Streamlining the Verification of Radiotherapy Contours by Identifying Clinically Relevant Subsections

by

Choo Run Kang Neville

DKU Medical Physics
Duke University

Defense Date: March 28, 2024

Approved:

Justus Adamson, Supervisor

William Giles

Fang-Fang Yin

Manju Liu

Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in DKU Medical Physics of
Duke University
2024

ABSTRACT

Streamlining the Verification of Radiotherapy Contours by
Identifying Clinically Relevant Subsections

by

Choo Run Kang Neville

DKU Medical Physics
Duke University

Defense Date: March 28, 2024

Approved:

Justus Adamson, Supervisor

William Giles

Fang-Fang Yin

Manju Liu

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in DKU Medical Physics of
Duke University
2024

Abstract

Purpose: New developments in radiation therapy such as AI based contouring and online adaptive radiotherapy have led to an increase in the number of structures and normal tissues being delineated with less human oversight, requiring review in a shorter timeframe. Addressing this problem, we aim to develop and validate a novel method for QA by streamlining the verification of radiotherapy contours. This method identifies subsections of organs at risk that could result in clinically relevant dose constraints being violated, should the contour be inaccurate.

Methods: Structures with planning constraints are evaluated by adding increasing margins and checking against dose constraints. Structures that violate constraints when expanded with margins are flagged for manual review; The smallest margin that violates constraints is prioritized, with the location necessitating manual review presented in a suitable manner, such as the maximum dose point for the contours specified in this study.

We applied this method in a retrospective analysis of 92 stereotactic radiosurgery plans, evaluating brainstem, optic nerves, and optic chiasm. Margins of 0, 1, 3, and 5mm were applied to define risk levels (very high, high, medium, low) for manual review. Contours and associated MR images were independently reviewed by 2 physicists with contouring experience using the locations flagged for review by our method to determine whether clinically relevant contouring errors existed.

Results: 12 contours (brainstem n=10, optic nerve n=1, chiasm n=1) from 11 plans were flagged for manual review at risk levels of very high (n=1), high (n=4), medium (n=2), and low (n=5). Review by the 2 physicists indicated a mean offset value of 0.45mm, with a mean difference of 0.19mm (SD = 0.57mm) and a correlation value of 0.7 between the two sets of

observations. Notably, one case exhibited a mean contouring error of 1.75mm, significantly beyond the standard tolerance for SRS of 1mm, suggesting a critical area of concern.

Conclusion: Our results indicate that the method described here has potential to improve both the efficacy and efficiency of the plan review process. When applied to radiosurgery, efficacy improved as a number of previously unidentified contouring errors were identified in critical locations among the 12% of cases flagged for manual review, suggesting potential to reduce medical errors. Potential improvements in efficiency are highlighted by the 88% of cases for which the tool indicated that contouring errors would not have a clinically relevant dosimetric effect, indicating review is not necessary. Further investigation is warranted to explore the application of this method to other treatment sites.

Contents

List of Figures

Acknowledgements

I would like to extend my thanks to the people who have contributed to this project, and everyone who has supported me in my graduate journey.

Firstly, I want to express my deep gratitude to Dr Adamson for his guidance and patience while taking time from his busy clinical schedule, and for giving me the opportunity to work on this fascinating project for my thesis.

I would also like to show my appreciation for Dr Giles, who provided his valuable expertise and insight into the development of the tool and the project.

I want to also thank Brett and Ke Lu for their help with the assessment of contours that were instrumental in the final analysis, as well as useful feedback for the abstract.

To my friends at Duke Kunshan University and Duke University, it has been an amazing journey together and I wish you all the best wherever you go!

To my parents, thank you for enabling me to pursue my goals. I appreciate all your support in getting me where I am today.

1. Introduction

1.1 Importance of Radiation Therapy

Radiation Therapy is a cornerstone of modern cancer treatment, delivering highly precise doses of radiation to tumors while sparing surrounding healthy tissues. Whether employed as a standalone treatment or in conjunction with other modalities, radiotherapy offers an improvement in treatment outcomes[1]. Compared to more invasive traditional modalities, radiation therapy not only improves the therapeutic ratio but also by substantially reducing the overall treatment burden for patients[2,3]. With evidence-based indications for its use in over 50% of cancer cases[4], continual advances in radiation therapy underscore its ever-growing importance in the role of cancer treatment.

1.1.1 Contemporary Radiation Therapy

1.1.1.1 Treatment Techniques

Contemporary radiotherapy leverages advanced techniques like Intensity-Modulated Radiation Therapy (IMRT)[5,6] and Image-Guided Radiotherapy (IGRT)[7] to achieve more precise dose delivery, reducing unnecessary irradiation of healthy tissues. These techniques facilitate Stereotactic Radiosurgery (SRS) and Stereotactic Body Radiation Therapy (SBRT), which deliver higher doses of radiation in fewer treatment sessions. This approach enhances the biologically effective dose (BED) while protecting surrounding normal tissues, leading to improved treatment outcomes[8–11].

Despite these technological advances, treatment planning remains a highly iterative process. The quality of treatment largely depends on the planners' experience and time invested[12–14], whereby even minor inaccuracies in dose distribution can lead to significant patient

risks such as incomplete treatment or increased toxicity[15–20]. Greater oversight is thus essential to harness these technologies effectively and minimize potential adverse effects.

1.2 Normal Tissue Toxicity

Radiation toxicity refers to the adverse side effects arising from the irradiation of healthy tissues and organs during radiation therapy, which are generally proportional to the dose they receive[21]. Common side effects include pain, bleeding, ulceration, and other organ-specific chronic conditions[6], which may necessitate specialized management or medication[22]. A fundamental challenge in radiation therapy is achieving effective tumor control while minimizing the dose to organs at risk (OARs) to reduce these toxicities[21–23].

Recent advances in radiotherapy, such as IMRT and IGRT, have improved the accuracy of treatments. These technologies allow for more precise and uniform dose delivery to the target volume, thereby decreasing the risk of long-term toxicities[6]. By increasing the number of beam orientations, IMRT and IGRT provide a more variable dose distribution across surrounding tissues, giving planners greater flexibility to selectively irradiate or spare specific areas[24]. This strategic distribution is often guided by established dose constraints, such as those recommended by QUANTEC[25], to optimize the balance between effective tumor treatment and the preservation of healthy tissue.

1.2.1 QUANTEC and Dose Constraints

The Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) offers guidelines that correlate dose and volume with clinical outcomes, aiding physicians in assessing the safety of radiation treatment plans. Dose-volume histograms (DVHs) are pivotal in this

process, despite their assumption of uniform organ function, providing a basis for predicting tissue complications and setting dose constraints[26].

Advances in radiation techniques like SRS and SBRT have prompted the development of HyTEC, a QUANTEC offshoot, which addresses the challenges posed by hypofractionation and the exposure of large volumes of normal tissue to low doses of radiation[3,24].

The need to assess the relative importance of different adverse effects has led to the development of several approaches such as quality of life adjusted tumor control probability and figures of merit combining tumor control probability (TCP) and normal tissue complication probability (NTCP), to balance the risk of toxicities against effective tumor coverage[27].

However, while avoiding toxicities is highly valued, target coverage remains the utmost priority, with normal tissue complications typically preferred over a marginal miss of the target, as the failure to accurately target the tumor can lead to recurrences that are challenging to manage and even potentially life-threatening[26]. To this end, ensuring precise delineation of both tumors and critical organs thus remains paramount.

1.2.2 Contour Delineation and Verification

Contour delineation is a crucial initial step in radiotherapy treatment planning[15], with subsequent dose optimization and analysis relying on the assumption that the contours accurately represent their respective structures[28]. Accurate identification and delineation of OARs is vital for planners to be able to design optimal treatment strategies and limit OAR dose exposure effectively[29].

Inaccuracies in contour delineation can lead to dose-dumping, where normal tissues might receive excessive radiation, resulting in unexpected toxicities[28]. Such risks are heightened by the adoption of sophisticated treatment techniques and inverse optimization algorithms[30].

A consensus paper by the American Society for Radiation Oncology (ASTRO) highlights a problematic trend of errors and significant discrepancies in how contours are standardized across different institutions and among planners, exacerbated by inter- and intra-observer variability[20,31,32]. This lack of uniformity underscores the necessity for rigorous contour review and verification during the QA and chart check process. Despite the importance of these reviews, they are typically conducted manually, making them both time-intensive and laborious.

1.3.2 Evolution of Radiotherapy

The future of radiotherapy, heralded by the integration of artificial intelligence (AI) and increased automation, have significantly enhanced the potential for online adaptive therapy (oART)[12]. This method adjusts treatment plans in real-time to accommodate changes in patient anatomy. While these innovations promise greater efficiency and tailored patient care, they also complicate treatment management, underscoring the critical need for enhanced quality assurance (QA) tools. Effective QA is essential to leverage these technologies fully while ensuring treatment safety and effectiveness.

1.3.2.1 Artificial Intelligence (AI) and Autocontouring

Artificial Intelligence (AI) and automation significantly enhance the efficiency of radiation therapy treatment planning by automating repetitive tasks and accelerating the contouring process. Deep learning (DL) autocontouring tools, for example, have matched the performance of human experts in some areas, reducing contouring time by up to 77%[15,18,30,33–37].

These tools also facilitate more comprehensive contouring tasks, such as delineating complex glands and substructures[38]. However, they remain a burgeoning area of research[28], with their accuracies varying across different anatomical regions[39-41], sometimes producing inaccurate

or incomplete contours due to overfitting when faced with anatomical variations not present in the training data[39]. All of these challenges lead to frequent inaccurate or incomplete contours, requiring manual post-processing which is often time-consuming and labor-intensive, potentially negating the benefits of autocontouring[39,40].

Despite these challenges, the adoption of autocontouring is increasing[33,39], which leads to a higher output of contours with less human oversight. This rise underscores the need for rigorous oversight and ongoing evaluation of autocontouring tools to ensure their effective integration into clinical practice, especially as these technologies continue to evolve[28,39,40].

1.3.2.2 Online Adaptive Radiation Therapy and Personalization

Online adaptive radiation therapy (oART) offers a dynamic approach to radiation therapy by adjusting to real-time changes in patient anatomy, enhancing dose conformity and mitigating toxicity risks[42–46]. Essential components of oART include daily planning and quality assurance (QA), which involve contouring, treatment planning, and dose calculations performed while the patient is on the treatment couch.

This expedited process is crucial not only for maintaining patient comfort but also for minimizing anatomical changes during treatment[13,44]. Traditional patient-specific QA, often limited under these conditions[45,47], highlights the pressing need for advanced automation and reliable QA solutions[43]. Accurate contour verification becomes particularly critical in this streamlined workflow, ensuring the efficacy and safety of treatments delivered through oART.

1.3.2.3 Problems and Challenges

Advancements in radiotherapy have delivered numerous benefits but also introduced significant challenges. Key among them is the proliferation of OAR (Organs at Risk) contours

produced with diminished human oversight, with less time given to review them. Given the demands of increased workload and efficiency, maintaining a high level of OAR contour fidelity is imperative to prevent unnecessary increases in toxicity.

1.4 Approach

The growing complexity of radiation therapy continues to be driven by advancements such as oART and autocontouring tools[48], highlighting the critical need for more robust contour verification solutions[17,20,43,49]. Traditional manual verification methods are not only susceptible to human error but also struggle to keep pace with the volume of contours generated by autocontouring technologies and the demands of advanced treatment modalities. This underscores the necessity for innovative quality assurance (QA) tools that combine precision with efficiency.

1.4.1 Absence of Tools

Current contour verification practices typically depend on manual review, which can be an inefficient and inconsistent process. It relies on planner expertise, alertness, and takes up to an hour, while remaining prone to errors, problems which are further exacerbated by the challenges of autocontouring and oART[50].

Although various contour verification methods have been proposed, their adoption has been limited due to challenges in generalizability and ease of implementation. These methods fall into two main categories:

Reference-Based Methods: These compare contours against a standard derived from previous cases or an independent segmentation method[18–20,51]. However, geometry-based metrics

like the dices coefficient and Hausdorff distance often fail to save time effectively and may not correlate well with clinical outcomes[16,18,33,52].

Knowledge-Based Methods: These identify discrepancies based on quantitative features like size, shape, and image intensities. While useful, they tend to overlook minor but clinically significant errors along the boundaries or extremities of contours, where dosimetric constraints are most critical, such as for points of maximum dose[19,39].

Both overfitting issues and insensitivity to minor deviations can greatly affect dose optimization. Under-contouring of the OAR leads to inferior OAR sparing with potential for increased or unanticipated toxicity, while over-contouring could result in unnecessary compromises to the target volume[30,53].

The limitations of these methods highlight the urgent demand for a tailored QA approach that ensures timely and accurate verification of clinically relevant contour sections, adapting to the diverse anatomical variations seen in patients.

1.4.2 Our Solution

We propose a methodology that relies on dose constraints in order to streamline the verification of OAR contours. This approach employs margins added to specific structures, allowing us to focus on critical areas determined by these constraints, thereby enhancing the accuracy and efficiency of contour verification. Unlike traditional methods, our strategy does not depend on extensive training datasets but rather on planner-specified dose constraints and margins. This makes it highly adaptable and easy to implement across different clinical settings.

2. Methods

2.1 Overview

In radiotherapy treatment planning, margins are employed to mitigate uncertainties and ensure precise dose delivery. These margins are crucial for maintaining high target coverage and for safeguarding organs at risk (OARs) by expanding volumes of interest.

Our methodology leverages margins to proactively identify and rectify potential inaccuracies in contours, while accounting for uncertainties from treatment planning and patient setup. By expanding OAR contours by specific margins and comparing these against established dose constraints, our approach can anticipate and flag areas that might violate these constraints, focusing the verification process on subsections of critical clinical importance. This strategy not only streamlines the manual review process by directing attention to the most pertinent areas but also aids in preventing significant errors, thereby optimizing both safety and efficiency.

The upcoming sections will detail the development and application of a tool that implements this method, demonstrating its practicality and effectiveness for potential clinical use.

2.2 Tool Development

2.2.1 Design Philosophy

The core components of our method and tool are the application of margins and dose constraints, features inherent to many treatment planning systems. This enables us to develop our tool natively within the Eclipse Treatment Planning System, allowing for seamless integration.

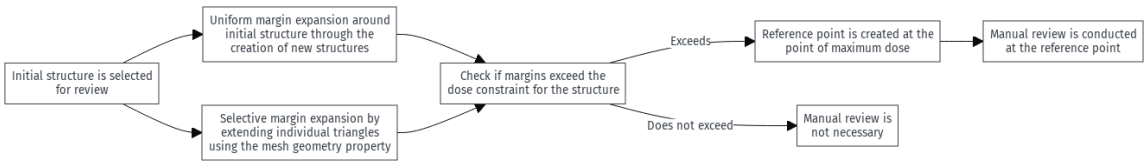

Figure 1: Workflow of different steps of the tool, with specific approaches towards each step geared towards its use in Eclipse for SRS cases

Figure 1 illustrates the workflow within the Eclipse environment, demonstrating how our tool applies different margin sizes to selected structures. The tool evaluates whether these margins exceed predefined dose constraints. When margins surpass these constraints, a reference point is marked at the maximum dose location to indicate areas needing manual review. This ensures that clinical focus is maintained only on areas where dose limits are breached, enhancing the efficiency of the review process by prioritizing critical treatment planning areas where manual intervention is necessary. If no constraints are exceeded, the tool allows planners to allocate their time and resources to other vital aspects of treatment planning, thus optimizing overall workflow efficiency.

2.2.2 Software Specifications

Our tool was specifically developed for the Varian Eclipse Treatment Planning System (v16) utilizing the Eclipse Scripting API (ESAPI v11) using C# in Visual Studio (2022). For manual contour reviews, we primarily used CT and MR images from Eclipse; however, in instances where MR images were unavailable, iPlanRT (v4.1.2) served as an alternative imaging resource.

2.3 Tool Application

2.3.1 Patient Case Selection

92 patient cases treated with SRS at Duke University Medical Center between 2013-2016 were retrospectively analyzed. Contouring for these patients was done manually. This diverse cohort, covering both single and multi-fraction SRS treatments, provided a comprehensive dataset for validating the tool's effectiveness, and utilized both Volumetric Modulated Arc Therapy (VMAT) and Intensity-Modulated Radiation Therapy (IMRT). However, the choice of treatment technique did not influence the application of dose constraints in our analysis and thus is not a focus of our evaluation.

2.3.2 Study-Specific Workflow

Our method, while adaptable, was specifically tailored and optimized for a retrospective dataset of SRS patients using our tool, currently dubbed "ContourCheck." Once run, this tool applies predefined margins (0, 1, 3, and 5mm) to brain structures such as the brainstem, optic nerves, and optic chiasm, as depicted in Figures 3 and 4. Should these margins exceed the dose constraints established by QUANTEC and HyTEC, the affected structure is flagged, assigned a risk level based on the exceeded margin, and a reference point is marked at the location of maximum dose for targeted manual review (Figure 5). This process ensures focused verification only where necessary, enhancing the efficiency of our workflow. If no constraints are exceeded, the associated minor discrepancies are deemed clinically irrelevant, negating the need for further manual assessment.

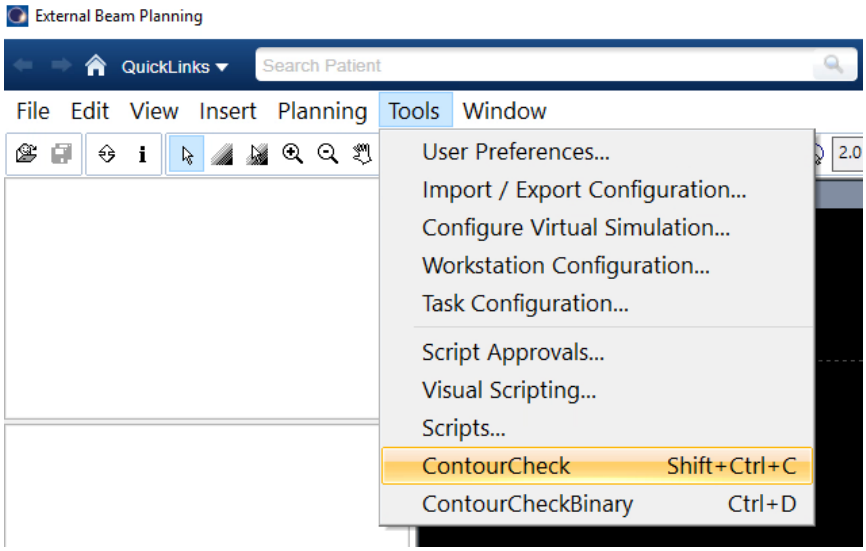

Figure 2: Use of the Tool in Eclipse

Figure 3: Treatment Plan View of Selected Structures

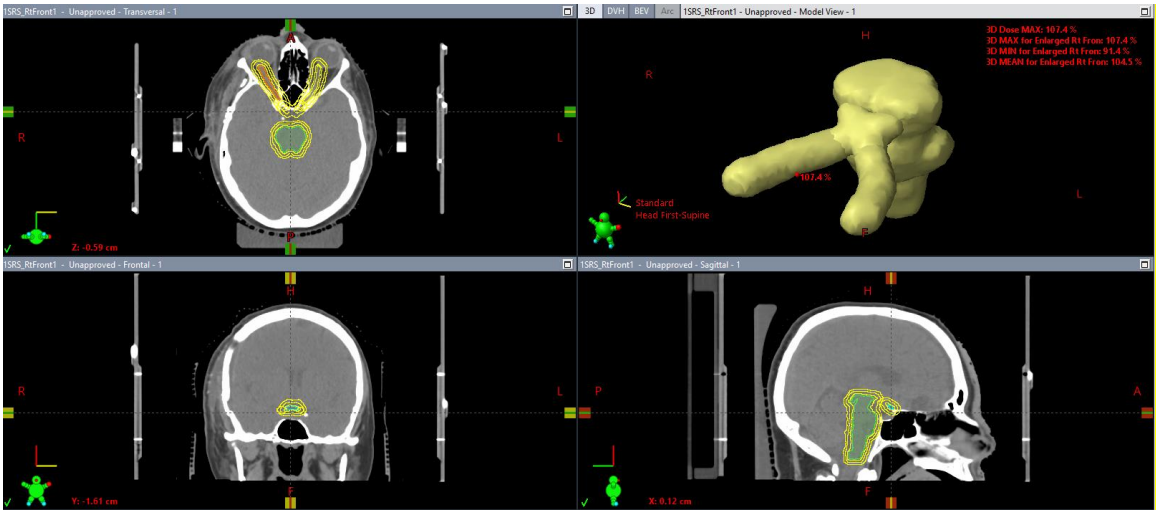

Figure 4: Margins are added to selected structures

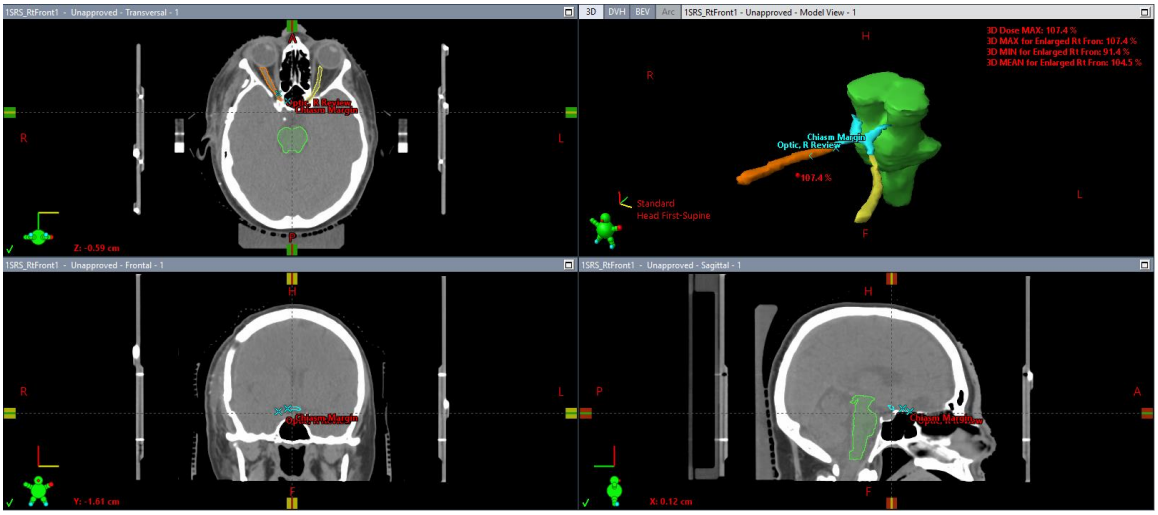

Figure 5: Reference points are created for easy localization

2.4 Structures

2.4.1 Structure Selection

The decision to use the SRS patient dataset was due to both the minimal motion from the brain as well as the high dose gradients of the technique, which helped informed margin

selection. The brainstem, optic nerves, and optic chiasm were selected as they had relevant QUANTEC dose constraints, highlighting their critical anatomical significance.

2.4.2 Dose Limits

Dose limits, crucial for ensuring patient safety while maximizing treatment efficacy, were standardized using established guidelines provided by QUANTEC and HyTEC. HyTEC provides greater nuance for hypofractionated treatments and offers recommendations for multiple fraction SRS schemes instead of just single fraction SRS[3]. It recommends specific dose limits for the optic nerves and chiasm to prevent radiation-induced optic neuropathy, a critical consideration in our tool's design. Similarly, the dose constraint for the brainstem involved careful consideration of literature values and the application of radiobiological principles to determine the most appropriate limits for single and multifractionated SRS treatments.

While clinical dose limits would better reflect the specific preferences of the institution, the variation across different patients and lack of standardization in their recording pose challenges to its retrospective use in this study.

For limits involving maximum dose to a point, checking if the structure dose constraint simply consists of finding the maximum dose within the structure, and checking if it was above the dose limit. Meanwhile, limits which involve mean dose require repeated checks of the whole structure to see if the dose constrain has been exceeded whenever the contour or treatment plan has been changed.

Volumetric constraints, whereby a certain volume or percentage of the structure has to be remain below a specified dose level, are similar to simple dosimetric constraints, in which the maximum dose can be used to determine clinical relevance. However, determining the area of

greatest clinical relevance is slightly more complex as there are multiple approaches to reducing the overall volume that contributes to the constraint being exceeded.

2.4.2.1 Optic Nerves and Chiasm

The optic nerves and optic chiasm are critical organs for vision, sharing the same dose limits of maximum dose to a point of 12 Gy in 1 fraction, and 25 Gy in 5 fractions for a less than 1% incidence of radiation-induced optic neuropathy[54].

2.4.2.2 Brainstem

The brainstem is one of the most critical organs in the body; it is not recommended to risk exceeding its constraints to meet tumor control targets[55]. For single fraction Stereotactic Radiosurgery (SRS), a maximum dose of 12.5 Gy to the brainstem correlates with a low risk (<5%) of adverse effects, as outlined by QUANTEC [56]. Due to a scarcity of data on multifractionated SRS, we initially established a provisional dose constraint of 26 Gy, derived by applying a scaling factor (25/12) analogous to the constraints for optic structures. It was also relatively close to the dose constraint of 25 Gy used in several of the retrospective 5-fraction cases for which we could find clinical dose constraints.

Subsequent refinement employed a linear-quadratic (LQ) model to calculate the Biologically Effective Dose (BED), incorporating a conservative α/β ratio of 2[54]. Although the direct application of this LQ model remains to be fully validated[56], it offers a more physiologically pertinent framework for 5-fraction SRS by considering the fractionation effect on tissue response. The detailed BED calculation, based on a single fraction dose of 12.5 Gy and an α/β ratio of 2, is as follows:

$$BED = Total\ Dose \times (1 + \frac{Fraction\ Dose}{\alpha\beta})$$

$$90.625\ Gy = 12.5 \times (1 + \frac{12.5}{2})$$

$$90.625\ Gy = x \times (1 + \frac{x/5}{2})$$

$$x = 25.52\ Gy$$

Given the close proximity between the refined value (25.5 Gy), the clinical constraint commonly used (25 Gy), and our initial estimate (26 Gy), we retained the original scaled dose constraint. However, conducting a sensitivity analysis with these alternative values in future studies could potentially allow for additional insight into the applicability of the LQ model or

conducting a sensitivity analysis with the refined 25.5 Gy constraint in future studies could potentially enhance the robustness and validity of our evaluation, offering a more nuanced understanding of the dose-response relationship in multifractionated SRS for the brainstem.

2.5 Margins and Risk Assessment

2.5.1 Margins

The systematic addition of margins serves to stratify risk, with structures that exceed dose constraints at smaller margins identified as higher risk, reflecting a greater likelihood of clinically significant contouring errors. Conversely, exceedances observed with larger margins are deemed to be at lower risk or reduced priority, as they suggest a lower probability of clinically significant errors. This stratification facilitates focused attention on areas where precision in contouring has the most substantial impact on treatment outcomes.

While the selection of which specific margins to use is somewhat arbitrary in nature, the patient's anatomy, structure selected, treatment techniques used, as well as the planner's

preferences are all potential factors in determining what margins might be used. In particular, the tolerance limits of the treatment technique can be used to determine tighter margins that correspond to areas of highest potential risk, acting as a safeguard against critical errors and uncertainties.

2.5.1.1 Margin Selection

In our approach, we systematically apply margins of 0, 1, 3, and 5mm to contours in the brain (designated as “very high”, “high”, “medium”, and “low” risk categories, respectively). If a dose deviation occurs with no margin, then the accuracy of the contour is critical as inaccuracies in the contour, especially near the tumor, would directly affect the violation of clinically relevant dose constraints.

A 1mm margin is chosen to align with the standard tolerance levels in Stereotactic Radiosurgery (SRS), reflecting the high degree of precision required in the treatment planning process. Margins of 3 and 5mm, which exceed this baseline, indicate a tiered priority system for review. They enable us to effectively check for clinically significant effects when contouring errors of larger magnitude occur, even when the probability of those contouring errors is relatively low.

In SRS treatments, precision is paramount due to the high doses delivered in a small number of fractions. However, it's essential to balance precision with robustness to uncertainties, such as setup errors and organ motion. While using smaller margins for our tool better correlates with the tolerances expected of the technique, larger margins may provide increased robustness to uncertainties, particularly in cases with greater anatomical variability or complex tumor geometries.

2.5.1.2 Margin Expansion

Our study explores two main approaches to the application of margins: uniform expansion of the whole structure and selective expansion using Eclipse's mesh geometry. These techniques add additional volume to the structure of interest in different ways, allowing these new volumes to be checked against the dose constraints, enabling the identification of critical areas that might not be immediately obvious. The choice of method depends on the specific clinical scenario, with each offering unique insights into the treatment planning process.

Uniform Expansion

This straightforward approach involves the expansion of contours in a uniform manner by predefined margins (0, 1, 3, and 5mm). This is particularly effective in situations where maximum point dose limits are the primary concern. Uniform expansion efficiently identifies whether the maximum dose is surpassed in the entire volume, facilitating quick risk assignment.

Selective Expansion

In contrast, selective expansion leverages the mesh geometry of contours within Eclipse for a more detailed analysis. This 3D representation of the structure through numerous triangles that make up the surface of the contour enables the identification of specific areas where dose constraints are exceeded by extending individual triangles outward into prisms. It incorporates the spatial ordering of triangle indices (x: left to right, y: bottom to top, z: feet to head) to map out the precise locations of potential exceedances, allowing for a focused examination of separate subsections beyond the set dose constraints.

While no volumetric constraints were used in this study, selective expansion was theorized to enable accurate assessment of the effect of margin expansion on volumetric

constraints. By assessing incremental volumes resulting from individual margin extensions (prisms) of the triangle in the mesh geometry, this method has the potential for evaluating how changes in contour dimensions and volume impact dose distribution.

Applications in Our Study

In our study, both methods were investigated to explore their advantages and capabilities. In the case of uniform expansion, uniform margins of 0, 1, 3, 5mm were added to the structure through the creation of new structures with the specified margin added. Meanwhile, selective expansion was done through the outward extension of individual triangles, while the average value of the dose profile along the center of the prism created was obtained for dose constraint comparison.

2.5.2 Dose Constraint Comparison and Risk Level Assignment

The maximum dose of the selected structures and their margins were compared against their maximum point dose constraints, and categorized into risk levels to guide clinical decision-making.

Risk classification is determined based on how the margins of a structure relate to dose constraints. If the dose received by the original structure (with a 0mm margin) surpasses the dose constraint, it's flagged as very high risk. This classification system extends to encompass the structures with added margins of 1, 3, and 5mm, with each increasing margin corresponding to descending levels of risk (high, medium, and low, respectively). The same principle applies to selective expansion; if an individual prism of a structure’s mesh geometry with an average dose value exceeds constraints, the structure is flagged.

2.5.3 Highlighting Critical Areas

Both techniques of margin expansion facilitated our ability to pinpoint clinically significant areas while providing different benefits through their different approaches. However, due to time constraints, not all of them were explored in detail.

In our study, identification and highlighting of critical areas were primarily accomplished using the selective expansion method. By focusing on individual prisms (extensions of triangles) for a given structure's mesh geometry and identifying those that exceed the dose constraint, we can potentially identify all potential points on the structure that are of interest. By comparing the index values of the triangles, we can also determine their proximity to one another within the mesh, identifying outliers.

The central point of the triangle on the mesh geometry of the structure receiving the highest dose while exceeding dose constraints is designated as a reference point for subsequent detailed assessment. For simplicity and to facilitate ease of review within this study, the reference point was created at the center of the highest dose triangle in the initial structure and utilized as the sole point for contour evaluation in each case.

The following details additional approaches to highlighting critical areas that were theorized based on our proposed margin expansion techniques, as well as the functionalities available through Eclipse and ESAPI, but were not used in the final iteration due to time constraints.

Extension Inwards

A technique theorized later in our study involves tracing the maximum dose within the margin-expanded structure back towards its origin using selective margin expansion (or contraction in this case), more effectively pinpointing the source of the highest dose within the

original structure for error identification. This inward extension from the margin structure may offer a more refined method for identifying critical areas that warrant close examination while counteracting the potential problem of having "blind spots" for areas between individual triangles facing different directions.

Overlapping Structures

Another strategy that was contemplated for highlighting critical areas involves analyzing the overlap between isodose lines—representing the dosimetric constraint—and the expanded margins of structures. This method aims to comprehensively map the entire region exceeding dose constraints, potentially offering a more complete view of areas requiring attention. This would also enable the use of the volume of the new structure to be used as a method to classify risk, with the potential to be used directly in inverse optimization strategies.

Although these techniques were not used, they hold promise for future studies, allowing for different approaches to the study of margins and dose constraints.

2.6 Validation and Testing

There are two parts of our tool that require evaluation. The first is the ability to reliably flag contours which exceed specified dose constraints after given margins. Evaluation of these contours is then performed by qualified personnel who have undergone training in contouring and treatment planning.

2.6.1 Statistical Analysis

To evaluate the efficacy of our contour verification tool, we analyzed the maximum and mean dose values for each structure, employing statistical correlation analysis. The Mann-

Whitney U test and Kruskal-Wallis test were used for assessing two independent samples and multiple groups respectively due to the non-parametric nature of our data in an effort to elucidate the relationship between dose levels and the structures' assigned risk levels.

Data analysis was conducted using Microsoft Excel, with significance levels were established at p-value thresholds commonly accepted in medical research (e.g., $p < 0.05$), allowing us to identify statistically meaningful correlations.

2.6.2 Assessment by Trained Personnel

To assess the accuracy of contours at the reference point identified by our tool, two physicists with training in OAR contouring independently verified every flagged contour's accuracy, determining whether it was over or under-contoured and measuring the offset in the plane with the largest discrepancy. They were briefed on the method and tool use, including how to locate the maximum dose coordinates, but were not informed of the risk levels to ensure unbiased assessment. To minimize subjectivity, the mean offset value was calculated between the two sets of measurements, alongside their correlation, mean, and standard deviation.

Both CT and MR images were used in the identification of errors at the coordinates of max dose in the structure given by the tool, with MR being the primary modality due to the ability to better distinguish precise boundaries of targeted structures due to its superior contrast. Slice thickness and resolution of MR image slices were 1mm. For a comprehensive evaluation, axial, coronal, and sagittal views were analyzed, allowing us to identify and measure the largest offset errors for each of these planes. The largest measured contour error over was used for analysis, which might have led to an underestimation of the true contouring error. However, this also reduces the effects of any human error in using the measurement tool.

This approach helps to provide us with a comprehensive evaluation of the tool's accuracy, its effectiveness in identifying clinically significant contouring errors, and its precision in pinpointing areas requiring attention. While primarily designed to enable faster and more accurate contour verification, our tool does not provide specific suggestions on corrective actions to take after errors are found, but merely provides the necessary information to do so.

3. Results

3.1 Performance Evaluation

The ability of the tool to perform as we designed can be shown through the following figures, whereby we see that the isodose line (pink) that matches the dose constraint used intersects with the margins (yellow) without intersecting with the original structure (green), resulting in a low risk level in Figure 6. This is in contrast to Figure 7 whereby the dose constraint directly intersects with the structure, categorized as very high risk, as well as Figure 8, whereby the dose constraint does not intersect with any margins.

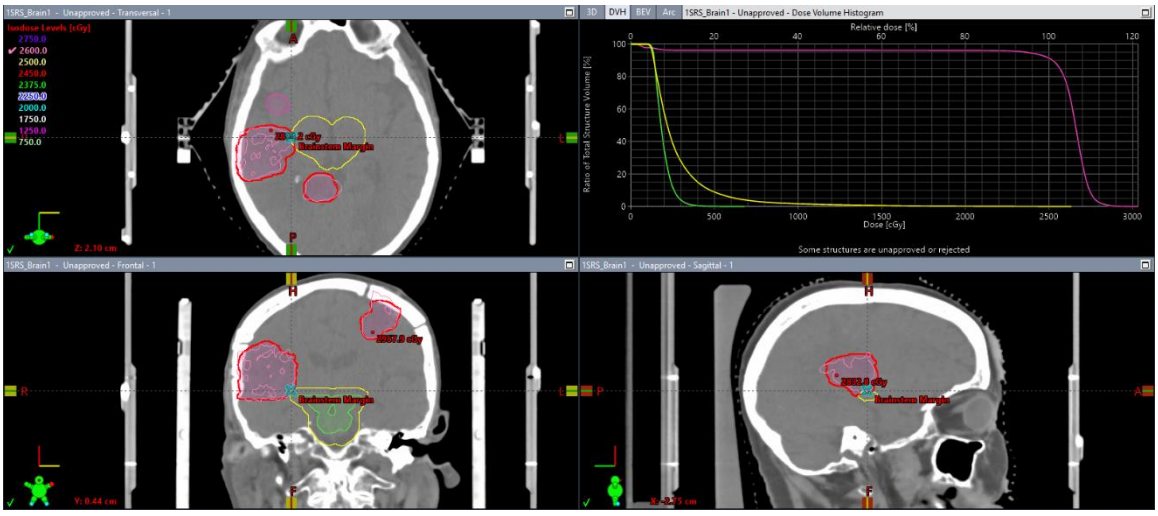

Figure 6: Low Risk, Isodose Line intersects with Margin only

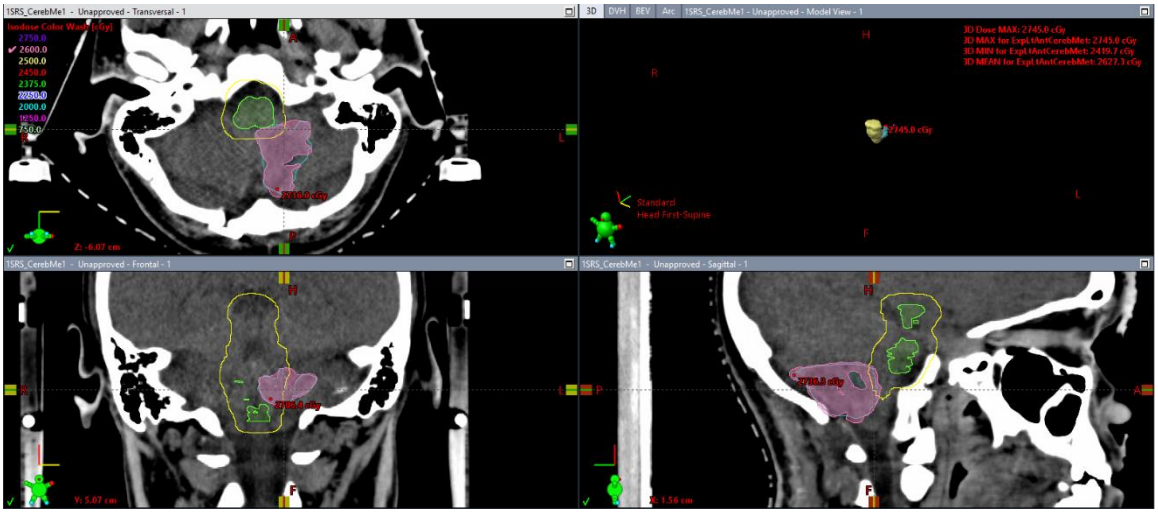

Figure 7: Very High Risk, Isodose Lines intersects with Margins and Structure

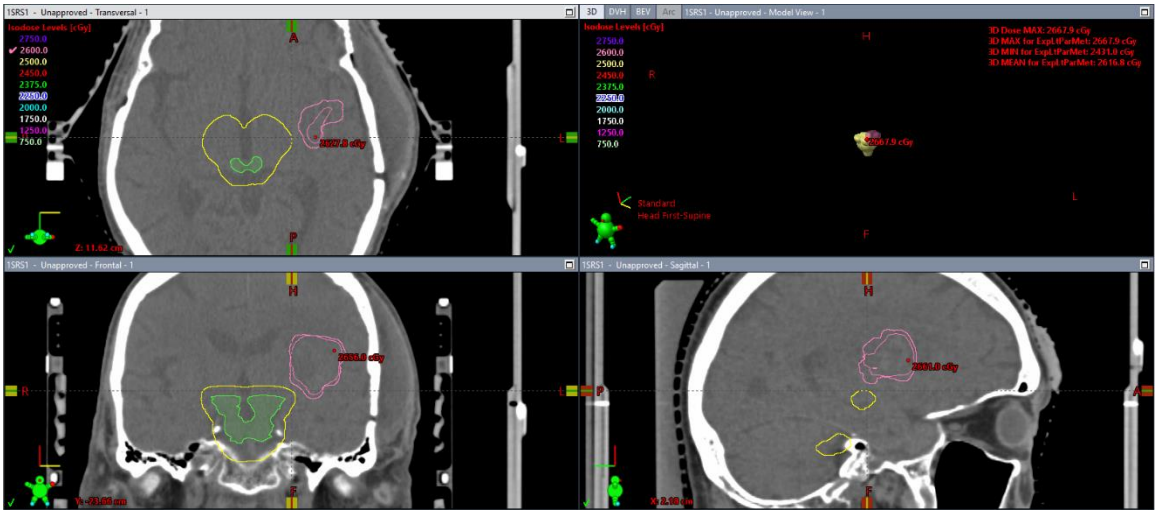

Figure 8: Not Flagged, Isodose Line dose not intersect with Margin

In our study of 92 stereotactic radiosurgery (SRS) cases, our tool identified 12 contours requiring manual review due to potential dose constraint exceedances, with the majority involving the brainstem (n=10), followed by the optic nerve (n=1) and chiasm (n=1). Statistical analysis with the Mann-Whitney U test revealed a significant correlation between both the

maximum (p = 2.78E-06) and mean doses (p = 7.18E-03) for the brainstem in 5-fraction SRS cases and the propensity for exceeding predefined dose constraints.

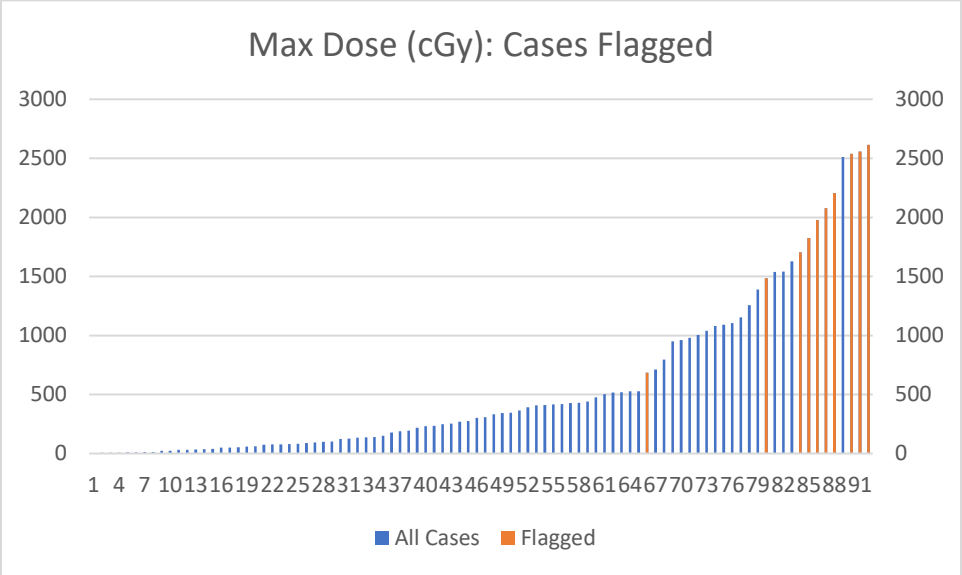

Figure 9: Max Dose (cGy) of the brainstem for all cases, highlighting those flagged

Figure 9 illustrates the distribution of maximum doses across all examined SRS cases, with highlighted cases indicating those that exceeded dose constraints, giving us a visual representation of outliers and trends observed. While there is an association between high max dose and being flagged by the tool, this relationship is not strictly so, and exceptions do occur. Thus, the maximum dose of a structure alone is not a sufficient indicator of whether the contour should be reviewed.

However, the relationship between these dosimetric parameters and the breakdown of the contours flagged by their risk levels for review using the Kruskal-Willis Test indicated weaker correlations (p = 7.86E-02 for maximum dose and p = 8.19E-01 for mean dose), suggesting that while dosimetric parameters effectively flag potential issues, they may not significantly inform

the assigned risk levels. This points to our risk levels providing insight that cannot simply be demonstrated by a direct analysis of dosimetric measures.

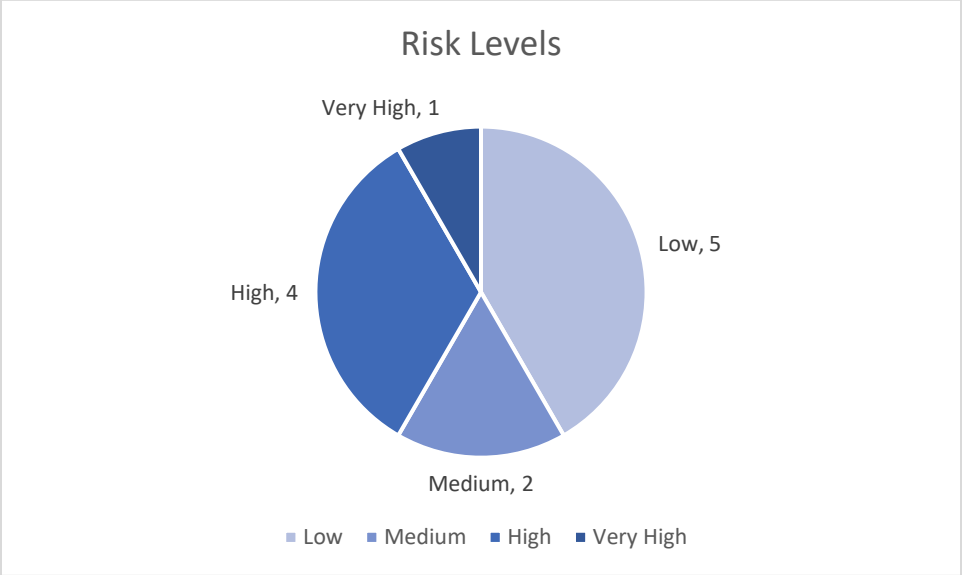

Figure 10: Risk Level Distribution

In Figure 10, we observe the distribution of risk levels for brainstem contours as follows: Very High Risk (n=1), High Risk (n=4), Medium Risk (n=2), and Low Risk (n=5), with the risks corresponding to exceedances at margins of 0mm, 1mm, 3mm, and 5mm, respectively, demonstrating the method's capability in risk stratification.

3.2 Assessment by Trained Personnel

Independent assessment of contouring errors by the trained physicists were recorded, with a mean difference of 0.19mm (SD = 0.57mm) with a correlation value of 0.7 between them. The average mean value offset over all flagged cases was 0.45mm, but increases to 0.76mm when only including those which had non-zero error measurements.

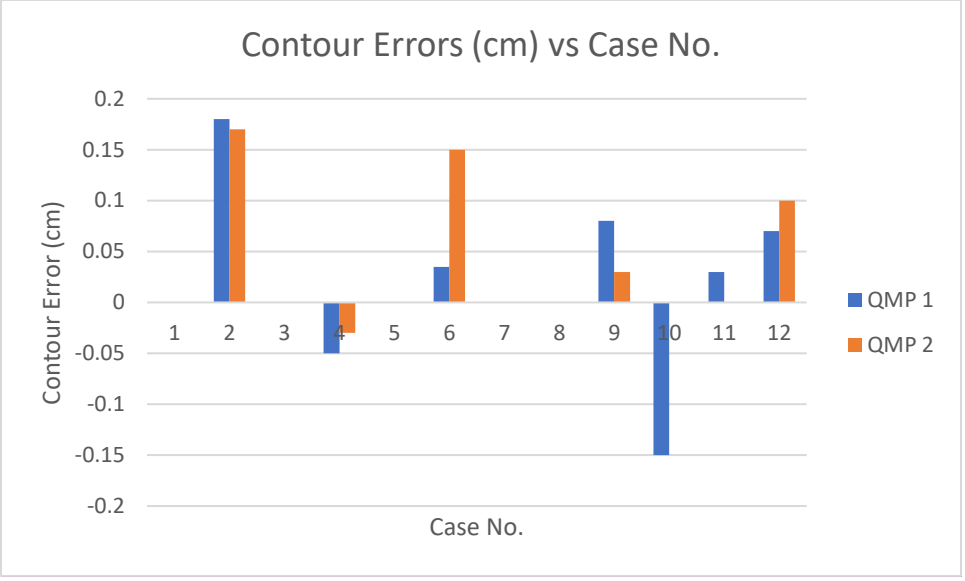

Figure 11: Errors from Manual Assessment

Figure 11 shows the contour errors as assessed by our two trained physicists. Positive values indicate that the structure was under-contoured, and may have received dose that would have been unaccounted for if no corrective actions are taken, while negative values indicate over-contoured structures, which might have contributed to non-optimal tumour coverage. The mean offset between the two assessment for each case provided the basis for our analysis. Notably, one case exhibited a mean contouring error of 1.75mm (case 2 in Figure 11), significantly beyond the standard tolerance for SRS of 1mm. This was an under-contoured structure, and thus may have been exposed to high levels of dose exceeding its constraints, suggesting a critical area of concern.

3.3 Case Studies

The following case studies present some of the study's most significant findings, showcasing results of flagged cases and errors that deviate from the rest in significant and interesting ways.

Large Offset Case

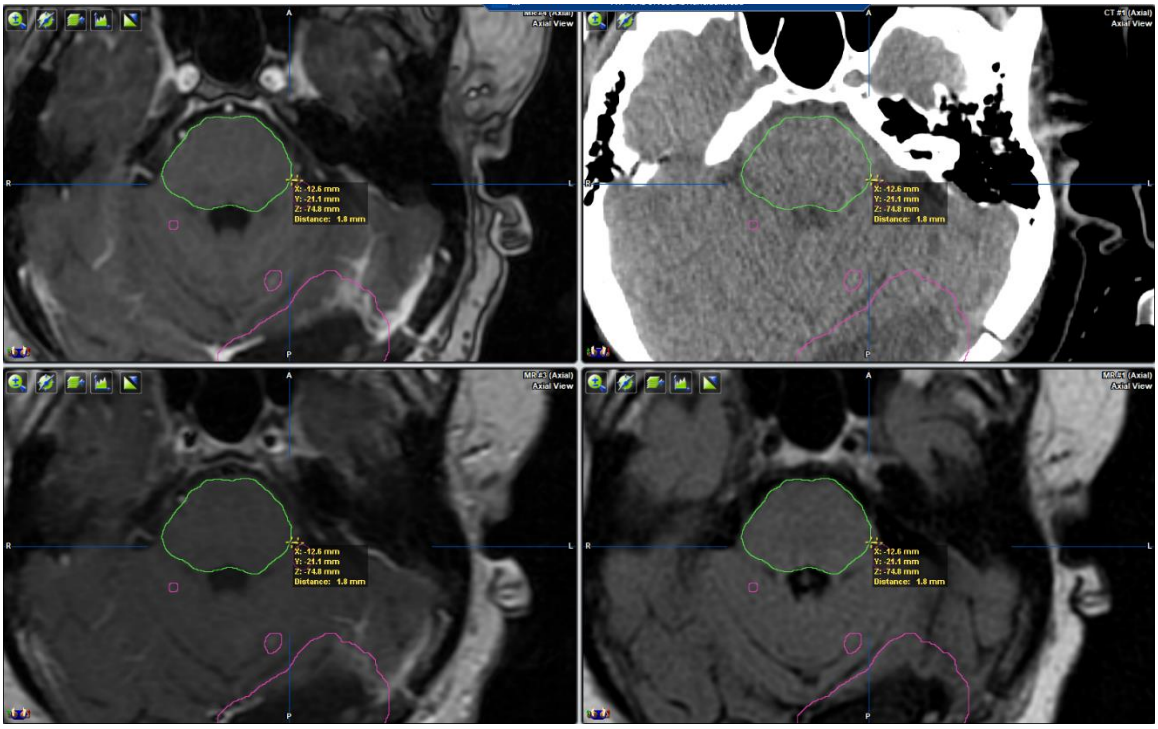

Figure 12: Large Offset Case, Axial View

Figure 12 illustrates a significant contouring error in the brainstem detected by our tool using both CT and specific MRI sequences in iPlanRT. The right side of the brainstem contour (green) was under-contoured by 1.75mm, as determined by the physicists in the axial plane. This finding underscores the precision of our tool in pinpointing clinically relevant errors, which are

crucial for further clinical evaluation and physician review to ensure optimal treatment planning and patient safety.

Low Maximum Dose and Sharp Edges

In a divergent case, the brainstem received a maximum dose of 684 cGy, significantly below the general trend where flagged cases exceed 50% of the dose constraint. However, after applying margins, the dose unexpectedly escalated by 1955 cGy.

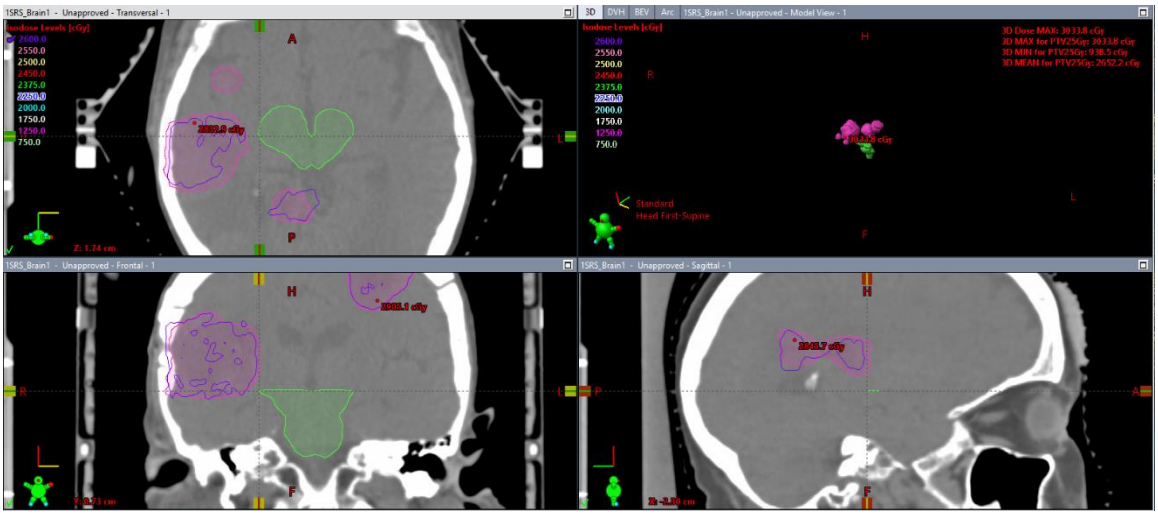

Figure 13: Low Maximum Dose and Sharp Edge

As illustrated in Figure 13, the point of maximum dose is positioned along a sharp contour edge in the coronal view (bottom-left), adjacent to the nearest tumor. This sharp edge, possibly resulting from interpolation between planes, eluded detection through selective expansion due to its location in a geometric "blind spot." It was only through uniform expansion that this exceedance was identified, indicating that the sharp edge significantly influenced the notable increase in maximum dose.

Decrease in Mean Dose

In our analysis, margin expansions generally resulted in an increase in mean dose values. However, a notable exception was observed in one case where the mean dose decreased, illustrating the nuanced impact of margin adjustments on dosimetric outcomes. This anomaly suggests that in certain scenarios, applying selective margins that allow for minor, direction-specific adjustments could more accurately reflect changes in dose distribution.

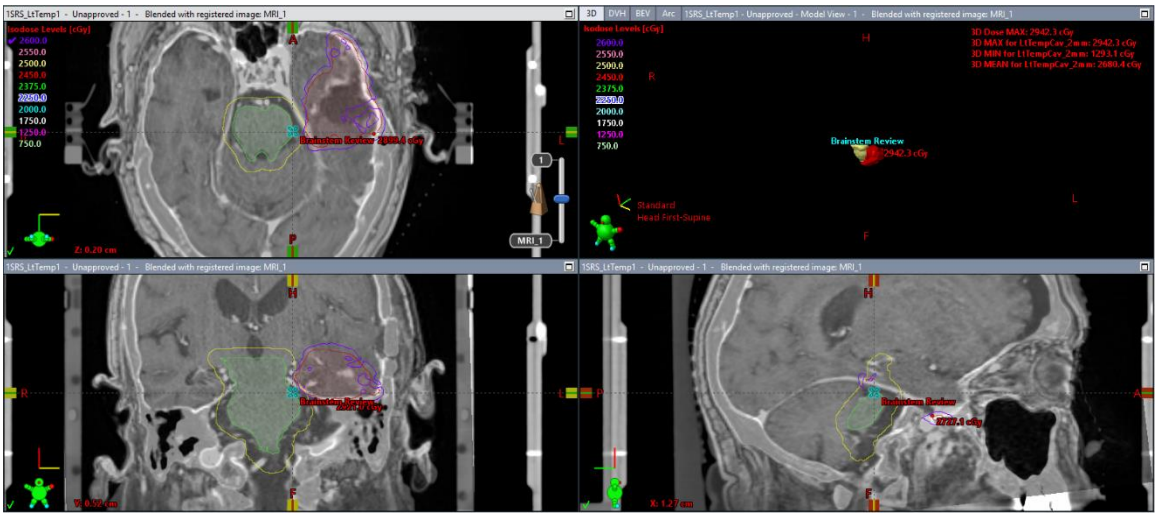

Figure 14: Decrease in Mean Dose, Uniform Margins

Figure 14 visually represents this phenomenon, showing the brainstem (green) and applied margins (yellow) in relation to the Planning Target Volume (PTV) (red) and the 26 Gy isodose line representing the brainstem dose constraint (purple), across three planes. This graphical depiction underscores the potential of selective margins to provide insights into the complex interactions between margin size and dose variability.

Optic Structures: Right Optic Nerve and Optic Chiasm

This case study examines a patient's treatment plan where both the right optic nerve and optic chiasm were flagged due to proximity to the tumor with a dose constraint of 25 Gy in a 5-fraction SRS treatment.

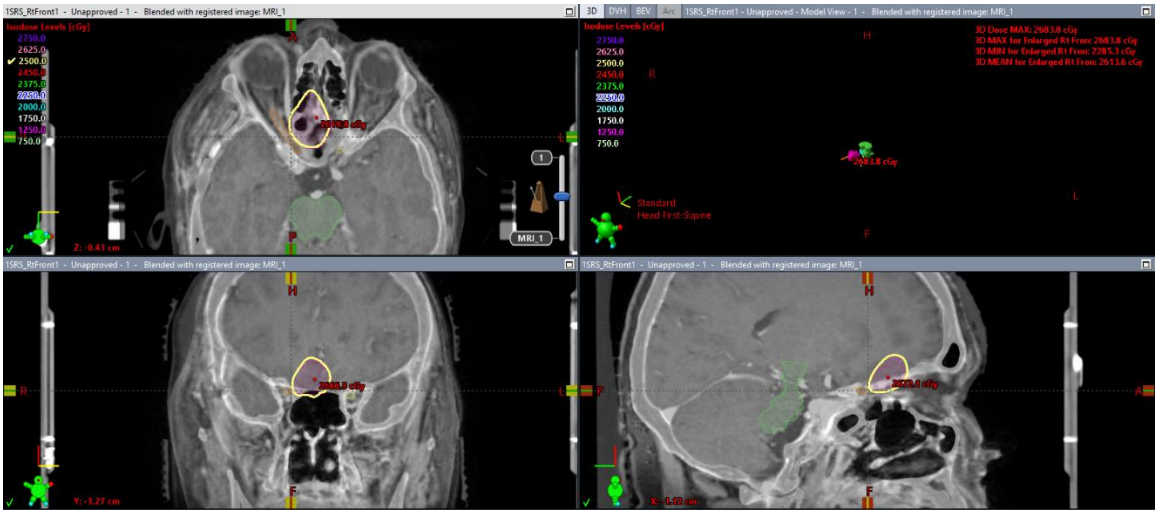

Figure 15: Right Optic Nerve, 25 Gy Isodose

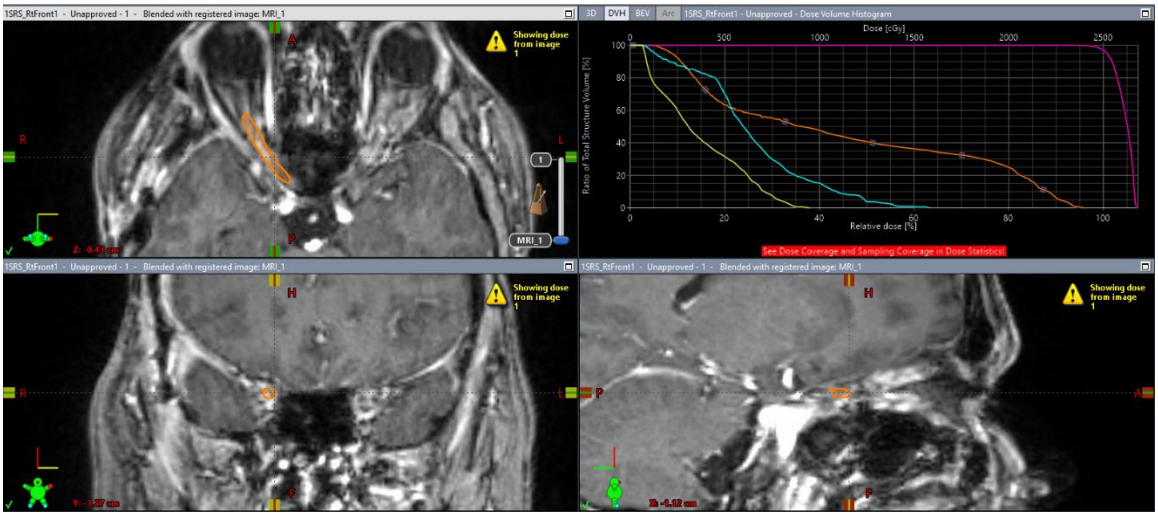

Figure 16: Right Optic Nerve, MRI

Figures 15 and 16 illustrate the right optic nerve with CT and MR images, highlighting the 25 Gy isodose line in yellow to delineate the dose limit.

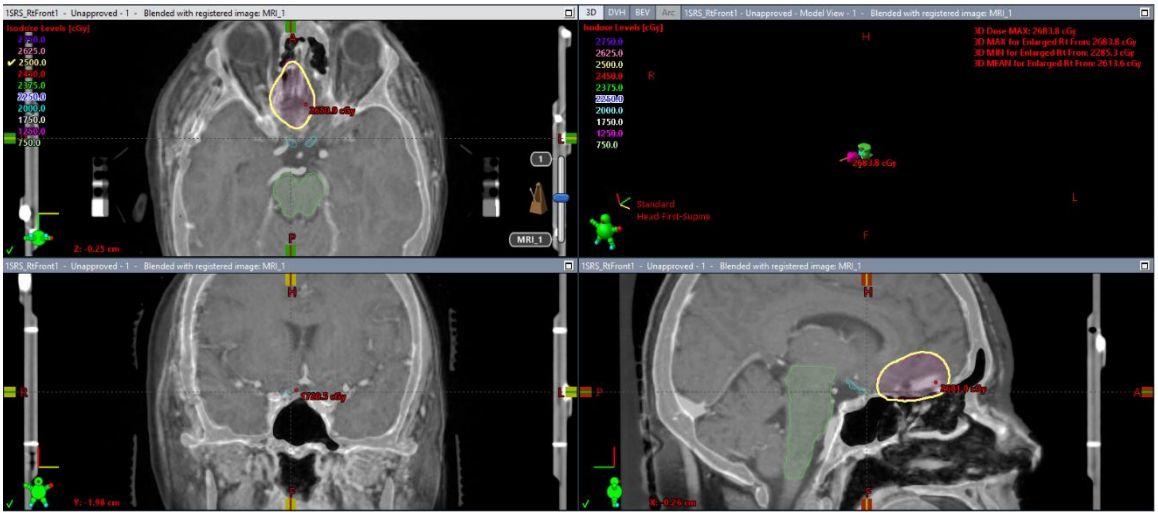

Figure 17: Optic Chiasm, 25 Gy Isodose

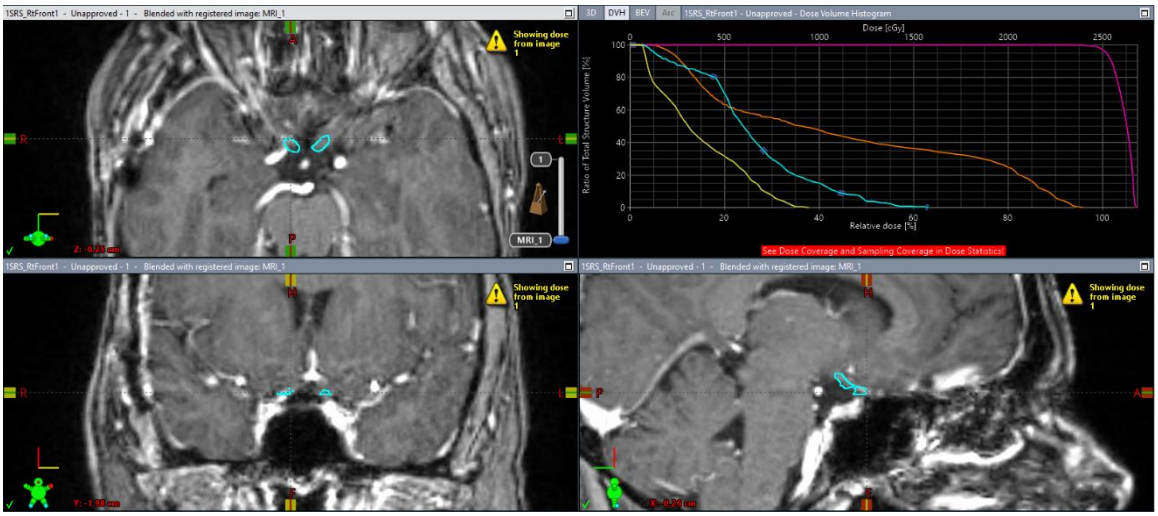

Figure 18: Optic Chiasm, MRI

Similarly, Figures 17 and 18 depict the optic chiasm with corresponding isodose lines and MR images. The visual representation from these figures shows the tumor's unusual proximity to

these critical structures, which likely influenced their flagging for potential risk in this treatment plan.

4. Discussion

4.1 Clinical Implications

Our findings present a novel method for streamlining contour verification in radiation therapy, demonstrating its utility through the development and application of a tool, accurately identifying and prioritizing critical areas of contours for manual review.

SRS-specific Observations

The dataset, including both single and five-fraction SRS cases, showed that flagged cases were exclusively from the five-fraction cohort, reflecting specific institutional preferences for fractionation strategies in treating large target volumes near organs at risk (OARs).

Notably, the majority of cases flagged for exceeding dose constraints involved the brainstem (n = 10), highlighting its sensitivity to dosimetric variations in SRS treatments. This might be a reflection of its size and shape, or proximity to common tumour locations. Despite employing a non-standardized dose limit for the brainstem, the potential severity of incorrect treatment delivery to this area cannot be understated.

The lack of flags for cases containing the optic nerves or optic chiasm is likely due in part to the conservative dose constraint of 20 Gy used in multiple plans, as compared to the higher 25 Gy recommended for these structures in literature. This is in contrast to the 25 Gy dose limit used for the brainstem in certain plans we were able to obtain dose limits for, which was relatively close to the 26 Gy value we used.

Contouring Errors

A mean offset value of 1.75mm was observed in one case, which exceeded the standard uncertainty margin of 1mm for SRS, suggesting significant potential for harm. While other mean

offset values remain within the tolerance limit of 1mm, contouring errors to OARs must be carefully evaluated with respect to their risk level and extent of error to determine the necessary corrective actions needed to minimize risk to the structure. Out of 5 cases which had a mean offset value greater than 0.5mm, 4 of them were assessed to be undercontoured, while 1 was overcontoured.

Applying margins to an undercontoured structure would get us results closer to what the structure would get in practice if it were contoured correctly. Such structures stand to be at greater risk, as there might be dose given to the structure that is not accounted for, which would likely need to be rectified and evaluated before treatment. Overcontoured structures on the other hand are ‘safer’, as the margins are being added on top of some amount of buffer already. However, there is a risk that the plan might have been inadvertently optimized to give a lower PTV coverage than otherwise possible to meet the increased volume of the brainstem constraint.

Another key observation was that contouring errors predominantly occurred in the axial plane, indicating this dimension's importance in contour verification. However, this might also be a reflection of the shape of the brainstem. In our study, offset values are measured in-plane of the greatest error, with just a single plane instead of considering the combined error in all 3 planes, which might lead to an underestimation of the true contouring error.

4.2 Comparisons with Existing Tools

Tools proposed for contour verification commonly employ reference-based and knowledge-based methods, which require extensive datasets to train their models or identify outliers effectively. Such methods are often subject to overfitting, potentially limiting their applicability to patients with atypical anatomies. Unlike these tools, our method does not depend on large historical datasets or extensive training. Instead, it utilizes dose constraints to verify

contours, ensuring that the focus remains on the most clinically relevant regions without the risk of overfitting. This approach not only enhances the tool's versatility across varied patient anatomies but also simplifies clinical integration and usage.

Furthermore, our method uniquely prioritizes contours based on their clinical significance rather than consistency alone, streamlining the verification process by concentrating on areas that are most likely to affect treatment outcomes. This is in contrast to other QA tools that may detect gross errors irrespective of their clinical impact. While those tools are effective in identifying significant deviations, our tool complements them by pinpointing finer, clinically significant errors that directly relate to exceeding dose constraints. By combining the strengths of both approaches, we could enhance overall treatment accuracy, focusing on critical errors without assuming the absence of gross errors.

4.3 Limitations

Our study, based on data from a single institution between 2013 and 2016, offers specific insights but has limited generalizability. Institutional variations, such as differences in prescribed doses, contouring techniques, and dose constraints, might affect the performance of our tool. Furthermore, the use of selective expansion and mesh geometry to verify contours presents technical challenges. Irregular shapes and sharp edges can impede accurate extension of triangles, potentially creating "blind spots" where critical areas may go undetected. This issue is particularly problematic for structures with complex geometries where sharp points might not align with the main faces, thereby eluding detection.

Additionally, the mesh geometry used occasionally produced degenerate triangles with collinear coordinates, unsuitable for selective expansion. These triangles were excluded from our analysis to preserve its integrity. The inherent complexities of mesh geometry, combined with

irregular structural nuances, necessitate ongoing refinement to ensure the tool's accuracy and reliability in identifying critical areas across diverse anatomical variations. Another significant challenge was the inconsistency in obtaining tumor volume data and the distances between tumors and structures due to varied institutional naming conventions in stereotactic radiosurgery (SRS)[57]. Future iterations of our tool will aim to address these limitations, enhancing its clinical utility and accuracy.

4.4 Future Directions

This study lays the groundwork for multiple advancements in contour verification and radiation therapy planning. We identify three main areas for future exploration: clinical integration of the tool, a detailed analysis of the relationship between margins and dose constraints, and a long-term outlook on the method's impact and integration with emerging technologies.

4.4.1 Clinical Integration

Clinical Studies

Our tool has highlighted critical errors in retrospective analyses, suggesting the need for institutional improvements. Future studies should focus on correlating these errors with clinical outcomes to validate the tool's effectiveness. Comparative studies on the time efficiency of our tool across varied patient cohorts and treatment modalities are essential to ascertain its utility and adaptability.

Tracking the decision to make corrections or not based on errors found, as well as the number and amount of modifications made, will help identify trends in contouring errors, guiding targeted improvements in contouring and institutional practices.

Features and Functionality

Currently, the tool is configured for specific brain structures, and not directly applicable to prospective clinical use. Future developments should focus on adapting the tool for diverse anatomical sites and integrating a user interface that enhances ease of use and displays relevant outputs for planners. This adaptation includes expanding the scope of selective margin expansion to encompass additional dosimetric parameters such as mean dose and volumetric constraints. Given further modifications, this could also see use in other treatment planning system environments.

4.4.2 Relationship between Margins and Dose Constraint

Investigating the relationship between applied margins and dose constraints offers significant potential for enhancing radiation safety and efficacy. Future research should investigate not only maximum and mean doses but also include tumor volume, proximity to the isocenter, and the impact of treatment techniques and beam angles and orientations. These factors could provide deeper insights into the institutional variability in contouring practices and the effectiveness of Planning Risk Volumes (PRVs) in managing patient-specific uncertainties.

Planning Risk Volumes (PRVs)

Further exploration into the use of PRVs could be greatly enhanced with the use of our method, encouraging more research into the use of margins on organs at risk and how they might best be used to improve treatment planning and patient outcomes.

Even though our primary purpose of this study is to establish the method's viability in the region of contour errors, further insights into the correlation between different margin levels and

the likelihood of clinically significant errors, can allow us to potentially use PRVs to account for uncertainties like patient movement and anatomical changes while striking a balance between tumor control and toxicity risks.

4.4.3 Long-term Outlook

Our novel methodology introduces a quality assurance layer that enhances the contour verification process, with potential for broader application in treatment planning. Integrating AI could transform its scope of use, allowing for preemptive contour adjustments before dose optimization[12], enhancing treatment plan reliability and saving time. Meanwhile, the prediction of potential errors, as well as suggestions of relevant corrective actions to be taken would bring us much close to full automation. Future iterations of the tool should focus on seamless integration with other potential QA tools, applying their strengths to the specific regions identified by our method, minimizing OAR contour errors and their related toxicities.

5. Conclusion

In this thesis, we introduced a novel method designed to streamline the contour verification process through the identification of clinically relevant subsections. To demonstrate its viability, a tool was developed for use with critical structures such as the brainstem in a retrospective study of stereotactic radiosurgery (SRS) treatments.

Our tool successfully flagged contour subsections in past cases, revealing contour inaccuracies and highlighting the prevalence of such errors in previously approved treatments, including an error of 1.75mm, exceeding the tolerance limits for SRS. Such findings suggest that our tool will be able to help planners identify meaningful inaccuracies they otherwise might have missed. When the tool detected no errors, it allowed planners to concentrate their efforts on other aspects of the treatment plan, enhancing confidence in the treatment's safety and efficacy. This efficiency not only saves time but also reduces the labor involved in scrutinizing non-critical areas, potentially offering substantial clinical benefits.

Significant outcomes of this study include improved patient safety through meticulous identification and mitigation of potential contouring errors, increased efficiency by focusing on critical adjustments, and enhanced support for informed clinical decision-making. These achievements highlight the critical role of advanced QA tools in modernizing clinical workflows within radiation therapy.

Looking ahead, we aim to validate the application of this method across different anatomical structures and integrate it with emerging technologies in radiation therapy to further evaluate its impact on treatment outcomes and provide support for planners. This research addresses a crucial need for sophisticated contour QA tools, particularly against the backdrop of advancements in AI, autocontouring, and oART, thereby paving the way for further automated and precise radiation therapy modalities.

References

1. Abshire, D. & Lang, M. K. The Evolution of Radiation Therapy in Treating Cancer. *Semin. Oncol. Nurs.* 34, 151–157 (2018).

2. Hata, M. Radiation therapy for elderly patients with uterine cervical cancer: feasibility of curative treatment. *Int. J. Gynecol. Cancer* 29, 622–629 (2019).

3. Grimm, J. *et al.* High Dose per Fraction, Hypofractionated Treatment Effects in the Clinic (HyTEC): An Overview. *Int. J. Radiat. Oncol.* 110, 1–10 (2021).

4. Delaney, G., Jacob, S., Featherstone, C. & Barton, M. The role of radiotherapy in cancer treatment: Estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer* 104, 1129–1137 (2005).

5. Bortfeld, T. IMRT: a review and preview. *Phys. Med. Biol.* 51, R363–R379 (2006).

6. El-Bared, N., Wong, P. & Wang, D. Soft Tissue Sarcoma and Radiation Therapy Advances, Impact on Toxicity. *Curr. Treat. Options Oncol.* 16, 19 (2015).

7. Dawson, L. A. & Jaffray, D. A. Advances in Image-Guided Radiation Therapy. *J. Clin. Oncol.* 25, 938–946 (2007).

8. Mayo, Z. S. *et al.* Limited Toxicity of Hypofractionated Intensity Modulated Radiation Therapy for Head and Neck Cancer. *Anticancer Res.* 42, 1845–1849 (2022).

9. Tuleasca, C. *et al.* Biologically effective dose correlates with linear tumor volume changes after upfront single-fraction stereotactic radiosurgery for vestibular schwannomas. *Neurosurg. Rev.* 44, 3527–3537 (2021).

10. Benedict, S. H. *et al.* Stereotactic body radiation therapy: The report of AAPM Task Group 101. *Med. Phys.* 37, 4078–4101 (2010).

11. Zhang, T., Chi, Y., Meldolesi, E. & Yan, D. Automatic Delineation of On-Line Head-And-Neck Computed Tomography Images: Toward On-Line Adaptive Radiotherapy. *Int. J. Radiat. Oncol.* 68, 522–530 (2007).

12. Wang, C., Zhu, X., Hong, J. C. & Zheng, D. Artificial Intelligence in Radiotherapy Treatment Planning: Present and Future. *Technol. Cancer Res. Treat.* 18, 153303381987392 (2019).

13. Nguyen, D. *et al.* Advances in Automated Treatment Planning. *Semin. Radiat. Oncol.* 32, 343–350 (2022).

14. Vaassen, F. *et al.* Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy. *Phys. Imaging Radiat. Oncol.* 13, 1–6 (2020).

15. Li, G., Wu, X. & Ma, X. Artificial intelligence in radiotherapy. *Semin. Cancer Biol.* 86, 160–171 (2022).

16. Sherer, M. V. *et al.* Metrics to evaluate the performance of auto-segmentation for radiation treatment planning: A critical review. *Radiother. Oncol.* 160, 185–191 (2021).

17. Valentini, V., Boldrini, L., Damiani, A. & Muren, L. P. Recommendations on how to establish evidence from auto-segmentation software in radiotherapy. *Radiother. Oncol.* 112, 317–320 (2014).

18. Claessens, M. *et al.* Machine learning-based detection of aberrant deep learning segmentations of target and organs at risk for prostate radiotherapy using a secondary segmentation algorithm. *Phys. Med. Biol.* 67, 115014 (2022).

19. Hui, C. B. *et al.* Quality assurance tool for organ at risk delineation in radiation therapy using a parametric statistical approach. *Med. Phys.* 45, 2089–2096 (2018).

20. Nourzadeh, H. *et al.* Knowledge-based quality control of organ delineations in radiation therapy. *Med. Phys.* 49, 1368–1381 (2022).

21. Berlin, E. *et al.* Acute and long-term toxicity of whole pelvis proton radiation therapy for definitive or adjuvant management of gynecologic cancers. *Gynecol. Oncol.* 172, 92–97 (2023).

22. Grewal, A. S., Jones, J. & Lin, A. Palliative Radiation Therapy for Head and Neck Cancers. *Int. J. Radiat. Oncol.* 105, 254–266 (2019).

23. Bentzen, S. M. Preventing or reducing late side effects of radiation therapy: radiobiology meets molecular pathology. *Nat. Rev. Cancer* 6, 702–713 (2006).

24. Bentzen, S. M. *et al.* Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): An Introduction to the Scientific Issues. *Int. J. Radiat. Oncol.* 76, S3–S9 (2010).

25. Van Den Bosch, L. *et al.* Comprehensive toxicity risk profiling in radiation therapy for head and neck cancer: A new concept for individually optimised treatment. *Radiother. Oncol.* 157, 147–154 (2021).

26. Marks, L. B. *et al.* Use of Normal Tissue Complication Probability Models in the Clinic. *Int. J. Radiat. Oncol.* 76, S10–S19 (2010).

27. Miften, M., Gayou, O., Parda, D. S., Prosnitz, R. & Marks, L. B. Using Quality of Life Information to Rationally. in *Late Effects of Cancer Treatment on Normal Tissues* (eds. Rubin, P., Constine, L. S., Marks, L. B. & Okunieff, P.) 83–89 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008). doi:10.1007/978-3-540-49070-8_11.

28. Wright, J. L. *et al.* Standardizing Normal Tissue Contouring for Radiation Therapy Treatment Planning: An ASTRO Consensus Paper. *Pract. Radiat. Oncol.* 9, 65–72 (2019).

29. Yang, H. *et al.* Provision of Organ at Risk Contouring Guidance in UK Radiotherapy Clinical Trials. *Clin. Oncol.* 32, e60–e66 (2020).

30. Mir, R. *et al.* Organ at risk delineation for radiation therapy clinical trials: Global Harmonization Group consensus guidelines. *Radiother. Oncol.* 150, 30–39 (2020).

31. Nelms, B. E. *et al.* Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Pract. Radiat. Oncol.* 2, 296–305 (2012).

32. Qazi, A. A. *et al.* Auto-segmentation of normal and target structures in head and neck CT images: A feature-driven model-based approach. *Med. Phys.* 38, 6160–6170 (2011).

33. Mackay, K., Bernstein, D., Glocker, B., Kamnitsas, K. & Taylor, A. A Review of the Metrics Used to Assess Auto-Contouring Systems in Radiotherapy. *Clin. Oncol.* 35, 354–369 (2023).

34. Chen, W. *et al.* A comparative study of auto-contouring softwares in delineation of organs at risk in lung cancer and rectal cancer. *Sci. Rep.* 11, 23002 (2021).

35. Ng, C. K. C., Leung, V. W. S. & Hung, R. H. M. Clinical Evaluation of Deep Learning and Atlas-Based Auto-Contouring for Head and Neck Radiation Therapy. *Appl. Sci.* 12, 11681 (2022).

36. Gooding, M. J. *et al.* Comparative evaluation of autocontouring in clinical practice: A practical method using the Turing test. *Med. Phys.* 45, 5105–5115 (2018).

37. Nikolov, S. *et al.* Clinically Applicable Segmentation of Head and Neck Anatomy for Radiotherapy: Deep Learning Algorithm Development and Validation Study. *J. Med. Internet Res.* 23, e26151 (2021).

38. Maffei, N. *et al.* Radiomics classifier to quantify automatic segmentation quality of cardiac sub-structures for radiotherapy treatment planning. *Phys. Med.* 83, 278–286 (2021).

39. Mancosu, P. *et al.* Applications of artificial intelligence in stereotactic body radiation therapy. *Phys. Med. Biol.* 67, 16TR01 (2022).

40. Thompson, R. F. *et al.* Artificial intelligence in radiation oncology: A specialty-wide disruptive transformation? *Radiother. Oncol.* 129, 421–426 (2018).

41. Lappas, G. *et al.* Automatic contouring of normal tissues with deep learning for preclinical radiation studies. *Phys. Med. Biol.* 67, 044001 (2022).

42. García-Figueiras, R. *et al.* How Imaging Advances Are Defining the Future of Precision Radiation Therapy. *RadioGraphics* 44, e230152 (2024).

43. Shelley, C. E., Barraclough, L. H., Nelder, C. L., Otter, S. J. & Stewart, A. J. Adaptive Radiotherapy in the Management of Cervical Cancer: Review of Strategies and Clinical Implementation. *Clin. Oncol.* 33, 579–590 (2021).

44. Byrne, M. *et al.* Varian ethos online adaptive radiotherapy for prostate cancer: Early results of contouring accuracy, treatment plan quality, and treatment time. *J. Appl. Clin. Med. Phys.* 23, e13479 (2022).

45. McComas, K. N., Yock, A., Darrow, K. & Shinohara, E. T. Online Adaptive Radiation Therapy and Opportunity Cost. *Adv. Radiat. Oncol.* 8, 101034 (2023).

46. Glide-Hurst, C. K. *et al.* Adaptive Radiation Therapy (ART) Strategies and Technical Considerations: A State of the ART Review From NRG Oncology. *Int. J. Radiat. Oncol.* 109, 1054–1075 (2021).

47. Bessieres, I. *et al.* Online adaptive radiotherapy and dose delivery accuracy: A retrospective analysis. *J. Appl. Clin. Med. Phys.* 24, e14005 (2023).

48. Lim-Reinders, S., Keller, B. M., Al-Ward, S., Sahgal, A. & Kim, A. Online Adaptive Radiation Therapy. *Int. J. Radiat. Oncol.* 99, 994–1003 (2017).

49. Baroudi, H. *et al.* Automated Contouring and Planning in Radiation Therapy: What Is 'Clinically Acceptable'? *Diagnostics* 13, 667 (2023).

50. Isaksson, L. J. *et al.* Quality assurance for automatically generated contours with additional deep learning. *Insights Imaging* 13, 137 (2022).

51. Beasley, W. J., McWilliam, A., Slevin, N. J., Mackay, R. I. & Van Herk, M. An automated workflow for patient-specific quality control of contour propagation. *Phys. Med. Biol.* 61, 8577–8586 (2016).

52. Cha, E. *et al.* Clinical implementation of deep learning contour autosegmentation for prostate radiotherapy. *Radiother. Oncol.* 159, 1–7 (2021).

53. Delaney, A. R., Dahele, M., Slotman, B. J. & Verbakel, W. F. A. R. Is accurate contouring of salivary and swallowing structures necessary to spare them in head and neck VMAT plans? *Radiother. Oncol.* 127, 190–196 (2018).

54. Milano, M. T. *et al.* Single- and Multifraction Stereotactic Radiosurgery Dose/Volume Tolerances of the Brain. *Int. J. Radiat. Oncol.* 110, 68–86 (2021).

55. Trifiletti, D. M. *et al.* Stereotactic Radiosurgery for Brainstem Metastases: An International Cooperative Study to Define Response and Toxicity. *Int. J. Radiat. Oncol.* 96, 280–288 (2016).

56. Mayo, C., Yorke, E. & Merchant, T. E. Radiation Associated Brainstem Injury. *Int. J. Radiat. Oncol.* 76, S36–S41 (2010).

57. Narayanasamy, G., Smith, A., Van Meter, E., McGarry, R. & Molloy, J. A. Total target volume is a better predictor of whole brain dose from gamma stereotactic radiosurgery than the number, shape, or location of the lesions. *Med. Phys.* 40, 091714 (2013).