**Editors' comments:**

This paper reports on an important issue and presents a potentially very useful solution which can impact practices around the world. It is appreciable to see a physics submission with a well-designed software that is made available open-source to the users.

Thank you for the kind words!

We agree with the reviewer that the manuscript is not very easy to read and even a bit dull because it reads like a manual (even though not as detailed). The authors claim it is scalable and usable on multiple platforms so the manuscript would benefit a lot from more explanations and results on these particular aspects. Even if it is a brief report, the potential impact is large, and this should be highlighted more in the manuscript.

 We agree and have placed many of the manual aspects of the publication as supplementary. We have helped to clarify the text by adding the following at the end of the Introduction: *"Our aim in this work was to lower the barrier to adoption of TG-263 nomenclature in English, Spanish, or French by disseminating standardization that may facilitate data sharing. We have developed a tool which runs on any Windows system to easily create TG-263-compliant structure template libraries. Our tool can monitor folders and automatically add patient-specific structure sets, or create loadable RT structure/.xml templates and is a scalable solution focused on compatibility with all Treatment Planning Systems (TPS) utilizing the DICOM standard."*

Abstract: It is always difficult to fit software development in the Intro-methods&materials-results-conclusion structure but here it should be rewritten with more fitting sub-sections. We suggest having sub-section "methods and results" as in the paper. The aim is actually the last sentence in the methods and material sub-section. In fact the entire methods and materials describes rather the motivation and the purpose. Similarly, the Conclusion is more of a continuation of the results.

 We have rewritten the abstract as suggested to include Purpose, Methods and Results, and Conclusion sections as per below:

***Purpose:*** *Consistency of nomenclature within radiation oncology is increasingly important as big data efforts and data sharing become more prevalent. Automation of radiation oncology workflows depends on standardized contour nomenclature which enables toxicity and outcomes research, while also reducing medical errors and facilitating quality improvement activities.  Recommendations for standardized nomenclature have  been published in the American Association of Physicists in Medicine (AAPM) report from Task Group 263. Transitioning to TG-263 requires creation and management of structure template libraries, and retraining of staff, which can be a considerable burden on clinical resources. Our aim is to develop a program that allows users to create TG-263 compliant structure templates in English, Spanish, or French to facilitate data sharing.*

***Methods and Results****: 53 pre-made structure templates were arranged by treated organ based on an American Society for Radiation Oncology (ASTRO) consensus paper. Templates were further customized with common target structures, relevant OARs (e.g., Spleen for anatomically relevant sites such as gastroesophageal junction or stomach), sub-site specific templates (e.g. partial breast, whole breast, intact prostate, postoperative prostate, etc.) and the addition of brachytherapy templates from the AAPM brachytherapy working group. An informal consensus on OAR and target coloration was also achieved, though color selections are fully customizable within the program.  The resulting C# program is usable on any Windows system and generates template files in practice-specific DICOM or XML formats, extracting standardized structure nomenclature from an online database maintained by members of the TG-263U1 Task Group which ensures continuous access to up-to-date templates.*

***Conclusions:*** *We have developed a tool which runs on any Windows system to easily create TG-263-compliant structure template libraries for all planning systems utilizing the DICOM standard. The program and source code are publicly available via GitHub.  Feedback from community users is encouraged to identify opportunities for improvement and guide further development.*

Introduction: We suggest (e.g. "Brian" instead of "Brain") and (e.g. "Lung\_R" or "Right Lung") in a consistent manner and with "e.g." instead of "i.e."

Thank you for these comments. We have corrected these statements as suggested above.

We agree with the reviewer: this is not really a study; we suggest something along the lines of "our aim was to develop... to lower the barrier..."

Thank you for your comments. We have corrected these statements and listed the last sentence of the Purpose in the abstract as follows: *"Our aim is to develop a program that allows users to create TG-263 compliant structure templates in English, Spanish, or French to facilitate data sharing."* and the last paragraph of the introduction as follows *"Our aim in this work was to lower the barrier to adoption of TG-263 nomenclature in English, Spanish, or French by disseminating standardization that may facilitate data sharing. We have developed a tool which runs on any Windows system to easily create TG-263-compliant structure template libraries. Our tool can monitor folders and automatically add patient-specific structure sets, or create loadable RT structure/.xml templates and is a scalable solution focused on compatibility with all Treatment Planning Systems (TPS) utilizing the DICOM standard."*

Figure 1: We agree with the reviewer here. The figure is very difficult to read due to the small font size. This figure should really illustrate the workflow in a simple schematic way. The videos are not that easy to find on the github page. 1-2 videos would be a very valuable addition as supplementary material to the article.

Thank you, we will adjust the Figure accordingly [new figure stuff here, probably should be a new graphical abstract as discussed below]. We will continue to add videos and have made the link more obvious.

We did not find a table S1, please make sure supplementary files are uploaded in the revision.

Thank you for drawing this to our attention. We have now included the original Table S1 in the supplementary.

For screenshot figures (2,3,4), adding a few arrows, boxes, or numbers referred to in the legend would be helpful for the reader.

Thank you, we have improved the figures are requested.

**Reviewer 1 Comments to Author: Open RT Structures: A Scalable Solution for TG-263 Accessibility**

General comments:

Thank you for the opportunity to read this paper. The submission is interesting and can potentially add significant value to the radiation oncology community.

1. Please comment on the importance of this question and the originality of the findings for the readers of Red Journal.

The paper's question is how standardized naming conventions can be efficiently and consistently obtained and used in radiation treatment planning. While the extent of this problem and the repercussions of mislabeled structures is not well described, the problem is significant enough to warrant an AAPM Task Group on the problem.

Thank you! We agree there should be a way facilitate buy-in to help address the problem of mislabeled structures and to encourage adoption of TG-263 compliant nomenclature. We hope this software will lower the barriers to adoption.

The originality of the work stems from a few things. First, the authors provide open-source code to create these structures. It needs to be noted that not every clinic will have the skills necessary to deploy such software in their clinic. Second, the structures themselves abide by standardized structure nomenclature and provide DICOM RT Structure Set data. Finally, although less clear to the reader, the datasets generated are theoretically vendor agnostic as they would adopt DICOM (SS) standards.

Thank you for this comment. We do agree that not every clinic will have the skills to deploy the software on a case-by-case basis within the clinic. However, every clinic should have the ability to generate these structure sets and import them into Eclipse/TPS. We hope that the videos will provide more step-by-step help for those less technically savvy.

2. Please comment on the appropriateness of the study approach and experimental design. (Examples: retrospective or prospective cohort, case-control, cross-sectional, ecological, case series; clinical trial or secondary analysis of clinical trial; registry-based; critical review; metaanalysis or systematic review; experimental, based on cell cultures, animal models, physical models, or method/technique development.)

The submission is not a study, and thus details for experimental design and study approach are not needed; however, details of the range of usefulness of the proposed software is lacking. Has it been tested and validated in common vendor platforms? If so, what is the software version this has been tested on? These details would be valuable for users who do not have ready access to the 'state-of-the-art' versions of treatment planning software. Also, while I don't expect the users to provide intricate details of the software platforms these routines were created, some basic information on the software version should be provided, along with the operating software for users to understand what tools they need to have in place if they chose to use this software.

Thank you for this comment. This software has been tested and confirmed working with Eclipse (v15.6), Raystation (v12.1) and Pinnacle (16.2.1), but since the program generates DICOM/XML files, it should work with all versions of software programs which rely on the DICOM standard.

3. Please comment on the appropriateness and reproducibility of the data collection and experimental techniques. (If applicable, does the study comply with the CONSORT, PRISMA and/or REMARK statements? If applicable, was the study IRB-approved or registered on [clinicaltrials.gov](https://urldefense.com/v3/__http:/clinicaltrials.gov__;!!LLK065n_VXAQ!j7OJcW2P1nXV-uVyb732wBTpOqdZTsEaJbRMe79wmMeBF2k2Mm599y-0WwLnKCMOy6VVbUUfZZfOHgouEEHgCw$)?

See Point 2.

4. Please comment on the analysis and interpretations of the data. Do you agree with the proposed conclusions?

Overall, I think the presentation of the workflow is great and I think would be a potentially useful tool for the community. However, there are a number changes to the document that should be made to make it more readable and highlight the shortcomings, in addition to the issues raised above.

Generally, the paper reads as to 'what' is done, and in several places, it needs to describe 'why' it is done. I encourage the authors to explain why a step was undertaken in creating these datasets when there is no explanation provided.

Thank you for this comment. We agree and have moved the focus away from 'what' is done to describe 'why' it was done. We have worked to explain why steps were taken and are open to additional feedback. Please see the sections below which we hope focus more on the ‘why’ aspect of our decision making process.

“The program was written (BMA) using C#3, ensuring it’s computability with windows systems.”

“The program was piloted at multiple sites with Eclipse (JR, KS, DH), Pinnacle (RZ) and Raystation (CE) to ensure compatibility with multiple treatment planning systems (TPS). We wanted to make the model output compatible with as many TPS as possible, and so ensured the model output follows the DICOM standard.”

“The 53 pre-made structure templates are arranged by treated organ and include treatment sites based on an American Society for Radiation Oncology (ASTRO) consensus paper2.”

“We wanted to ensure that users could benefit from the previously created templates that follow TG-263 nomenclature, but also have the ability to create their own templates as desired within their clinic.”

“If the user has pre-existing templates in Varian .xml file format, they can be easily added to allow for template modification within our program. This plugin was created to remove any headache of exporting templates from the Varian system to our program.”

5. Please comment on weaknesses or limitations of the study. (Examples are: selection biases, sample size limitations, missing data.)

The major limitation, as noted earlier, is the range of usefulness. The authors should identify what platforms this has been tested on (e.g., can it be used successfully in all major TPS, details of the additional computing platform needed). The authors should also provide information on how long the process takes to convert a patient's DICOM SS to another.

Thank you for this comment. We have tested this on Eclipse (v15.6), Raystation (v12.1) and Pinnacle (16.2.1) and this program should be compatible with all versions of any TPS utilizing the DICOM standard.

6. Please comment on the writing and organization of the paper. Is the paper overly wordy? Is the English language acceptable?

The paper is well written with few (if any) grammatical issues.

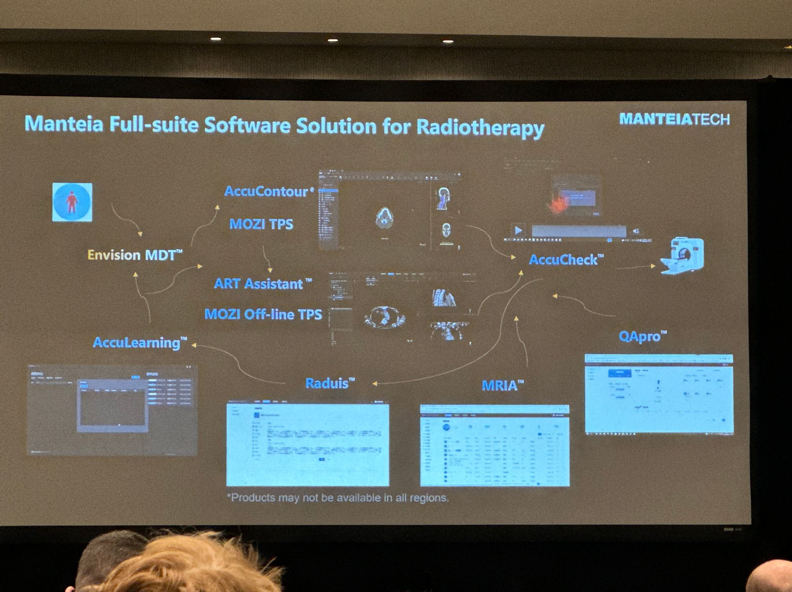
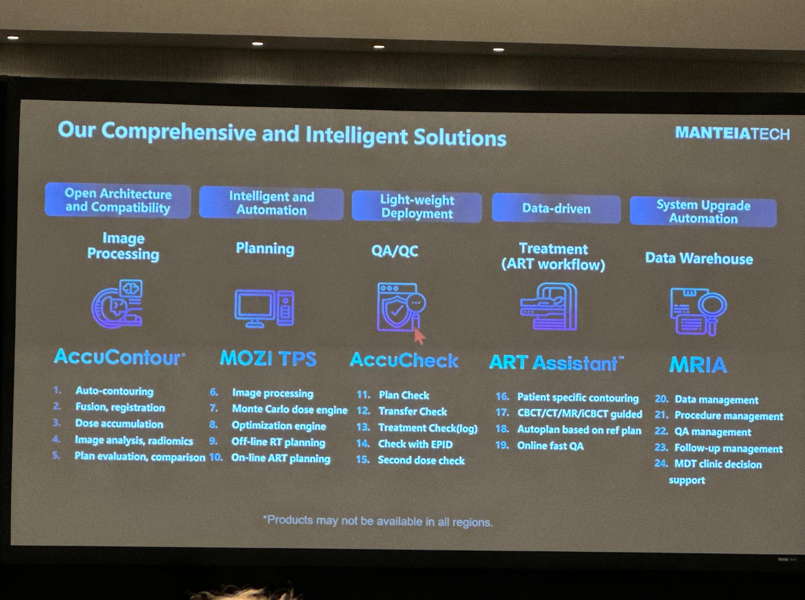
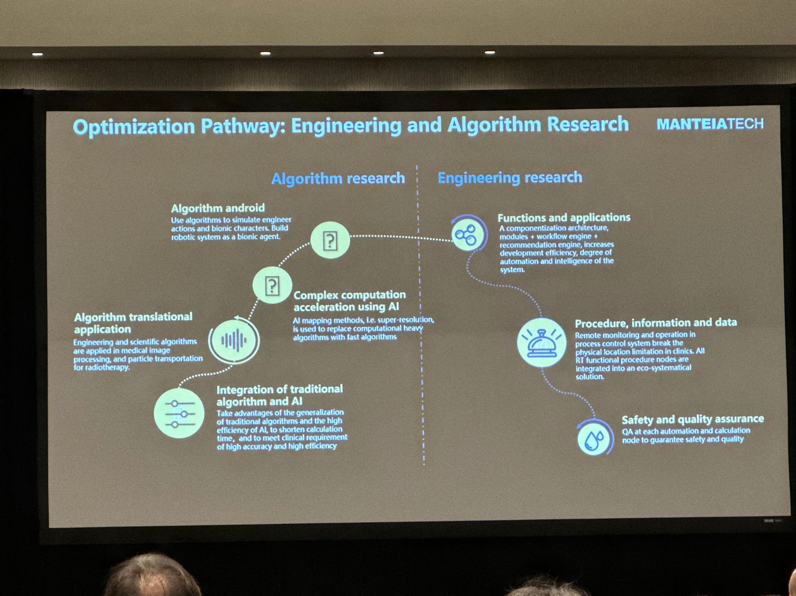
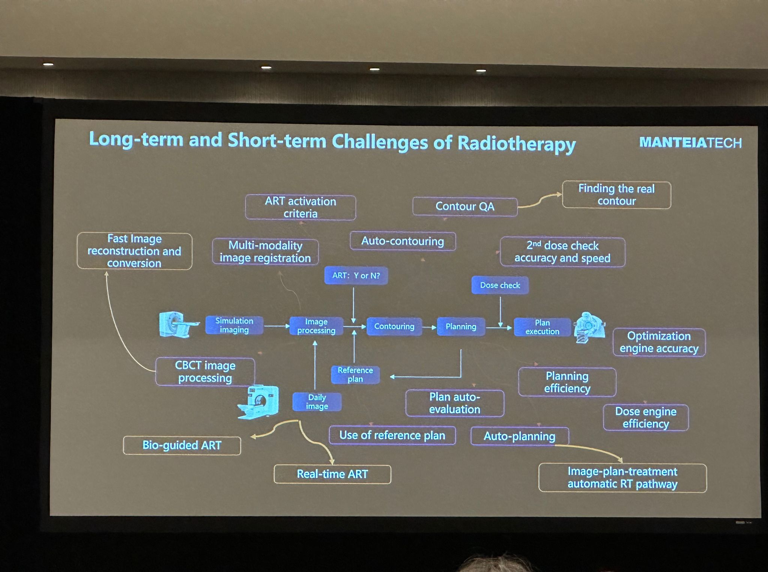
Thank you!

7. Please comment on the necessity and clarity of the figures and tables. Can they stand independently of the text?

Figure 1 is not helpful at all. I appreciate the authors want to display the interface of the software, but if the purpose of the figure is to demonstrate the workflow, it could be more easily done with a simplified figure with descriptive text. The font is too small to read and will not render well in published form. I strongly encourage the authors to create a simplified figure. The entire work could be synthesized in a well-designed graphical abstract (and frankly, this is probably the best way to transmit the work!)

Thank you, we will adjust the Figure accordingly [new figure stuff here - preferably, a graphical abstract].

Here are some graphical abstracts from MANTEIAtech - maybe we can use these as a blueprint for the graphical abstract?



8. Please comment on any need for formal statistical review.

None

Specific comments (no line numbers were provided in the revision version)

Title: The paper does not provide much evidence for the 'scalability' of the software.

Thank you. We hope this addition to the introduction will help to clarify: *"Our aim in this work was to lower the barrier to adoption of TG-263 nomenclature in English, Spanish, or French by disseminating standardization that may facilitate data sharing. We have developed a tool which runs on any Windows system to easily create TG-263-compliant structure template libraries. Our tool can monitor folders and automatically add patient-specific structure sets, or create loadable RT structure/.xml templates and is a scalable solution focused on compatibility with all Treatment Planning Systems (TPS) utilizing the DICOM standard."* If this is not agreeable, we will gladly remove "scalable solution" from the title of the manuscript.

Keywords: Is "RT" a keyword? TG-263 and RT are already in the title

We felt that the inclusion of ‘TG-263’ are still helpful in finding the work, although we have remoted ‘RT’

Abstract: no major comments currently.

Introduction:

"In a recent ..." is there a reference to support this claim?

"Our aim in this study..." this is not a study, and thus consider rephrasing.

Thank you for these comments. We have added the reference to support this claim. [Add reference here].

We have corrected this statement as described above, and said "Our aim of this work" in two places as follows: in the last sentence of the Purpose in the abstract as follows: *"Our aim is to develop a program that allows users to create TG-263 compliant structure templates in English, Spanish, or French to facilitate data sharing."* and the last paragraph of the introduction as follows *"Our aim in this work was to lower the barrier to adoption of TG-263 nomenclature in English, Spanish, or French by disseminating standardization that may facilitate data sharing. We have developed a tool which runs on any Windows system to easily create TG-263-compliant structure template libraries. Our tool can monitor folders and automatically add patient-specific structure sets, or create loadable RT structure/.xml templates and is a scalable solution focused on compatibility with all Treatment Planning Systems (TPS) utilizing the DICOM standard."*

Material and Methods:

(Step 3:

"A file system..." Why is this done? Presumably, to ensure the entire DCM dataset was intact prior to editing?

Thank you for the comment, we have clarified this further.

“This is performed since DICOM images are often uploaded to a server after acquisition on the CT. The upload process can take time, depending on the size of the scan and latency of the network. This file system watcher ensures the entire DICOM dataset is present before an RT structure is generated.”

"If the DICOM images... ", why provide this option again? (Readers will want to know).

Thank you for the comment! We have attempted to clarify this for the reader.

“If the DICOM images are consistently placed within the same folder (server location post acquisition where all acquired images are deposited), the users can define values within the Series Description or Study Description to indicate which template should be run automatically. For example, including the tag ‘Breast\_CW’ in the Series Description during acquisition could ping the program to automatically create the ‘Breast\_CW’ template.”

"This ensures..." This sentence highlights a risk that needs to be discussed in the discussion section (see later).

Thank you for the comment, we have added the following to the discussion section.

“The largest risk that we could foresee is that the program continually updates it’s own previously generated RT Structure files. To ensure this does not happen, the program internally tracks which images have been previously viewed (via Series Instance UID), and creates each RT Structure file with that same Series Instance UID. The program never opens or edits an already existing RT Structure file, and so presents no risk to work flows already present by the user.”

Results: no comments

Discussion:

A detailed paragraph describing the range of usefulness should be provided here, along with another few sentences on the possible shortcomings and errors that could be introduced with the software. Many users will be risk-averse to the use of such a system without an understanding of risks.

Thank you for this comment, we have addressed any potential errors in the paragraph above.

References:

References 5 and 13 look odd. Are these HTML?

These are from visited web-pages. The references have been expanded upon

Figures:

Figure 1 should be revised as stated earlier

Thank you - we have created a graphical abstract.

Figure 2: A more descriptive caption would be helpful here. What am I looking at, and where do you want me to look?

We have revised the caption and provided a revised Figure 2.

Figure 3: no comments

Figure 4: no comments

Tables: na

Sincerely,

A Reviewer

**Reviewer 2 Comments to Author: 1.** Please comment on the importance of this question and originality of the findings for the readers of Red Journal.

     Tool sets for enabling implementation of TG-263 and TG-263U1 are valuable and important to improving data use within radiation oncology. The present treatment has three primary deficits which should be addressed. 1) It needs to be completely clear from statements in the publication, figures, results  that the proposed software tool is strictly following TG-263 and TG-263U1 guidelines for nomenclature of structures, DVH metrics, brachy reference points etc. TG-263 is endorsed by multiple professional societies.  Deviations from the consensus guidelines undermines the professional society based approach. 2) There is no reporting of results from implementation of the tool at multiple centers for the range of applicable technologies. This demonstration is important proving the viability of the approach, 3) TG-263 structure naming has been endorsed in the HL7-FHIR standard as the basis for structure naming, including mappings to SNOMED CT. This mapping and coding will be the basis of information transformation systems incorporated into vendor systems. The present treatment is promoting an alternative standardization, FMAID which is inconsistent with current standardization efforts.

Thank you for these comments. We have added a few things to the text in the discussion as follows: "In this paper, we describe the first reported effort to create open-source software to create and maintain libraries of patient-specific treatment planning structure templates to lower the barrier to adoption of TG-263 standardized nomenclature and facilitate data sharing for toxicity and outcomes research. All outputs are compatible with TG-263 and TG-263U1 guidelines for nomenclature of structures, which is endorsed by multiple professional societies." "This software was tested at multiple sites and ensured to be compatible with Pinnacle v16.2.1, Raystation v12.1, and Eclipse v15.6, although output should be compatible with all TPS utilizing the DICOM standard."

We have also added SNOMED CT codes to all structures at the reviewers request. We are currently creating a templated list between FMA and SNOMED and aim to include it as a part of the program. FMAID is the current standard at the clinics where the program was evaluated. Since there are many structures (over 600) we want to ensure that we take the time to carefully curate the list of FMAID and SNOMED codes. For this reason, we have translated the 200 which are included in the templates already, and are actively translating the entire list provided by TG-263 in a methodical manner.

2. Please comment on the appropriateness of the study approach and experimental design. (Examples: retrospective or prospective cohort, case-control, cross-sectional, ecological, case series; clinical trial or secondary analysis of clinical trial; registry-based; critical review; metaanalysis or systematic review; experimental, based on cell cultures, animal models, physical models, or method/technique development.)

Lacking in clinical use validation study

Thank you for this comment. We did informally test this program at multiple institutions (see previous comment above). We intend to complete a follow up survey on the program for a more formal clinical use validation study.

3. Please comment on the appropriateness and reproducibility of the data collection and experimental techniques. (If applicable, does the study comply with the CONSORT, PRISMA and/or REMARK statements? If applicable, was the study IRB-approved or registered on [clinicaltrials.gov](https://urldefense.com/v3/__http:/clinicaltrials.gov__;!!LLK065n_VXAQ!j7OJcW2P1nXV-uVyb732wBTpOqdZTsEaJbRMe79wmMeBF2k2Mm599y-0WwLnKCMOy6VVbUUfZZfOHgouEEHgCw$)?

No data evident of use of the tool in multi-institutional settings for data collection.

4. Please comment on the analysis and interpretations of the data. Do you agree with the proposed conclusions?

5. Please comment on weaknesses or limitations of the study. (Examples are: selection biases, sample size limitations, missing data.)

Missing data, inconsistent of ambiguous recommendations on standardizations

Thank you for this comment. We attempted to clarified the ambiguous recommendations on standardization as in our first reply to this reviewer.

6. Please comment on the writing and organization of the paper. Is the paper overly wordy? Is the English language acceptable?

Acceptable

7. Please comment on the necessity and clarity of the figures and tables. Can they stand independently of the text?

8. Please comment on any need for formal statistical review.

Thank you all so much for the commentary. We have worked hard to address all concerns and believe the paper is stronger and less ambiguous as a result. Please let us know if anything else is needed, and we would be happy to address any additional  issues.