European University Cyprus

INPICTURA

German Oncology Center

DETECTING AUTOMATION BIAS IN AI AUTOCONTOURING

P. Doolan[1], A. Karousiou[2], D Boukerroui[3], C Zamboglou[1,4], K Ferentinos[1,4] M J Gooding[3,5]

[1]German Oncology Center, Cyprus; [2]University of Surrey, Surrey; [3]Inpictura Ltd, Abingdon, UK; [4]European University Cyprus, Cyprus, [5]Division of Cancer Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, United Kingdom

Paul.Doolan@goc.com.cy

@PJDOOLAN

PURPOSE

- Autocontouring a critical part of radiotherapy pathway, **due to its consistency and time-saving**.[1]
- When the system is deployed in the clinical workflow, it is important the **system continues to be corrected** with the same level of attention as at commissioning and initial evaluation. **Automation bias**, where the user trusts the automated result more than their own judgement, is a real concern.
- Re-evaluation of test cases or periodic audits may not detect automation bias or review fatigue, so recent **RCR guidelines recommend ongoing monitoring**.[2] This pilot study tested AIQUALIS for detecting automation bias.

METHODS & MATERIALS

- Autocontouring is conducted in RayStation (RaySearch, Stockholm). As shown in Figure 1, auto-contours are sent to AIQUALIS (www.aiqualis.com) after creation ('automatic' contours) and after plan completion ('clinical' contours).
- **The difference between automatic and clinical contours shows how much editing is performed.**
- Tested **400+ patients** in 2023-2024. Anatomical sites: **Brain, H&N,** plus others.
- **Four different radiographers performed OAR editing.**
- Geometric similarity between autocontours and clinical contours was measured using the **surface Dice Similarity Coefficient** (sDSC)[3], with a **3D map** highlighting edits per structure.
- Intermediate results were presented to RTTs in July 2024 to try to reduce automation bias.

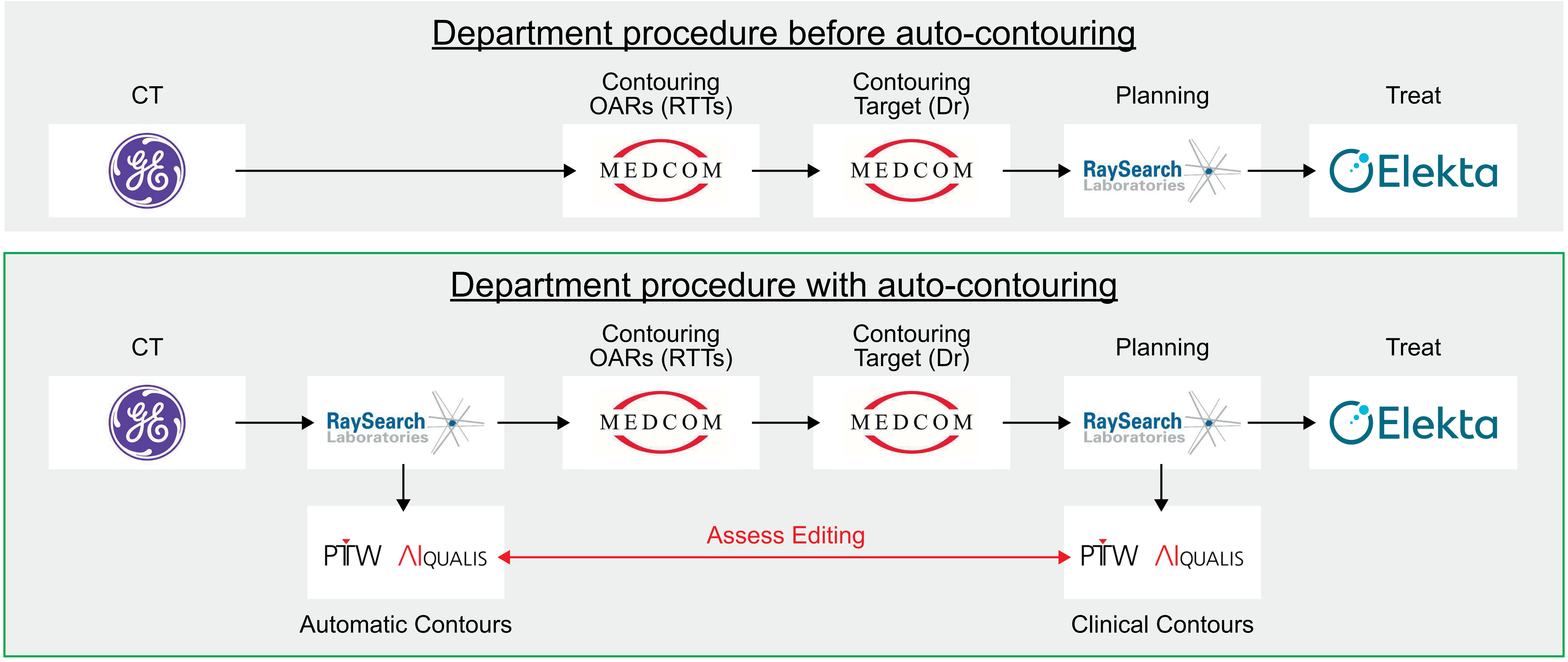

Figure 1. Radiotherapy pathway without (top, grey) and with (bottom, green border) auto-contouring.

RESULTS

- Compared to commissioning, higher sDSC values suggest **less editing during routine practice**. This is evident in 13/14 structures assessed (Table 1).
- Increasing sDSC values over time suggest automation bias. Figure 2 shows the brain, which demonstrates this behaviour. Compared to commissioning (green line), only RTT1 performs similar corrections. **RTTs 2, 3 and 4 correct the autocontours less than at commissioning.**
- In the 10 days after presenting findings, the average sDSC for RTT4 decreased from 0.95 to 0.66, suggesting reduced automation bias.

Structure	Commissioning sDSC	Monitoring sDSC	Difference sDSC
Brain	0.70	0.77	-0.07
Brainstem	0.51	0.77	-0.26
Bone_Mandible	0.79	0.95	-0.16
Eye_L	0.76	0.97	-0.21
Eye_R	0.76	0.97	-0.20
Lens_L	0.91	0.93	-0.02
Lens_R	0.91	0.95	-0.04
Lips	0.71	0.68	0.03
OpticNrv_L	0.77	0.80	-0.03
OptiNrv_R	0.71	0.80	-0.09
Pituitary	0.79	0.84	-0.05
Spinal Cord	0.58	0.90	-0.32
Kidney_L	0.35	0.97	-0.62
Kidney_R	0.35	0.95	-0.60

Table 1. Geometric similarity during commissioning and routine use. Negative difference values (orange/red) indicate less editing during routine use.

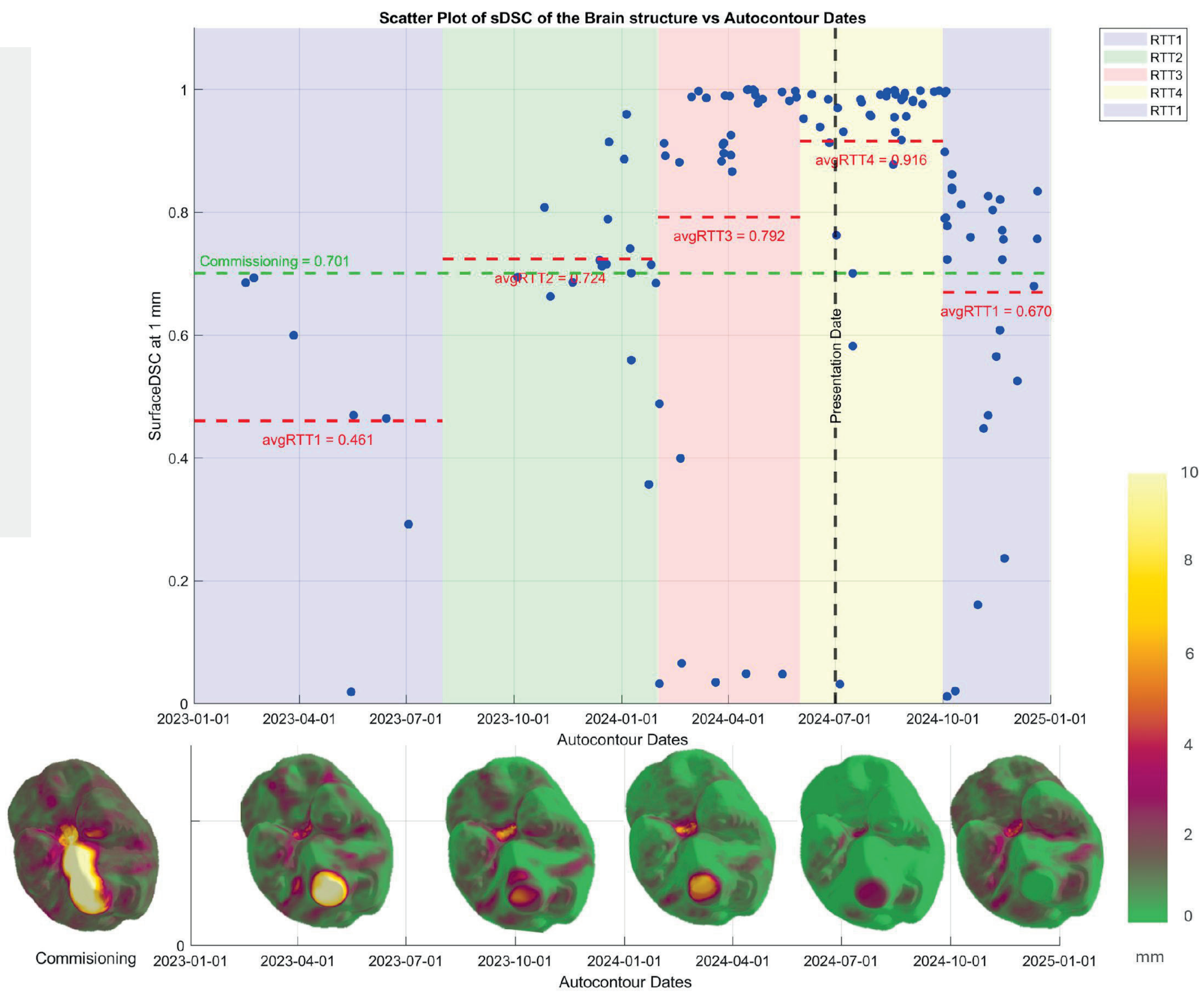

Figure 2. Similarity between autocontours and clinical contours for the brain. Commissioning data from a previous study[1] is shown by the green dotted line, while the mean disagreements for each radiation therapist is shown by the different red dotted lines. 3D renders highlighting the median location of edits for each radiation therapist is positioned at the bottom of each section of the graph, with the colour scale in mm on the right.

CONCLUSION

- Monitoring of autocontouring is crucial and can detect **differences in staff education** and **automation bias.**
- Ongoing surveillance can detect differences in autocontouring use between commissioning and routine, between **different staff members**, and for **different structures.**

References

- Doolan, P. J. et al. A clinical evaluation of the performance of five commercial artificial intelligence contouring systems for radiotherapy. Front. Oncol. 13, 1–13 (2023).
- Royal College of Radiologists. Clinical Oncology: Guidance on auto-contouring in radiotherapy. (2024).
- Nikolov, S. et al. Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy. 1–31 (2018).