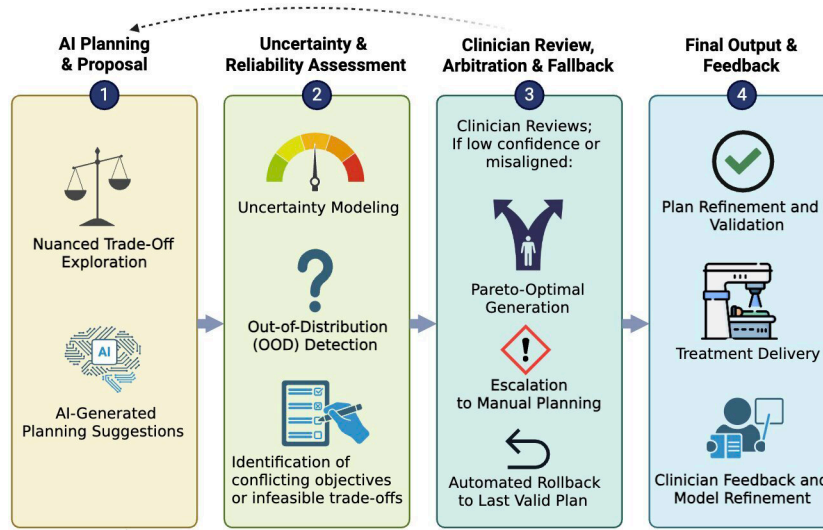


Figure 2 Framework for safe and trustworthy planning and dosimetry AI. AI, artificial intelligence. Created in BioRender. Choi, S. (<https://BioRender.com/svxqq07>) is licensed under CC BY 4.0.

One common failure mode is misprioritisation of competing objectives, such as overemphasising target coverage or a single OAR constraint at the expense of other clinically relevant structures. A second failure mode involves convergence to locally optimal but clinically undesirable solutions. These solutions happen when gradient-based or heuristic optimisation settles on solutions that are mathematically efficient yet misaligned with clinical intent or functional preservation goals.⁶⁴ A further critical risk is overconfidence and silent failure, whereby AI systems generate recommendations without communicating uncertainty or recognising when they are operating outside their domain of reliability. In such cases, errors may propagate downstream without triggering clinician review.⁶⁵

These failure modes underscore the need for clinician-in-the-loop feedback mechanisms as a central design principle for safe and effective planning AI. By enabling clinicians to review, modify, and annotate AI-generated planning suggestions, such systems preserve clinical accountability while supporting bidirectional learning and continuous improvement. Complementary to this approach, uncertainty modelling and out-of-distribution detection are essential components of trustworthy planning AI, enabling systems to flag low-confidence recommendations, defer to human judgement when appropriate, and avoid overconfident extrapolation.⁶⁶ These design principles are also aligned with emerging regulatory expectations for clinical AI, which emphasise human oversight, auditability and accountability in high-risk medical decision support systems. In addition, rigorous commissioning and clinical evaluation testing should be viewed as integral components of this safety framework, ensuring that system performance, failure modes, and boundary conditions are well characterised prior to clinical deployment.^{43 67}

Beyond failure detection, safe planning AI should incorporate explicit recovery pathways triggered by predefined signals.^{43 68} Such signals could be elevated predictive uncertainty, deviation from learnt population-level feasibility bounds or detection of out-of-distribution anatomies or constraint combinations.⁶⁹ Recovery mechanisms may include automated rollback to previously validated plans, generation of Pareto-optimal solutions and structured escalation to clinician review when algorithmic confidence is low or competing objectives cannot be reliably reconciled automatically.⁷⁰ Importantly, such recovery should be understood as a graceful degradation rather than system failure.

A commonly cited concern surrounding advanced automation is the potential for clinical deskilling. In the context of treatment planning, however, this risk is not inevitable but instead reflects design choices. Planning AI that operates as a black-box optimiser may erode engagement, whereas systems that make trade-offs explicit, expose underlying rationale and require clinician arbitration can reinforce expertise rather than replace it.⁷¹ Human–AI collaborative frameworks that emphasise transparency, interpretability, and active clinician participation are therefore more likely to preserve skill development while improving consistency and safety in planning workflows.^{72 73}

At a broader systems level, the safe and equitable deployment of planning and dosimetry AI depends on the representativeness of the data and environments in which these systems are developed and validated. Broader analyses of clinical AI have demonstrated that training data sets, model development and authorship are disproportionately concentrated in high-income countries and large academic centres, raising concerns that AI systems may underperform in underrepresented populations or lower-resource settings.^{74 75} In radiation oncology, these

risks may be amplified by variation in imaging protocols, contouring practices, treatment planning systems and available hardware across institutions and geographic regions. Without diverse external validation and adaptation to local workflows, planning AI may inadvertently perpetuate or exacerbate existing disparities in access to high-quality radiotherapy rather than reduce them. In addition, implementation in resource-limited settings may be constrained by limited digital infrastructure, fragmented data systems and lack of local technical expertise, even when algorithmic performance is strong. Addressing these structural and global limitations will be essential if planning AI is to fulfil its potential as a scalable and equitable component of modern radiotherapy delivery.

System-level value and strategic importance

Beyond individual patient outcomes, investment in planning and dosimetry AI offers system-level advantages across the radiotherapy care continuum. Whereas auto-segmentation primarily improves efficiency at a single upstream step, AI-driven treatment planning has the capacity to streamline the entire planning process.^{76 77} By reducing the number of optimisation cycles, shortening time from simulation to treatment and improving the predictability of plan quality, AI-driven treatment planning has the potential to translate into meaningful improvements in workflow efficiency and departmental throughput. These benefits may be especially relevant in complex disease sites such as head and neck cancer, where planning is highly iterative and resource-intensive.⁷⁸ In a multicentre evaluation across three institutions and five disease sites, a hybrid deep learning-based automated planning system integrating dose prediction with clinical-goal-guided optimisation generated directly deliverable plans within 5 min, with over 80% meeting institutional clinical criteria and 60% preferred over manually optimised plans in blinded review.⁴⁵

Planning and dosimetry AI may also reduce dependence on highly specialised local expertise. Head and neck treatment planning requires extensive experience to balance competing dosimetric objectives across numerous critical structures, and substantial interinstitutional and inter-planner variability persists when standardised guidelines are followed.^{79 80} Chen *et al* found that beyond planning variability, interobserver differences in target volume and OAR delineation among radiation oncologists have been shown to reduce prescription dose coverage and increase doses to OARs, with relative volume differences in target delineation strongly correlating with decreased tumour control probability in nasopharyngeal carcinoma.⁸⁰ By embedding expert-level planning strategies into automated or semiautomated optimisation frameworks, planning AI has potential to improve consistency and quality of care across diverse practice settings, thereby addressing inequities in access to high-quality head and neck radiotherapy.⁸¹

From a strategic perspective, planning AI is also central to the scalability of emerging radiotherapy paradigms.

Adaptive radiotherapy has been increasingly relevant in head and neck cancer due to pronounced anatomical and volumetric changes over the course of treatment. However, widespread adoption of adaptive workflows has been limited by the time and expertise required for repeated plan re-optimisation. AI-driven planning and dosimetry infrastructure enables rapid, reproducible reoptimisation while preserving target coverage and OAR sparing, positioning planning AI not merely as an efficiency tool but as an enabling infrastructure for next-generation radiotherapy delivery.⁸²

Towards integrated and predictive radiotherapy: digital twins and future directions

Sustained investment in planning and dosimetry AI enables a broader reconceptualisation of radiotherapy as a predictive and adaptive process, rather than a static, one-time intervention. A long-term vision emerging from this trajectory is the concept of a *digital twin*: a continuously updated computational representation of a patient that integrates anatomy, delivered dose, biological response and functional outcomes over time.⁸³ Although still aspirational, head and neck cancer represents a particularly compelling use case given the sensitivity of outcomes. Importantly, digital twins cannot arise from segmentation or prognostication alone; they depend on robust planning infrastructure capable of iteratively adjusting treatment. Chaudhuri *et al* demonstrated a predictive digital twin framework for high-grade gliomas that integrates longitudinal imaging with mechanistic modelling and Bayesian calibration to optimise radiotherapy regimens under uncertainty, illustrating the type of robust computational infrastructure needed to support adaptive and personalised treatment planning.⁸⁴ In this context, planning and dosimetry AI function as enabling technologies, providing the decision-making substrate on which more comprehensive predictive models may eventually be built.⁷¹

Complementing this vision, emerging evidence supports the use of unsupervised and self-supervised methods to reveal latent relationships between anatomy, dose and function that are not explicitly encoded in current planning paradigms. Such methods may reveal previously unrecognised patterns of toxicity or recovery, further informing adaptive and personalised treatment strategies. For example, unsupervised clustering of radiotherapy dose data has been shown to identify distinct patient subgroups with differing toxicity profiles that are not explained by standard dose-volume parameters or clinical grading alone.⁸⁵

CONCLUSION

Planning and dosimetry AI have the potential to influence multiple aspects of head and neck radiotherapy due to their ability to synthesise complex clinical, anatomical and dosimetric information. Emerging applications of planning AI in head and neck cancer range from dose

prediction and optimisation support to adaptive replanning and decision-support tools. When coupled with advances in multimodal learning and integration within compound AI frameworks, planning AI systems are increasingly positioned to address long-standing clinical and translational challenges.

However, several challenges remain to the development of planning and dosimetry AI in routine clinical practice. These include ensuring reliability across heterogeneous patient anatomies, maintaining transparency and interpretability in optimisation decisions, safeguarding patient data and addressing ethical and medicolegal considerations associated with AI-assisted decision-making. Validation of planning AI systems therefore requires demonstration of meaningful benefit in real-world clinical settings, with evaluation frameworks that extend beyond geometric accuracy. While current evidence does not support the replacement of clinician expertise in head and neck treatment planning, AI may serve as a valuable decision-support tool under clinician oversight, augmenting rather than supplanting human judgement in complex radiotherapy workflows.

Ultimately, the value of planning AI lies not in replacing clinical expertise but in operationalising it; translating the dosimetric trade-offs, prognostic insights and adaptive strategies discussed throughout this review into reproducible, efficient workflows. By compressing planning timelines from days to minutes and enabling rapid reoptimisation, these systems can shorten time to treatment initiation while freeing clinical staff to focus on plan review, quality assurance and patient care. This consistency also creates a foundation for meaningful quality assurance metrics, because when plans are generated through uniform optimisation frameworks rather than variable manual approaches, institutions can establish benchmarks, track performance across patients and planners, and identify deviations that warrant review. In this way, planning and dosimetry AI can tie together the advances in segmentation, prognostication and adaptive delivery discussed in this review, moving head and neck radiotherapy closer to a more consistent, accurate, accessible and individualised standard of care.

Contributors AE contributed to the conception and design of the review, conducted the literature review and drafted the majority of the manuscript. SJC contributed to conducting the literature review, critical revision of the manuscript and the research, conceptualisation and creation of the figures. ECD contributed to the conception and design of the review, development of the central idea and critical revision of the manuscript. MSP, YW, NR, HV, DL, SE and NYL contributed to review and revision of the manuscript and refinement of the ideas presented. NYL provided senior oversight and supervision as principal investigator. All authors critically reviewed the manuscript for important intellectual content, approved the final version submitted for publication and agreed to be accountable for all aspects of the work. AE, SJC, ECD and NYL are the guarantors of this paper.

Funding All authors are funded in part through the Cancer Center Support Grant from the National Cancer Institute (P30 CA008748). ECD is funded in part through the Prostate Cancer Foundation Young Investigator Award.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Abdulla Elkhadrawy <https://orcid.org/0009-0009-9108-0723>

Sean J Choi <https://orcid.org/0009-0005-4034-9551>

Edward Christopher Dee <https://orcid.org/0000-0001-6119-0889>

REFERENCES

- Huynh E, Hosny A, Guthrie C, et al. Artificial intelligence in radiation oncology. *Nat Rev Clin Oncol* 2020;17:771–81.
- Iyer A, Thor M, Onochie I, et al. Prospectively-validated deep learning model for segmenting swallowing and chewing structures in CT. *Phys Med Biol* 2022;67:024001.
- Jiang J, Elguindi S, Berry SL, et al. Nested block self-attention multiple resolution residual network for multiorgan segmentation from CT. *Med Phys* 2022;49:5244–57.
- Podobnik G, Ibragimov B, Tappeiner E, et al. HaN-Seg: The head and neck organ-at-risk CT and MR segmentation challenge. *Radiother Oncol* 2024;198:110410.
- Reinders FCJ, Savenije MHF, de Ridder M, et al. Automatic segmentation for magnetic resonance imaging guided individual elective lymph node irradiation in head and neck cancer patients. *Phys Imaging Radiat Oncol* 2024;32:100655.
- Paudyal R, Jiang J, Han J, et al. Auto-segmentation of neck nodal metastases using self-distilled masked image transformer on longitudinal MR images. *BJR Artif Intell* 2024;1:ubae004.
- Wei Z, Ren J, Eriksen JG, et al. An interactive deep-learning workflow for head and neck gross tumour volume segmentation. *Phys Imaging Radiat Oncol* 2025;35:100820.
- Dong Y, Gong K. Head and neck tumor segmentation from [18F] F-FDG PET/CT images based on 3D diffusion model. *Phys Med Biol* 2024;69:155001.
- van Dijk LV, Van den Bosch L, Aljabar P, et al. Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring. *Radiother Oncol* 2020;142:115–23.
- Heilemann G, Buschmann M, Lechner W, et al. Clinical Implementation and Evaluation of Auto-Segmentation Tools for Multi-Site Contouring in Radiotherapy. *Phys Imaging Radiat Oncol* 2023;28:100515.
- Chlap P, Min H, Martin J, et al. Implementation of an automated contour quality assurance tool within the TROG 18.01 NINJA trial. *Radiother Oncol* 2026;214:111269.
- Doolan PJ, Charalambous S, Roussakis Y, et al. A clinical evaluation of the performance of five commercial artificial intelligence contouring systems for radiotherapy. *Front Oncol* 2023;13:1213068.
- Young T, Butterworth V, Misson S, et al. Real-world performance evaluation of commercial autocontouring software for head and neck cancer radiotherapy. *Br J Radiol* 2025;98:1632–41.
- Vorbach SM, Putz F, Ganswindt U, et al. Contouring in transition: perceptions of AI-based autocontouring by radiation oncologists and medical physicists in German-speaking countries. *Strahlenther Onkol* 2025;201:1151–61.
- Vrtovec T, Močnik D, Strojani P, et al. Auto-segmentation of organs at risk for head and neck radiotherapy planning: From atlas-based to deep learning methods. *Med Phys* 2020;47:e929–50.
- Brouwer CL, Steenbakkers RJHM, Bourhis J, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. *Radiother Oncol* 2015;117:83–90.
- Cardenas CE, Yang J, Anderson BM, et al. Advances in Auto-Segmentation. *Semin Radiat Oncol* 2019;29:185–97.
- Kosmin M, Ledsam J, Romera-Paredes B, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. *Radiother Oncol* 2019;135:130–40.
- Wong J, Huang V, Wells D, et al. Implementation of deep learning-based auto-segmentation for radiotherapy planning structures: a workflow study at two cancer centers. *Radiat Oncol* 2021;16:101.

- 20 Gibbons E, Hoffmann M, Westhuyzen J, *et al.* Clinical evaluation of deep learning and atlas-based auto-segmentation for critical organs at risk in radiation therapy. *J Med Radiat Sci* 2023;70 Suppl 2:15–25.
- 21 Wiens J, Saria S, Sendak M, *et al.* Author Correction: Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019;25:1627.
- 22 Alabi RO, Elmusrati M, Leivo I, *et al.* Artificial Intelligence-Driven Radiomics in Head and Neck Cancer: Current Status and Future Prospects. *Int J Med Inform* 2024;188:105464.
- 23 Oh JH, Apte AP, Katsoulakis E, *et al.* Reproducibility of radiomic features using network analysis and its application in Wasserstein k-means clustering. *J Med Imag* 2021;8:031904.
- 24 Oh JH, Apte A, Veeraraghavan H, *et al.* Radiomic clustering using graph network techniques coupled with unbalanced optimal transport. *Comput Struct Biotechnol J* 2025;27:4772–8.
- 25 Tian R, Hou F, Zhang H, *et al.* Multimodal fusion model for prognostic prediction and radiotherapy response assessment in head and neck squamous cell carcinoma. *NPJ Digit Med* 2025;8:302.
- 26 Vollmer A, Hartmann S, Vollmer M, *et al.* Multimodal artificial intelligence-based pathogenomics improves survival prediction in oral squamous cell carcinoma. *Sci Rep* 2024;14:5687.
- 27 Spratt DE, Tang S, Sun Y, *et al.* Artificial Intelligence Predictive Model for Hormone Therapy Use in Prostate Cancer. *Res Sq* 2023.
- 28 Iancu RI, Zara AD, Mirestean CC, *et al.* Radiomics in Head and Neck Cancers Radiotherapy. Promises and Challenges. *Maedica (Buchar)* 2021;16:482–8.
- 29 Bang C, Bernard G, Le WT, *et al.* Artificial intelligence to predict outcomes of head and neck radiotherapy. *Clin Transl Radiat Oncol* 2023;39:100590.
- 30 van den Bosch L, van der Schaaf A, van der Laan HP, *et al.* Comprehensive toxicity risk profiling in radiation therapy for head and neck cancer: A new concept for individually optimised treatment. *Radiother Oncol* 2021;157:147–54.
- 31 van der Laan HP, van der Schaaf A, van den Bosch L, *et al.* Quality of life and toxicity guided treatment plan optimisation for head and neck cancer. *Radiother Oncol* 2021;162:85–90.
- 32 Lambin P, Leijenaar RTH, Deist TM, *et al.* Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–62.
- 33 Nelms BE, Robinson G, Markham J, *et al.* Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Pract Radiat Oncol* 2012;2:296–305.
- 34 Eisbruch A, Ten Haken RK, Kim HM, *et al.* Dose, volume, and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer. *Int J Radiat Oncol Biol Phys* 1999;45:577–87.
- 35 Huang W, Liu T, Shen Y, *et al.* Automatic radiotherapy planning for deliverable plans using deep learning dose prediction and dose rings optimization in cervical cancer. *J Appl Clin Med Phys* 2025;26:e70353.
- 36 Wang C, Zhu X, Hong JC, *et al.* Artificial Intelligence in Radiotherapy Treatment Planning: Present and Future. *Technol Cancer Res Treat* 2019;18.
- 37 Krayenbuehl J, Zamburlini M, Ghandour S, *et al.* Planning comparison of five automated treatment planning solutions for locally advanced head and neck cancer. *Radiat Oncol* 2018;13:170.
- 38 Liu S, Pastor-Serrano O, Chen Y, *et al.* Automated radiotherapy treatment planning guided by GPT-4Vision. *Phys Med Biol* 2025;70:155002.
- 39 Domal SJ, Maniscalco A, Visak J, *et al.* Closing the gap in plan quality: Leveraging deep-learning dose prediction for adaptive radiotherapy. *J Appl Clin Med Phys* 2025;26:e70045.
- 40 Visak J, Inam E, Meng B, *et al.* Evaluating machine learning enhanced intelligent-optimization-engine (IOE) performance for ethos head-and-neck (HN) plan generation. *J Appl Clin Med Phys* 2023;24:e13950.
- 41 Mastella E, Calderoni F, Manco L, *et al.* A systematic review of the role of artificial intelligence in automating computed tomography-based adaptive radiotherapy for head and neck cancer. *Phys Imaging Radiat Oncol* 2025;33:100731.
- 42 Aristophanous M, Aliotta E, Lichtenwalner P, *et al.* Clinical Experience With an Offline Adaptive Radiation Therapy Head and Neck Program: Dosimetric Benefits and Opportunities for Patient Selection. *Int J Radiat Oncol Biol Phys* 2024;119:1557–68.
- 43 Hurkmans C, Bibault J-E, Brock KK, *et al.* A joint ESTRO and AAPM guideline for development, clinical validation and reporting of artificial intelligence models in radiation therapy. *Radiother Oncol* 2024;197:110345.
- 44 Jia X, Rong Y, Wu Q, *et al.* NRG Oncology Assessment of Artificial Intelligence for Automatic Treatment Planning in Radiation Therapy Clinical Trials: Present and Future. *Int J Radiat Oncol Biol Phys* 2025;123:282–95.
- 45 Yu L, Ni Q, Wang B, *et al.* Multicenter study on the versatility and adoption of AI-driven automated radiotherapy planning across cancer types. *Nat Commun* 2025;17:867.
- 46 Monz M, Küfer KH, Bortfeld TR, *et al.* Pareto navigation: algorithmic foundation of interactive multi-criteria IMRT planning. *Phys Med Biol* 2008;53:985–98.
- 47 Jee K-W, McShan DL, Fraass BA. Lexicographic ordering: intuitive multicriteria optimization for IMRT. *Phys Med Biol* 2007;52:1845–61.
- 48 Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. *Med Phys* 2019;46:2760–75.
- 49 Chen D, Parsa R, Swanson K, *et al.* Large language models in oncology: a review. *BMJ Oncol* 2025;4:e000759.
- 50 Tol JP, Dahele M, Delaney AR, *et al.* Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? *Radiat Oncol* 2015;10:234.
- 51 Fogliata A, Cozzi L, Reggiori G, *et al.* RapidPlan knowledge based planning: iterative learning process and model ability to steer planning strategies. *Radiat Oncol* 2019;14:187.
- 52 Kaderka R, Hild SJ, Bry VN, *et al.* Wide-Scale Clinical Implementation of Knowledge-Based Planning: An Investigation of Workforce Efficiency, Need for Post-automation Refinement, and Data-Driven Model Maintenance. *Int J Radiat Oncol Biol Phys* 2021;111:705–15.
- 53 Gronberg MP, Gay SS, Netherton TJ, *et al.* Technical Note: Dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-net architecture. *Med Phys* 2021;48:5567–73.
- 54 Zamanian M, Kavehpoor MA, Soltaninejad AM, *et al.* Deep learning for head and neck radiation dose prediction: a systematic review and meta-analysis. *Sci Rep* 2025;15:45766.
- 55 Chu H, de Vette SPM, Neh H, *et al.* Three-Dimensional Deep Learning Normal Tissue Complication Probability Model to Predict Late Xerostomia in Patients With Head and Neck Cancer. *Int J Radiat Oncol Biol Phys* 2025;121:269–80.
- 56 Xiong T, Ren G, Chen Z, *et al.* A generalizable dose prediction model for automatic radiotherapy planning based on physics-informed priors and large-kernel convolutions. *Med Phys* 2026;53:e70272.
- 57 Ploquin N, Lau H, Dunscombe P. Intensity modulated and three-dimensional conformal radiation therapy plans for oropharyngeal cancer: a comparison of their sensitivity to set-up errors and uncertainties. *Curr Oncol* 2006;13:61–6.
- 58 Gao Y, Kyun Park Y, Jia X. Human-like intelligent automatic treatment planning of head and neck cancer radiation therapy. *Phys Med Biol* 2024;69:115049.
- 59 Ferber D, El Nahhas OSM, Wölflein G, *et al.* Development and validation of an autonomous artificial intelligence agent for clinical decision-making in oncology. *Nat Cancer* 2025;6:1337–49.
- 60 Dong Z, Chen Y, Gay H, *et al.* Large-language-model empowered 3D dose prediction for intensity-modulated radiotherapy. *Med Phys* 2025;52:619–32.
- 61 Yap W-K, Cheng S-C, Lin C-H, *et al.* An External Validation Study on Two Pre-Trained Large Language Models for Multimodal Prognostication in Laryngeal and Hypopharyngeal Cancer: Integrating Clinical, Treatment, and Radiomic Data to Predict Survival Outcomes with Interpretable Reasoning. *Bioengineering (Basel)* 2025;12:1345.
- 62 Sendak MP, D'Arcy J, Kashyap S. A Path for Translation of Machine Learning Products into Healthcare Delivery. *EMJ Innov* 2020.
- 63 Boero IJ, Paravati AJ, Xu B, *et al.* Importance of Radiation Oncologist Experience Among Patients With Head-and-Neck Cancer Treated With Intensity-Modulated Radiation Therapy. *J Clin Oncol* 2016;34:684–90.
- 64 Tonneau M, Roos M, Cayez R, *et al.* Multicriteria optimization of radiation therapy: Towards empowerment and standardization of reverse planning for head and neck squamous cell carcinoma. *Cancer Radiother* 2024;28:317–22.
- 65 Cheng M, Nazarian S, Bogdan P. There Is Hope After All: Quantifying Opinion and Trustworthiness in Neural Networks. *Front Artif Intell* 2020;3:54.
- 66 Kompa B, Snoek J, Beam AL. Second opinion needed: communicating uncertainty in medical machine learning. *Npj Digit Med* 2021;4:4.
- 67 Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (artificial intelligence act). European Union; 2024. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689> [Accessed 10 May 2026].
- 68 Wahid KA, Kaffey ZY, Farris DP, *et al.* Artificial intelligence uncertainty quantification in radiotherapy applications - A scoping review. *Radiother Oncol* 2024;201:110542.

- 69 Abdar M, Pourpanah F, Hussain S, *et al.* A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion* 2021;76:243–97.
- 70 Olanrewaju A, Court LE, Zhang L, *et al.* Clinical Acceptability of Automated Radiation Treatment Planning for Head and Neck Cancer Using the Radiation Planning Assistant. *Pract Radiat Oncol* 2021;11:177–84.
- 71 Holzinger A, Biemann C, Pattichis CS, *et al.* What do we need to build explainable AI systems for the medical domain? *arXiv* 2017.
- 72 Amann J, Blasimme A, Vayena E, *et al.* Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020;20:310.
- 73 Holzinger A. Explainable AI and Multi-Modal Causability in Medicine. *I Com (Berl)* 2021;19:171–9.
- 74 Futoma J, Simons M, Panch T, *et al.* The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Health* 2020;2:e489–92.
- 75 Celi LA, Cellini J, Charpignon M-L, *et al.* Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. *PLOS Digit Health* 2022;1:e0000022.
- 76 Vaassen F, Hazelaar C, Vaniqui A, *et al.* Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy. *Phys Imaging Radiat Oncol* 2020;13:1–6.
- 77 Arjmandi N, Mosleh-Shirazi MA, Mohebbi S, *et al.* Evaluating the dosimetric impact of deep-learning-based auto-segmentation in prostate cancer radiotherapy: Insights into real-world clinical implementation and inter-observer variability. *J Appl Clin Med Phys* 2025;26:e14569.
- 78 Cao Y, Kunaprayoon D, Xu J, *et al.* AI-assisted clinical decision making (CDM) for dose prescription in radiosurgery of brain metastases using three-path three-dimensional CNN. *Clin Transl Radiat Oncol* 2023;39:100565.
- 79 Berry SL, Boczkowski A, Ma R, *et al.* Interobserver variability in radiation therapy plan output: Results of a single-institution study. *Pract Radiat Oncol* 2016;6:442–9.
- 80 Chen M, Peng Y, Chen R, *et al.* Interobserver variability in organ delineation on radiotherapy treatment planning for nasopharyngeal carcinoma: A dosimetric and prognostic analysis. *Front Oncol* 2025;15:1510568.
- 81 Krishnamurthy R, Mummudi N, Goda JS, *et al.* Using Artificial Intelligence for Optimization of the Processes and Resource Utilization in Radiotherapy. *JCO Glob Oncol* 2022;8:e2100393.
- 82 Khalifa A, Winter JD, Tadic T, *et al.* Machine learning automated treatment planning for online magnetic resonance guided adaptive radiotherapy of prostate cancer. *Phys Imaging Radiat Oncol* 2024;32:100649.
- 83 Abdollahi H, Yousefirizi F, Shiri I, *et al.* Theranostic digital twins: Concept, framework and roadmap towards personalized radiopharmaceutical therapies. *Theranostics* 2024;14:3404–22.
- 84 Chaudhuri A, Pash G, Hormuth DA II, *et al.* Predictive digital twin for optimizing patient-specific radiotherapy regimens under uncertainty in high-grade gliomas. *Front Artif Intell* 2023;6:1222612.
- 85 Padannayil NM, Sharma DS, Nangia S, *et al.* IMPT of head and neck cancer: unsupervised machine learning treatment planning strategy for reducing radiation dermatitis. *Radiat Oncol* 2023;18:11.