"A geometric analysis of Brainlab auto-contouring software for proton treatment planning of brain tumours"

Virginia Marin Anaya[1]; Alexander Grimwood[1]; Jaymisha Davda[1]; Caroline Thould[1]; Emma Dwyer[1]; Colin Baker[1]

(1) University College London Hospitals NHS Foundation Trust

Background:

Contouring of relevant structures in the vicinity of the tumour is currently performed manually. This is time-consuming, subjective and can delay the start of treatment. For brain patients, this can lead to poorer clinical outcome. Moreover, proton therapy is very sensitive to anatomical changes and re-planning, including re-outlining of structures, may be necessary. The aim of this study was to assess the feasibility of auto-contouring for proton treatment planning of brain tumours using Brainlab Elements version 1.6.1.38.

Methods:

Ten brain patients were selected retrospectively. The anonymised CT and MRI datasets were imported into Brainlab. For each patient, CT and MRI image fusion and distortion correction were performed. For brain, lenses, optic nerves, globes, cochleas and pituitary, CT was used for the generation of Brainlab auto-contours. MRI was selected for brainstem, chiasm, hippocampi, hypothalamus and cerebellum. Geometric analysis of the Brainlab contours was performed using several evaluation metrics such as the Dice Similarity Coefficient (DSC), the Mean Distance to Conformity (MDC) and the Target Registration Error (TRE). The manual contours on the planning CT by the oncologist were used as reference.

Results and discussion:

Table 1. Geometric analysis. Results expressed as median and range between brackets.

Structure	DSC	MDC (mm)	TRE (mm)	Volume Difference (%)	Sensitivity Index	Inclusiveness Index
Brain	0.97 (0.97, 0.98)	4.42 (3.55, 4.94)	1.01 (0.04, 4.02)	-4.18 (-4.57, -2.55)	0.95 (0.95, 0.97)	0.99 (0.99, 1.00)
Brainstem	0.89 (0.84, 0.90)	4.37 (3.10, 5.46)	1.19 (0.30, 3.34)	-0.38 (-8.82, 16.32)	0.88 (0.85, 0.91)	0.89 (0.78, 0.93)
Cerebellum	0.92 (0.91, 0.95)	3.48 (3.01, 3.80)	0.95 (0.43, 2.04)	-0.65 (-5.38, 2.46)	0.92 (0.90, 0.94)	0.93 (0.91, 0.95)
Chiasm	0.50 (0.22, 0.65)	4.10 (3.04, 8.69)	3.35 (2.00, 7.88)	-6.70 (-35.96, 68.42)	0.50 (0.19, 0.66)	0.51 (0.26, 0.75)
Cochlea Left	0.38 (0.20, 0.70)	3.37 (2.41, 4.06)	1.99 (0.61, 3.50)	26.79 (-33.33, 300.00)	0.48 (0.25, 0.67)	0.33 (0.13, 0.88)
Cochlea Right	0.52 (0.25, 0.78)	2.74 (1.95, 3.98)	1.24 (0.06, 4.17)	45.00 (-33.33, 350.00)	0.64 (0.43, 1.00)	0.44 (0.18, 0.88)
Globe Left	0.93 (0.83, 0.95)	2.36 (2.17, 3.06)	0.77 (0.07, 1.42)	-11.17 (-27.71, 23.54)	0.88 (0.72, 0.99)	0.98 (0.80, 1.00)
Globe Right	0.92 (0.87, 0.95)	2.43 (2.24, 2.87)	0.87 (0.28, 1.41)	-5.57 (-20.82, 6.64)	0.89 (0.78, 0.95)	0.96 (0.88, 0.99)
Hippocampus Left	0.65 (0.54, 0.73)	4.19 (3.38, 6.36)	2.29 (1.43, 6.81)	43.09 (5.07, 75.00)	0.80 (0.64, 0.89)	0.54 (0.46, 0.67)
Hippocampus Right	0.66 (0.51, 0.73)	3.92 (3.20, 6.44)	2.02 (0.77, 6.79)	34.50 (11.60, 73.95)	0.75 (0.65, 0.90)	0.59 (0.42, 0.65)
Hypothalamus	0.53 (0.11, 0.65)	3.82 (2.96, 5.10)	2.90 (1.35, 4.34)	53.80 (12.75, 1685.71)	0.70 (0.62, 1.00)	0.43 (0.06, 0.60)
Lens Left	0.74 (0.54, 0.83)	2.30 (1.39, 3.24)	1.09 (0.59, 2.35)	51.88 (27.78, 110.00)	0.92 (0.77, 1.00)	0.62 (0.41, 0.74)
Lens Right	0.73 (0.43, 0.85)	2.25 (1.01, 3.10)	1.04 (0.14, 2.08)	48.08 (13.33, 120.00)	1.00 (0.50, 1.00)	0.60 (0.38, 0.74)
Optic Nerve Left	0.61 (0.22, 0.68)	3.38 (2.73, 5.96)	2.07 (0.92, 9.23)	-42.07 (-59.70, -28.26)	0.46 (0.16, 0.58)	0.82 (0.36, 0.97)
Optic Nerve Right	0.54 (0.40, 0.65)	3.22 (2.57, 5.85)	2.84 (0.37, 5.35)	-57.52 (-66.67, -40.74)	0.38 (0.28, 0.51)	0.90 (0.70, 0.95)
Pituitary	0.40 (0.18, 0.46)	4.18 (3.86, 5.19)	2.75 (2.03, 3.44)	41.03 (-51.85, 900)	0.52 (0.30, 1.00)	0.37 (0.10, 0.62)

For brain, brainstem, cerebellum and globes, median DSC values were ≥0.89 and range DSC values were between 0.83 and 0.98. For all Brainlab contours, median MDC and TRE values were ~ 2-4mm and ~1-3mm, respectively.

Qualitative analysis of auto-contours by oncologists showed a preference towards editing auto-contours, if necessary, rather than outlining from scratch, saving overall contouring time.

Conclusion:

Brainlab is a promising tool for proton treatment planning of brain tumours. Its implementation could potentially improve contouring consistency; optimise clinical workflow, increasing patient throughput, whilst enabling effective use of staff resources and improving patients' outcome.

Feasibility of a simple KBP planning tool for head and neck radiotherapy planning.

Matthew Jones[1], Julia Handley[2], Gareth Baugh[1].

[1] Arden Cancer Centre, University Hospitals Coventry and Warwickshire, Coventry.
[2] The Christie NHS Foundation Trust, Manchester.

Background. Investigate the potential application, utilisation and clinical implementation of a simple knowledge-based planning solution for head and neck radiotherapy within a clinical radiotherapy department.

Methods. A knowledge base of 141 previously treated head and neck patients was created by extracting data using a Python data mining script and the existing scripting capabilities of the RayStation treatment planning system. This knowledge base was used to create three separate knowledge-based models to predict the optimal and mandatory achievable doses for the spinal cord, brainstem, and parotids respectively. The models were validated using a range of methods. A graphical user interface was developed and validated to display the predicted model doses from within the planning system.

Results and Discussion. It was demonstrated that the three models developed could accurately identify treatment plans in which the doses to the brainstem, spinal cord and parotids could be reduced without adversely affecting any other aspects of treatment plan quality. For a separate cohort of validation head and neck patients, it was shown that implementing the models could potentially reduce the maximum spinal cord, maximum brainstem and mean parotid doses by 5.42Gy, 3.62Gy and 5.93Gy respectively without adversely affecting plan complexity and surrounding organ at risk doses. It was also demonstrated that the developed GUI was accurate and could feasibly be introduced into routine clinical use.

Conclusion. Three simple knowledge-based models have been developed and validated which could be clinically implemented and potentially significantly reduce organ at risk doses for head and neck patients within the clinical radiotherapy department. These models present a low cost, accessible, and simple alternative to commercially available knowledge-based planning solutions.

Automating 4D Manual Delineation Treatment Pathways

M Tyyger [1] , A Clark [1], DW Smith [1], C Lai [1], M Nix [1], P Dickinson [1], B Al-Qaisieh [1], I Bond [1]

1. Leeds Cancer Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Background

The focus of automation in treatment planning is typically on radiotherapy planning or auto-contouring. However, the assistance of a script for even simple tasks can provide large benefits in treatment pathways. This is particularly true where many similar and repetitive actions are required, as this type of activity is prone to human error.

A local example is the workflow required for delineation of tumour volumes on 4D datasets for lung SABR treatments. The treatment planning system (TPS) did not provide suitable tools for a pathway without many repetitive actions, such as creating regions of interest (ROI) and copying their geometries between examinations. To perform this manually was deemed clinically unsuitable due to the high likelihood of errors, the training burden on oncologists, and the length of time required.

The aim was to produce a Python script which ran inside the TPS Python environment which could automate the non-delineation steps while guiding the user through the treatment pathway. Therefore, making the process more efficient and reducing the probability of errors.

Methods

Local Radiotherapy Physics and Clinical Scientific Computing teams collaborated to define a clinically robust treatment pathway and develop the necessary Python script. A Consultant Clinical Oncologist evaluated the suitability of the pathway and script.

To assess potential time saving impacts, two experienced RayStation users, a senior Dosimetrist and Medical Physicist, were timed performing the functional steps of the workflow both manually and using the script.

Results

A pathway (figure 1) and script were produced, verified, and validated as being clinically suitable.

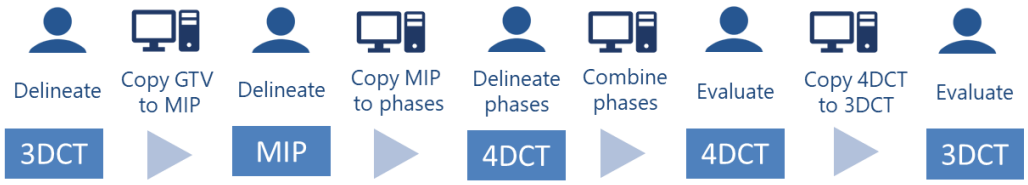


Fig 1. Developed delineation pathway, manual and automated steps shown.

The average time saved using the script was found to be approximately 6 minutes.

Discussion

Software development took longer than anticipated due to unexpected TPS behaviours, which required altering the pathway or recreating existing TPS functions to behave in the desired manner. The script went through several cycles of development to accommodate these adjusting behaviour requirements.

The developed script noticeably reduced the number of steps a user was required to perform, and pop-up notifications at each step informed the user exactly what to perform next.

Conclusion

The potential for improving radiotherapy treatment pathways via automation has been presented in the context of 4D delineation for lung SABR radiotherapy. In this scenario, a reduction in necessary staff training, error likelihood, and the time required has been shown.

Key references.

Reducing Region of Interest Export Errors Through Automation

M Tyyger [1], I Bond [1]

1. Leeds Cancer Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Background

Between the complexities of radiotherapy pathways and treatment planning systems (TPS), it is unsurprising that staff can frequently make minor mistakes when exporting treatment data for later use. Locally, it was identified that approximately 25% of reported Radiotherapy Physics errors over six months were related to incorrectly exported regions of interest (ROIs) from the TPS. Whilst it is rare these mistakes would lead to adverse patient outcomes, they can cause treatment delays and require staff to spend time fixing subsequent issues. Automation can mitigate these minor, but frequently problematic, TPS tasks.

Here, we discuss “SetExportSettings” a simple script designed to automatically set ROI export settings in a TPS to ensure only the correct structures are exported.

Methods

The requirements were a Python script, usable within a TPS Python environment which correctly sets a flag inside the TPS whether to export an ROI based on, its type (target, organ at risk, or other), and their name. The designed logic was meant to apply to all current and future treatment pathways. It was preferable to avoid using pathway-specific configuration files for the anticipated 40+ pathways due to the overhead of producing and maintaining those files.

Results

A script was developed, verified, and validated by Clinical Scientific Computing and Radiotherapy Physics staff members.

It used a single external configuration file to allow Radiotherapy Physics to change some behaviours of the script without requiring Scientific Computing input.

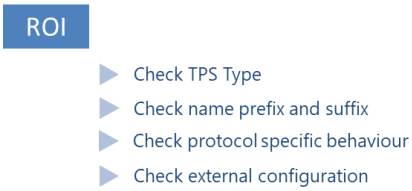


Figure 1. Performed checks for each ROI to set the export setting

An audit is on-going to assess the impact in reported errors since the clinical deployment of the script.

Discussion

During development it became clear that managing all treatment pathways without requiring the script to be altered was potentially an unachievable goal. Certain pathways were found to have conflicting behaviours for the same ROIs. Therefore, where necessary pathway-specific logic was added.

Clinical deployment of the script did not raise any unexpected issues, and initially appears to have reduced the number of errors seen. However, an on-going audit is being performed to ensure systematic errors have not been introduced for any individual treatment pathway.

Conclusion

This project has shown the potential to reduce common radiotherapy errors occurring from incorrect settings inside a TPS using simple automation.

Key references.

Title of Study: Evaluation and clinical implementation of deep learning auto-segmentation across all clinical sites

Josh Mason, Sarah Robinson, Ingrid Johnson, Jack Doherty, Jack Miskell, Meagan de la Bastide, Ruth McLauchlan
Imperial College Healthcare NHS Trust

Abstract no more than 1 page in Arial 11 point, presenting speaker underlined

Background.
Deep learning segmentation (DLS) can automate region of interest (ROI) delineation in radiotherapy treatment planning, offering the potential for significant time saving, improved efficiency and improved consistency/adherence to guidelines.

At Imperial College Healthcare NHS Trust DLS has been implemented for all planned radiotherapy treatments. This study describes the work done to evaluate DLS solutions from two commercial vendors and ensure safe implementation into the clinical pathway.

Methods.
5-10 patients per clinical site (45 patients in total) were evaluated retrospectively by comparing the manually contoured ROIs used in their clinical treatment to the respective DLS generated ROIs. Qualitative evaluation by experienced planners and clinical oncologists involved rating each DLS ROI on a 1-4 scale. Quantitative evaluation compared manual and DLS ROIs geometrically using DICE similarity coefficient (DSC) and dosimetrically by comparing dose volume histogram (DVH) statistics for the clinical plan calculated for manual and DLS ROIs. Automated scripts were used to assist evaluation and to simplify the process of adjusting DLS generated ROIs. A workflow for clinical implementation was developed and each clinical site is being audited a few months after implementation to ensure DLS ROIs are being reviewed and adjusted appropriately.

Results.
From qualitative evaluation, all ROIs were considered suitable for use with manual review and adjustment. Specific issues for users to look out for and differences from local contouring practice were identified. Quantitative DSC results varied especially due to differences in the superior-inferior extent that structures were contoured to. Dosimetric evaluation showed the differences between manual and DLS ROIs mostly had clinically insignificant impact on DVH values, though specific issues were identified for certain OARs particularly brainstem and optic pathway ROIs. Clinical implementation has been effective with the one issue identified being staff remembering to delete unwanted ROIs that otherwise have the potential to cause confusion later in the patient pathway. An audit has been completed for breast and thorax ROIs showing safe implementation although it also identified that use of DLS ROIs resulted in clinically insignificant changes to contouring practice due to the user being guided by the DLS ROI to some extent.

Discussion.
Evaluating and implementing DLS is a significant amount of work however both qualitative and quantitative evaluation are useful to identify potential issues with specific ROIs before proceeding to clinical implementation. Post-implementation audits are useful for better understanding the impact of clinical implementation.

Conclusion.
Deep learning auto-segmentation has been successfully implemented across all clinical sites. Further work will assess the impact of DLS ROIs in terms of time saving and impact on staff workload through regular user surveys.

Key references.
Automation, Deep learning segmentation

Failure rates and Quality Assurance of commercial AI auto-segmentation systems for head and neck cancer

Simon Temple, Carl Rowbottom
The Clatterbridge Cancer Centre NHS Foundation Trust

Background.
AI-based commercial software can be used to automatically delineate organs at risk (OAR) on CT scans, with the potential for significant efficiency savings in the radiotherapy treatment planning pathway, and simultaneous reduction of inter- and intra-observer variability. It is important that a suitable Quality Assurance (QA) program is implemented for such systems[1], which requires a good understanding of expected failure rates and the reason for these failures.

Methods.
A commercial AI auto-segmentation system was used to generate four commonly used OARs on 500 anonymised H&N patient datasets. Auto-segmented contours were compared to existing clinical contours, outlined by an expert human, and a failure rate was set at three standard deviations below the expected mean Dice Similarity Coefficient (DSC), based on a previous study[2]. Failures were classified into one of five groups (setup position, anatomical, image artefacts, suboptimal clinical contour and unknown). Failures relating to suboptimal contouring of the original clinical structure were removed, to produce a 'true failure' rate for each OAR.

Final true failure rates were used to inform recommendations for system QA.

Results.
The study resulted in consistently high quality AI auto-segmentation with a commercial system for H&N cancer patients, with few failures from a large sample size. A summary of results are given below.

Table 1. AI auto-segmentation failure rates for 500 patients

	Brainstem	**Mandible**	**Lt Parotid**	**Rt Parotid**
Total Failures	4	20	13	7
Failure Reason:				
Setup position	2	0	0	1
Anatomical	0	8	5	2
Dental artefacts	0	3	0	1
Clinical structure suboptimal	2	9	6	3
Unknown	0	0	2	0
True failures (Total – clnical error)	**2**	**11**	**7**	**4**
True failure rate	**0.4%**	**2.2%**	**1.4%**	**0.8%**

Discussion.
Where true failures of the auto-segmentation system were identified, there was often a non-standard element associated with the planning CT dataset, for example unusual setup position or unusual anatomy. It can be hypothesised that these non-standard elements were the cause of the failure, and further suggested that the patient datasets used to train the DL model did not contain sufficient heterogeneity of patient data.

Conclusion.
The true failure rate for AI auto-segmentation systems in the H&N region for the OARs investigated is extremely low, in the range 0.5-2%. Due to this very low failure rate, human inspection alone is unlikely to be effective or efficient in identifying failures. It is therefore advised that QA of auto-segmented OARs should utilise automated methods.

Keywords: AI auto-contouring, Quality assurance.

Key references.
1. Vandewinckele L., Claessens M., Dinkla A., Brouwer C., Crijns W., Verellen D., et al. Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance. Radiother Oncol 2020;153:55–66. Doi: 10.1016/j.radonc.2020.09.008.
2. Temple, S. (2022). *An evaluation of AI auto-segmentation for Head & Neck cancer*. DClinSci. The University of Manchester.

Title of Study

Automated **C**linical **T**reatment Planning: from manual to auto-planning in Clinical Practise to reduce the patient pathway.

Authors: Anna Vella, Aoife Gallagher, Laura Stubbs, Harkirat Singh, Maxwell Robinson, Sriram Padmanaban
Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom

Background: Breast and Prostate are the most common treatment sites in radiotherapy, representing approximately two thirds of all patients receiving radiotherapy. Planning automation for these sites is fundamental to reducing the patient pathway, increasing conformity of treatment quality, and reducing treatment planning times [2][3].

Methods: Automated **C**linical **T**reatment Planning (ACT) was conceived as a rapid and efficient tool to streamline breast and prostate radiotherapy treatment planning at local institution. ACT was developed using an in-house Eclipse [1][4] Scripting Application Programming Interface (ESAPI) for inverse planning with IMRT and VMAT technique to automate dose optimization and efficiently produce high-quality treatment plans. Plans were generated starting from a simple protocol which consisted of the constraints for PTV targets and organs at risk (OAR) such as lungs and heart for breast. The performance of the automatic approaches was evaluated in terms of treatment planning time, target coverage, target dose heterogeneity, and OAR sparing.

Results: ACT-Breast was retrospective tested and assessed on 20 breast patients before starting its clinical use. Following a local audit of subsequent clinical use, the initial release was improved to support planning with newly installed TrueBeam Linacs and latest Varian calculation algorithm. ACT-Breast has drastically reduced total treatment planning times to approximately 10 minutes, with the actual ACT plan creation time ~ 1 mins, in comparison to approximately 45min for manual planning. ACT-Prostate is currently a prototype and will be tested and assessed similarly to ACT-Breast. The prototype supports automatic optimisation with RapidPlan models and DVH Estimation and creates an acceptable initial dose plan.

Discussion: ACT automatically generates clinically suitable radiotherapy plans in a time efficient manner. In challenging cases where ACT may produce clinically sub optimal plans, ACT offers a base for further improving plans in a second optimisation run i.e. combining automated and manual planning where appropriate to maximise clinical care for patients. ACT offers the potential to significantly reduce the patients' care path.

Conclusion: Clinical use of ACT-Breast creates the basis for further auto-planning development with the aim to achieve general timesaving, consistent and conformal dosimetry in planning. ACT-Prostate prototype will be further developed by extending auto-planning to other conformal VMAT sites such as simple pelvis (rectum, gynae, bladder) or more complex planning supported by RapidPlan Models such as Head&Neck.

Key Words: Auto-Planning, Breast, Prostate, VMAT, IMRT, Radiotherapy, Treatment Planning, Eclipse ESAPI Scripting, RapidPlan.

Key references:

1. Varian, Palo Alto, CA
2. K. Spencer et al., The Lancet Oncology 22, 2021.
3. B. V. Offersen et al., Radiotherapy and Oncology 114, 2015.
4. Joakim Pyry and Wayne Keranen, Varian APIs, 2018

Figure 1: ACT-Breast Graphic User Interface.

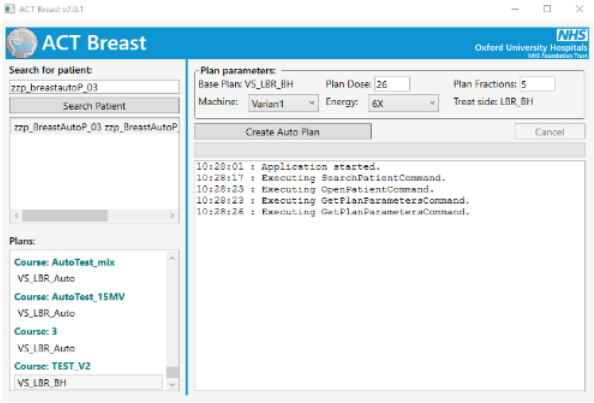


Figure 2: ACT-Breast Dosimetry Results – Clinical Goals. PTV_DVH coverage at D98%, Heart, Body, Heart and Lungs constraints for the same test patient are comparable for all the plans manual (1st column), first release of ACT-Breast (2nd column), second and current release of ACT-Breast with iX Clinacs AAA Calculation model, and with TrueBeam/Acuros (3rd and 4th colums) (all results within ~2%).

Overview

Clinical goals: All Plans — Evaluate Goals for All Plans

Plan			■ 1 / LBrea...	▲ 2 / LtBre...	● 3Test_iX /...	☼ 3Test_TB...
Total Dose			26.000 Gy	26.000 Gy	26.000 Gy	26.000 Gy
Clinical Goal Summary			0 1 9	0 1 9	0 1 9	0 1 9
PTV_DVH	P2	D 98.0 % > 95.0 %	96.20 %	95.31 %	95.21 %	95.04 %
	P2	V 105.0 % < 5.0 %	1.19 %	0.00 %	0.00 %	0.02 %
BODY	P3	D 0.0 % < 107.0 %	105.77 %	103.25 %	104.63 %	105.44 %
Heart	P1	V 1.50 Gy < 30.0 %	14.32 %	13.81 %	14.00 %	11.37 %
	P1	V 7.00 Gy < 5.0 %	0.03 %	0.10 %	0.07 %	0.07 %
	R	Dmean > 0.00 Gy	0.78 Gy	0.78 Gy	0.77 Gy	0.76 Gy
Left Lung	P1	V 8.00 Gy < 15.0 %	15.55 %	15.63 %	15.19 %	16.11 %
	R	Dmean > 0.00 Gy	4.18 Gy	4.10 Gy	4.02 Gy	3.94 Gy
Right Lung	P1	V 8.00 Gy < 15.0 %	0.00 %	0.00 %	0.00 %	0.00 %
	R	Dmean > 0.00 Gy	0.06 Gy	0.05 Gy	0.05 Gy	0.12 Gy

Title of Study: Implementing an automated treatment plan checking script

Submitters details: Ben Harris, Radiotherapy Physicist, Weston Park Cancer Centre.
Jonathan Hughes, Senior Radiotherapy Physicist, Weston Park Cancer Centre.

Background.

Independent checking of treatment plans by physics staff is time consuming and error prone [1]. It has been shown that by automating checks suitable for computer evaluation, the plan error rate and checking time can be reduced [2,3,4,5]. Based on these findings, we have implemented an automatic checking script to improve the efficiency of our planning and checking. We audited this process to identify further checks that can be automated, and to track our error rate over time.

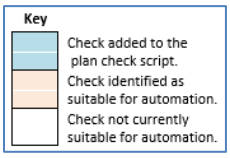
Methods.

An Eclipse script was developed, following best practices in software development. A 6-week audit of plan checking was carried out before the implementation of the script. A second audit was carried out a year later to identify further checks that could be automated. A software QA programme (including an automated self-test routine) was implemented to provide continuing confidence in the integrity of the script, and to maintain plan checker competency.

Results.

Table 1: Results from the 1st and 2nd audit.

	1st Audit			2nd Audit		
Plan Check	**Number of Errors**	**Average time lost / mins**	**Total time lost / mins**	**Number of Errors**	**Average time lost / mins**	**Total time lost / mins**
Plan quality and DVH stats	28	13.9	388.0	26	6.7	174.0
Correct dose rate and field times	21	6.5	136.7	0	0.0	0.0
Secondary dose calculation	21	6.0	126.0	15	7.6	114.0
Setting up distance & shifts	31	4.2	130.0	53	5.0	266.0
Appropriate linac booked	11	6.1	67.3	10	4.0	40.0
Reference point checks	10	6.8	68.0	8	4.9	39.0
Matching structures correct	8	11.8	94.7	11	6.6	73.0
Paperwork checks	38	8.8	334.3	35	5.2	183.0
Contours	18	24.0	432.3	28	36.7	1028.0
Field IDs	10	5.1	51.3	14	2.4	33.0
MLC jaws backed up with collimators	5	22.0	110.0	0	0.0	0.0
Plan normalisation method	5	5.8	29.0	0	0.0	0.0
Prescription checks	4	16.3	65.0	14	5.6	78
Image DICOM origin checks	4	11.3	45.0	0	0.0	0.0



Discussion.

The 1st audit demonstrated that there was significant time lost on checks well suited for automation. The script now catches these errors before the checking stage and so improves the efficiency of the process, saving an average of 1.1 minutes per plan. The 2nd audit revealed further errors that will be added to the next version of the script. Our software QA programme gives us confidence in the integrity of the script and has not identified any serious software errors.

Conclusion.

The plan check script has eliminated time lost in checking and resolving errors for several checks. This is an ongoing project which, coupled with regular plan-checking audit, aims to continuously improve our efficiency and reduce our error rate to improve patient safety. Automated plan checks offer significant prospects for resource saving and risk reduction, provided they are implemented according to best practices in software development and maintained and monitored with a rigorous QA programme.

Key references.

[1] Clouser E L, Chen Q, Rong Y. Computer automation for physics chart check should be adopted in clinic to replace manual chart checking for radiotherapy. *J App Clin Med Phys.* 2021; 22(2): 4-8.
[2] Liu S, et al. Optimizing efficiency and safety in external beam radiotherapy using automated plan check (APC) tool and six sigma methodology. *Rad onc Phys.* 2019; 20(8): 56-64.
[3] Dewhurst J M, et al. AutoLock_ a semiautomated system for radiotherapy treatment plan quality control. *J App Clin Med Phys*. 2015; 16(3).
[4] Covington E L, Popple R A, Cardan R A. Technical Note: Use of automation to eliminate shift errors. *J App Clin Med Phys.* 2020; 21(3): 192-195.
[5] Covington E L, et al. Improving treatment plan evaluation with automation. *J App Clin Med Phys.* 2016; 17(6): 16-31.

Automated Prostate Planning with ESAPI Scripting and RapidPlan

Gavin Orchin – Beatson West of Scotland Cancer Centre

Background

VMAT radiotherapy plans can be time consuming to create, regularly requiring several hours per treatment [1]. At the Beatson, approximately 1000 VMAT prostate treatments are planned each year and as such they take up a large portion of the departmental planning time. Since 2018, the Beatson has used a partially automated knowledge-based model (RapidPlan) to generate PTV and OAR dose objectives for prostate plans. However, it has been shown that scripting can further reduce the overall planning time while maintaining plan quality and reducing the rate of technical errors [2-5]. Therefore, we now aim to use ESAPI scripting to build upon pre-existing RapidPlan models and streamline the planning process further.

Methods

A plugin script has been developed that works from a CT scan with contoured GTVs and OARs and produces an external beam prostate plan, optimised using the approved RapidPlan model. The main tasks completed by the script are the following:

- Identify existing structures in the structure set
- Margin three PTVs from the prostate GTVs according to the CHHiP protocol
- Contour the gold fiducial markers within the prostate and assign them a density
- Add a treatment couch model
- Create a treatment plan in the correct course
- Select a suitable isocentre position
- Insert treatment fields and setup fields in a standard geometry
- Create a reference point at the centre of the high risk PTV
- Fit treatment field jaws to the PTVs
- Create DRRs
- Add dose estimates and optimisation objectives from the approved RapidPlan model
- Optimise the plan

Results

The script is currently being evaluated for department wide use and will soon be implemented clinically. The automatically performed contouring (PTV margining and high density segmentation) has been found to be highly consistent and near indistinguishable from current methods. Optimisation using a RapidPlan model allows the script to produce a clinically acceptable external beam plan in just a few minutes and the safety checks that the script performs are able to identify contouring and prescription errors at an early stage and should therefore reduce the probability of treatment delays.

Conclusion

The automated prostate planning script that has been developed is expected to create large time savings for the planning department and reduce the rate of repeat planning by preventing errors such as: violations of naming conventions, incorrect structure margining and incorrect structure assignments within RapidPlan.

Key references

1. Moore, K. L. (2019). Automated radiotherapy treatment planning. *Semin Radiat Oncol* 29, 209–18.
2. Gleeson, I., Bolger, N., Chun, H., et al. (2023). Implementation of automated personalised breast radiotherapy... *The British Journal of Radiology*, 95, 20220707.
3. Teruel, J. R., Malin, M., Liu, E. K., et al. (2020). Full automation of spinal stereotactic radiosurgery... *Journal of Applied Clinical Medical Physics*, 21(10), 122-131.
4. Yang, D., & Moore, K. L. (2012). Automated radiotherapy treatment plan integrity verification. *Medical physics*, 39(3), 1542-1551.
5. Yedekci, Y., Gültekin, M., Sari, S. Y., & Yildiz, F. (2023). Improving normal tissue sparing using scripting… *Radiation and environmental biophysics*, 1-8.

Automating the recalculation of clinical SABR treatment plans in an independent TPS to provide 3D dose evaluation at plan check

Aims and/or Background: Stereotactic ablative radiotherapy (SABR) treatments require accurate methods of independently verifying the treatment planning system (TPS) dose calculation. Often, simple dose check software does not adequately account for tissue inhomogeneities, resulting in inaccurate or unreliable verification. As a result, departments may resort to measurement-based patient-specific quality assurance (PSQA) of every patient to verify the TPS dose calculation. With an increasing number of patients receiving SABR treatments, this can place increased time and machine requirements on radiotherapy departments.

The aim of this work is to automate the recalculation of clinical SABR treatment plans on a second TPS in order to save time checking plans, reduce the requirements of PSQA measurements on the treatment machine, as well as providing a tool for evaluating a 3D dose distribution at physics check.

Methods: A 10FFF beam model was commissioned and verified in Raystation for checking Eclipse dose calculations. The model was verified against previous PSQA measurements and compared to Eclipse for 81 clinical SBRT treatment plans. A script was developed for Raystation that automatically recalculates the dose distribution for plans exported from Eclipse, and subsequently exports the DICOM Dose and Plan files for direct import into the Aria Database using the Varian DICOM Daemon. Once a clinically acceptable plan has been produced, the planner exports the plan using an export filter configured in Aria and the Raystation script automatically generates the check plan overnight. The recalculated check plan and its 3D dose distribution is then available within Aria for the plan checker the next day for comparison with the Eclipse clinical plan, as well as commercial independent dose check software.

Results: The mean ± standard deviation calculation error for Raystation point-dose PSQA plans was -0.9%±1.2% whilst for Eclipse it was 2.0%±2.3%. Similarly, the mean ± standard deviation PTV D95% (Gy) was -2.3Gy±1.0Gy for plans calculated using the Raystation model compared to Eclipse. Using the automation script reduces the time required to check a SBRT plan, and removes the repetitive tasks of importing, calculating, and exporting on Raystation. The check plan is available in Eclipse the following day, allowing a direct 3D dose comparison with Eclipse, whereas commercial independent dose check software often verifies a single point.

Discussion around results: The smaller PSQA calculation error using the checking (Raystation) beam model provides confidence in its use as an independent verification tool. The differences in the PTV D95% metric between Eclipse and Raystation can be used as a tolerance to help decide whether further PSQA is required. The automated recalculation of SABR plans using a second model provides a valuable resource for checking SABR plans, provides more information for the checker, including the ability to evaluate conformance to target and OAR constraints on the check plan, and reduces the time required to check. It is not dependent on machine time and thus reduces the burden of the physics team for PSQC.

Conclusion: 10FFF beam model was developed on a second TPS to act as independent dose check of SABR plans. Using a script to automate this verification check is an efficient way to verify the dose distribution and relieves some of the burden on the physics QC for machine time for PSQC.

Key Words:

- SABR, Eclipse, Raystation, Plan verification, Plan checking, independent dose check

An overview of treatment planning automation used for proton beam therapy at The Christie

Samuel Ingram[1,2], Matthew Clarke[1], Matthew Lowe[1,2], and The Christie PBT Physics Team[1].
[1]Christie Medical Physics and Engineering, The Christie NHS Foundation Trust, Manchester, UK.
[2]The University of Manchester, Manchester, UK.

Background: Proton beam therapy (PBT) is a specialised form of radiotherapy can offer that dosimetric benefits to a selection of patient sites. As PBT is a specialised form of radiotherapy and less widespread there are several features that the commercial treatment planning systems are not well equipped to provide yet. In this work, we will discuss how we've used a range of scripting solutions to account for these missing features along with other solutions to improve areas of technical contouring and plan checking.

Methods: Scripting work has been carried out using Varian's Eclipse Scripting Application Programming Interface (ESAPI) for v16.1 of the Eclipse Treatment Planning System. These scripts are written in C# and use a range of user interfaces, config files and higher-level input files to ensure widespread adoption across the whole Physics team. Our approach to scripting, when possible, is to design solution frameworks that are not dependent on programmers to expand allowing us to achieve our clinical aims through the efforts of the wider team. Thus, allowing the programmer time to be focused on the continuation of the development of new solutions and minimisation of scripting feature updates. In this overview we will discuss the following automation scripts: (1) Plan Assessment Forms – automated dosimetric (including worst-case values in robustness scenarios) extraction for a range of clinically agreed metrics; (2) Worst Case Scenario Plans – a voxel-wise 3D dose map of the maximum and minimum dose values across all robustness scenarios; (3) Contour Cook Book – a parser which allows simple user made scripts to be run to create technical volumes automatically; (4) Auto CSI – a tool to allow for the automation of technical structures, isocentre positioning and beams for Cranio-Spinal Irradiation (CSI) patients; (5) Plan Check Script – a tool for plan checkers to automatically collate the results of a range of standard plan checks.

Results: Each of the scripts mentioned have a clinical impact in improving our functionality and efficiency during treatment planning. Figure 1 outlines some of the key aspects of these scripts to how we have tried to maximise these impacts.

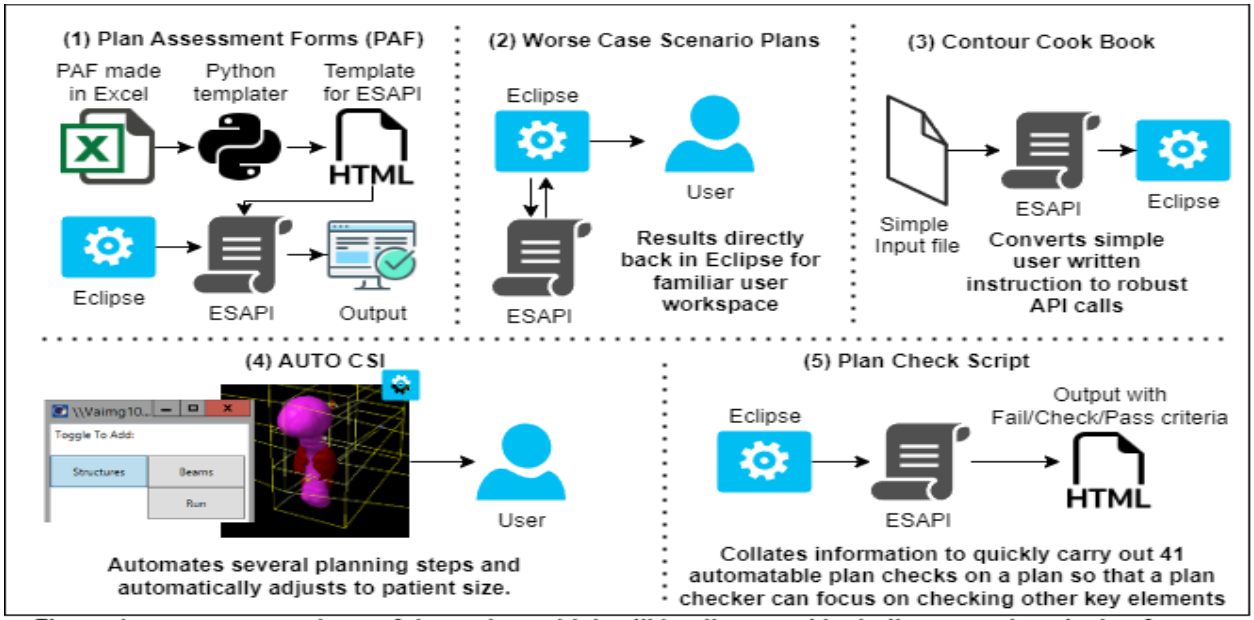


Figure 1 - process overviews of the scripts which will be discussed including some key design features used to improve accessibility and utility of each.

Discussion: We are continuing to develop scripts and utilise their potential clinically. We have worked focused on ensuring the sustainability of this work going forward as clinical pressures are likely to increase. Furthermore, to better understand the impact of these scripts we are aiming to introduce a range of collected metrics for each script which will help us to highlight their importance to both ourselves and the wider staff groups.

Conclusion: Scripting provides an invaluable tool for PBT treatment planning and has allowed us to compensate for missing features found in commercial treatment planning systems.

Comprehensive dosimetric evaluation of a CT scanner based deep learning auto-contouring solution for prostate radiotherapy

Berenato S.[1], Williams M.[1], Woodley O.[1], Möhler C.[2], Evans E.1, Millin A.E.[1], Wheeler P.A.[1]

[1]Velindre Cancer Centre, Cardiff, Wales, United Kingdom.
[2]Siemens Healthineers, Forchheim, Germany

Background: For extreme hypo-fractionated prostate radiotherapy, this study geometrically and dosimetrically evaluated DirectORGANS: a novel commercial AI solution that is natively integrated into a CT scanner and utilises dedicated reconstructions optimised and standardised for auto-contouring.

Methods: CT scans of 20 prostate patients were sequentially selected to evaluate AI contouring for rectum, bladder and proximal femurs. 5 plan generation 'pipelines' were considered. 3 used AI contours with differing levels of manual editing: nominally none (AI_{Std}), minor editing in specific regions e.g. target/OAR boundaries ($AIMin_{Ed}$), and fully corrected (AI_{FullEd}). The remaining 2 were manual delineations from different observers (MD_{Ob1},MD_{Ob2}). MD_{ob1} was defined as the reference contour set in all analysis. Contouring time was recorded and plans generated for each pipeline using a validated automated planning solution. The geometric and dosimetric agreement of contour sets AI_{Std}, $AIMin_{Ed}$, $AIFull_{Ed}$ and MD_{Ob2} were evaluated against the reference set MD_{Ob1}. The non-inferiority of the AI pipelines was assessed with the testing hypothesis that 'absolute deviations in geometry and dose metrics for AI contouring (vs MD_{Ob1}) were no greater than that from a second observer (MD_{Ob2})'. For dosimetric comparison the error in Reported Dose (RD) and Patient Dose (PD) was evaluated. RD was defined as DVH parameters that would be reported in patient records for a given pipeline.The dose distribution generated by each pipeline plan was evaluated on both the reference (RD_{Ref}) and pipeline ($RD_{Pipeline}$) contour sets, with the difference calculated to assess the impact of contour discrepancies on RD. PD, defined as the best estimate of the actual dose the patient would receive, was extracted from the pipeline plan's DVH using the reference contour set ($PD_{Pipeline}$). By comparing $PD_{Pipeline}$ with plans generated by and evaluated using MD_{Ob1} (PD_{Ref}), a contour set's influence on the optimisation process and hence final dose distribution, was assessed.

Results: Compared to MD_{Ob1}, overall delineation time for AI_{Std}, AI_{MinEd} and AI_{FullEd} was reduced by 24.9min (96%), 21.4min (79%) and 12.2min (45%) respectively. AI_{Std} contours exhibited good geometric alignment to MD_{Ob1} with median DSC results of 0.89, 0.95, 0.96 and 0.95 for rectum, bladder, femur_R and femur_L respectively. Minor editing led to marginal improvements but both AI_{Std} and AI_{MinEd} DSC results were statistically inferior to MD_{Ob2}. All pipelines exhibited generally good dosimetric agreement with MD_{Ob1}. For RD, median deviations were within ±1.8cm^3, ±1.7% and ±0.6Gy for absolute volume, relative volume and mean dose metrics respectively (Figure 1). For PD, agreement was improved with respective values within ±0.4cm^3, ±0.5% and 0.2Gy. Statistically AI_{MinEd} and AI_{FullEd} were dosimetrically non-inferior to MD_{Ob2}.

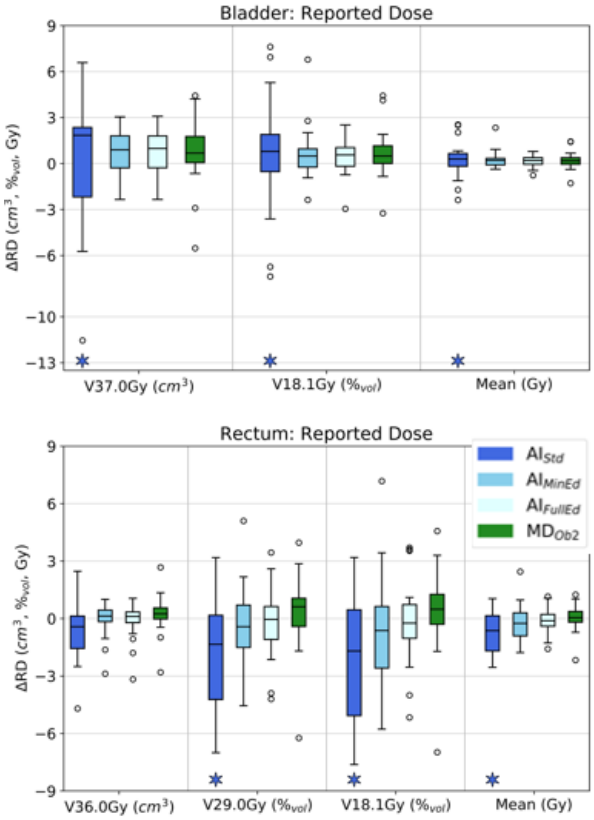


Figure 1: Results of the dosimetric assessment (RD) stars indicate statistical significant in terms of inferiority of the AI pipeline vs MD_{Ob2} ($p<0.05$)

Conclusion: Following minor editing (AI_{MinEd}), AI contours were dosimetrically non-inferior to manual delineations and reduced delineation time by 79%.

Assessing plan quality in the 'PLATO anal cancer trial 5' pilot phase with automated planning

Barrel M.J.[1], Abbott N.[1], Adams R.[1], Hawkins M.[2], Sebag-Montefiore D.[3], Millin A.[1], Wheeler P.A.[1]
1. Velindre Cancer Centre, Cardiff
2. University College of London, London
3. Leeds Cancer Centre, St James' University Hospital, Leeds

Background: Treatment efficacy relies on plan quality. Within trials, plan quality may vary due to training and equipment differences, which may influence treatment outcome or trial results. This study uses automated planning to assess plan quality and variation within the PersonaLlising rAdioTherapy dOse (PLATO) Anal Cancer Trial 5 (ACT5).

Methods: A protocol based automatic iterative optimisation (PBAIO) planning solution [1], implemented in RayStation, was calibrated for anal cancer using 5 pre-trial benchmark patient plans and 10 non-trial patients. Plans were generated for the pilot phase of PLATO ACT5; a dataset of 51 patients from 11 centres. Patients with prosthetic hips, replans, or unavailable suitable planning data were excluded (n=9). All trial plans were approved by the PLATO national trials QA team. The trial and automated plans were quantitatively compared using the ACT5 planning protocol parameters, small bowel V15Gy in cm^3, and planning target volume (PTV) conformity index (CI) and homogeneity index (HI). Statistical analysis was completed using a Wilcoxon signed rank test.

Results: At a population level, automation generally yielded higher quality plans with less variation when compared to trial plans. Automation reduced mandatory and optimal objective failures from 4 to 3 and 137 to 80 respectively.

34/46 metrics showed statistically significant ($p<0.05$) differences between automated and trial plans. Automation significantly reduced OAR dose (Table 1). Genitalia D50% and D35% reduced by >5.5Gy, femoral heads (FHs) by >2.5Gy and bladder D50% by 1.8Gy. Small bowel D200cc and D150cc reduced by 5.0Gy and V15Gy by $41cm^3$. These reductions did not adversely impact PTV D98%, D2%, HI or CI, which were within 0.6Gy, 0.6Gy, 0.018, and 0.017 respectively.

Structure	Metric	Objective		Trial		Automated	
		Man	Opt	Mean	StDev	Mean	StDev
Small Bowel	*D200cc (Gy)*	<35	<30	17.7	9.5	12.7	7.9
	D150cc (Gy)	<40	<35	20.3	10.0	15.2	9.2
	D20cc (Gy)	<50	<45	35.0	10.2	31.8	12.2
	D5cc (Gy)	<55	<50	40.4	7.9	38.6	10.2
	V15Gy (cc)	N/A	N/A	161.9	118.1	121.3	105.2
Left FH	*D50% (Gy)*	<45	<30	26.3	2.6	23.7	3.1
	D35% (Gy)	<50	<40	28.7	2.5	26.0	2.9
	D5% (Gy)	<55	<50	35.5	2.4	32.8	2.3
Right FH	*D50% (Gy)*	<45	<30	26.0	3.3	23.2	2.9
	D35% (Gy)	<50	<40	28.4	3.4	25.7	2.8
	D5% (Gy)	<55	<50	34.7	3.0	32.8	2.5
Genitalia	*D50% (Gy)*	<35	<20	23.2	5.3	17.3	3.3
	D35% (Gy)	<40	<30	27.1	5.5	21.4	5.5
	D5% (Gy)	<55	<40	42.6	9.0	39.8	11.3
Bladder	*D50% (Gy)*	<45	<35	32.6	5.1	30.7	5.4
	D35% (Gy)	<50	<40	37.1	4.7	36.2	5.2
	D5% (Gy)	<58	<50	48.2	5.3	48.4	5.5

Table1 - Trial and auto plan DVH data. *Italic* and underline indicate statistically significant differences. Man=Mandatory, Opt=Optimal PLATO ACT5 objectives.

At a per patient level, substantial variation in the difference between trial and automated plan metrics indicated noteworthy plan quality variability. For the genitalia and FHs, interquartile range (IQR) of the difference (trial-auto) was largest for D35%; 5.8Gy and 5.2Gy respectively. For the bladder, D50% IQR was 4.5Gy. The small bowel D200cc and V15Gy IQRs were 7.7Gy and $46cm^3$ respectively. Meaningful variations in PTV D98%, D2%, CI and HI were also observed with IQRs of up to 2.4Gy, 2.4Gy, 0.018, and 0.060 respectively.

Conclusion: Automated planning highlighted significant variations in plan quality within the pilot phase of PLATO ACT5. Evaluating plan quality in this manner may encourage improvements in training, QA and future trial approaches. This may reduce variation and improve overall plan quality.

Key references:

[1] P. Wheeler et.al, "Utilisation of Pareto navigation techniques to calibrate a fully automated radiotherapy treatment planning solution", *Phys Img Radiat Oncol,* vol. 16, no. 10, pp. 41-48, 2019

Development and Clinical Implementation of an Automated Radiotherapy Prostate Planning Script using the RayStation Scripting Interface

Authors: Richard Powis and Gareth Webster

Worcestershire Oncology Centre, Worcester Royal Hospital

Background:

VMAT prostate radiotherapy plan optimisation is dependent on the patient anatomy, the skill and experience of the planner and the time available. Scripts in Raystation TPS can be used to efficiently audit historic plan quality and have been employed locally as an effective tool to guide the manual VMAT plan optimisation process and reduce organ at risk doses (OAR) [1]. An in-house Raystation automatic planning script (AutoPlan) has been developed and implemented for prostate radiotherapy with a view to minimise manual input whilst producing high quality clinical plans.

Method:

AutoPlan fully automates the plan production process growing PTV(s) and plan optimisation structures, creating a VMAT arc and fully optimising the plan to produce a high quality dose distribution that meets all clinical goals for standard clinical cases. The script utilises an existing local knowledge-based planning (KBP) model [1] and an iterative plan optimisation process.

AutoPlan was implemented into an experienced team, following training and advice to consider manual intervention if worthwhile. Prostate plan quality was regularly audited using the RayStation scripting interface to monitor the performance of AutoPlan. Over subsequent audits it was noticed that planners were able to achieve modest improvements on the original knowledge base using the plan produced by the AutoPlan as a foundation. The knowledge base was re-baselined and incorporated into a second version of the script (AutoPlanV2) which was subsequently introduced into clinical use.

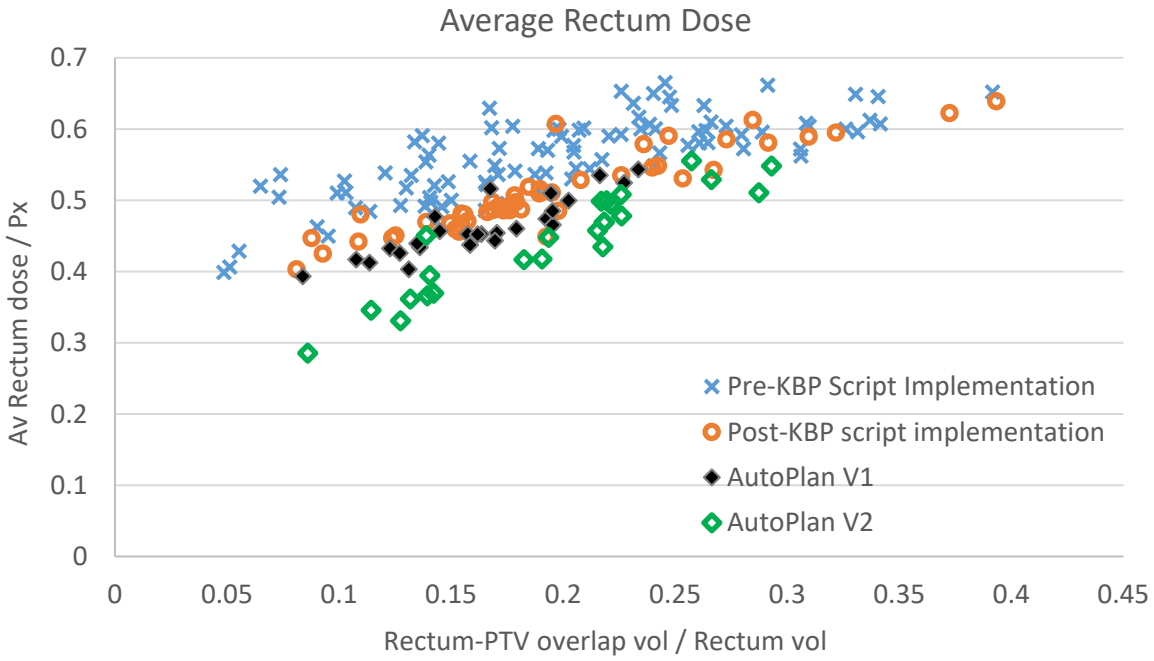
Results: Manual prostate plan optimisation guided by an existing local Raystation KBP script [1] (see hollow circles in figure below) has previously been shown to produce good quality plans with significantly lower average rectum doses compared to manual planning alone (see crosses in figure below). Introduction of AutoPlan was found to produce plans of comparable high quality with modest improvements to rectum average dose and minimal manual input (see solid diamonds in figure below). The introduction of AutoPlan2 was able to further improve cohort average rectum dose with minimum manual input (see hollow diamonds in figure below).

Conclusion:

An automated planning script has been developed and refined using the RayStation scripting interface to produce high quality clinical prostate plans with minimal user input.

Key References:

1] Clinical implementation of a knowledge based planning tool for prostate VMAT, Powis et al. Radiation Oncology (2017) 12:81


Average Rectum Dose
Av Rectum dose / Px
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
Rectum-PTV overlap vol / Rectum vol
Pre-KBP Script Implementation
Post-KBP script implementation
AutoPlan V1
AutoPlan V2

Automated Optimisation Structure Generation for Head and Neck Radiotherapy Planning

Henry Carver, Daniel Egleston, Russell Dawson, Simon Temple
The Clatterbridge Cancer Centre NHS Foundation Trust, Liverpool, United Kingdom

Abstract no more than 1 page in Arial 11 point, presenting speaker underlined

Invited talks - an abstract summarising you presentation is welcome including any images or tables.

Proffered papers - please follow the style below:

Background. Background to the study and aim of study including 5-10 key references.

Planning of complex radiotherapy treatments involves the generation of optimisation structures. These are grown from anatomical structures according to simple geometrical rules. The process of creating optimisation structures from planning structures is time consuming and prone to error, especially for complex sites such as radiotherapy to the head and neck.

Methods. Key methods used in the study including diagrams, images as necessary.

A C# script was developed leveraging the Eclipse Scripting API (ESAPI) to generate optimisation structures for inverse-optimised radiotherapy planning. This is achieved by pattern matching in the structure name to determine the type of structure. This matching method is robust for a range of treatment sites and structure names.
The script takes as input a structure set containing populated CTVs and OARs, it will then automatically populate any planning target volumes, hot structures, cold structures, opt structures and planning risk volumes in the structure set.

Efficiency will be measured using self-reported timing of structure creation by treatment planners. This has been done before and after script deployment for a set (N=50) of head and neck plans with a range of plan complexity. This will be supplemented by a retrospective audit of plan rejection rate following structure checking by an independent physicist. Feedback from beta testers has been recorded by questionnaire.

Results. Results of the study including diagrams, images, tables as necessary.

The results of the quality improvement audit will be presented at conference as this audit has not yet completed.

Initial feedback from beta testers has been very positive, citing time saving and efficiency as significant improvements.

Discussion. Discussion of the significance of the results

The script has the potential to improve the efficiency and quality of head and neck radiotherapy planning by automating the tedious and time-consuming task of optimisation structure generation. The script also reduces inter-planner variability and enhances standardisation of planning practices. The structure matching algorithm is flexible and robust for different sites and anatomies, and can be easily adapted for other regions of interest.

Conclusion. Conclusion relating to the aim of the study.
We have developed a C# script that automatically generates optimisation structures for head and neck radiotherapy planning using a novel structure matching algorithm.

The script has been well received by beta testers and has shown promising results in terms of planning time reduction and planner consistency. We aim to present audit results showing the impact of the script on clinical outcomes.

Key references. In alphabetical order, numbered.

Automation, ESAPI, optimisation, planning, radiotherapy, structure generation