Journal of Medical Radiation Sciences

Open Access

ORIGINAL ARTICLE

Improvement in male pelvis magnetic resonance image contouring following radiologist-delivered training

Bronwyn Oar, BSc, MRadTher, GCMagResonTech[1], Amy Brown, PhD, MAppSc, BAppSc (MRT)[1,2,3], Glen Newman, BAppSc (RT)[1], Alan Boles, MBBS, FRANZCR[4], Christopher N. Rumley, MBBS, FRANZCR[1,3], Rachel Doyle, BAppSc (RT)[1], John Baines, PhD, MSc, BSc[1,3], & Alex Tan, MBBS, FRANZCR[1,3]

[1]Townsville University Hospital, Townsville, Queensland, Australia
[2]Queensland University of Technology, Brisbane, Queensland, Australia
[3]James Cook University, Townsville, Queensland, Australia
[4]Queensland XRay, Townsville, Queensland, Australia

Keywords
magnetic resonance imaging, magnetic resonance linear accelerator, planning, radiation oncology

Correspondence
Amy Brown, Townsville Cancer Centre, Townsville University Hospital, PO Box 670, Townsville, QLD 4810, Australia. Tel: +61 7 44331664; Fax: +61 7 44331661; E-mail: amy.brown@health.qld.gov.au

Received: 6 December 2022; Accepted: 7 September 2023

J Med Radiat Sci **0** (2023) 1–8

doi: 10.1002/jmrs.727

Abstract

Introduction: The magnetic resonance linear accelerator (MRL) combines both magnetic resonance imaging and a linear accelerator, allowing for daily treatment adaptation. This study aimed to assess the impact of radiologist-delivered training in magnetic resonance (MR) contouring of relevant structures within the male pelvis. **Methods:** Two radiation oncologists, two radiation oncology registrars and seven radiation therapists completed contouring on 10 male pelvis MR datasets both pre- and post-training. A 2-hour MR anatomy training session was delivered by a radiologist, who also provided the 'gold standard' contours. The pre- and post-training contours were compared against the gold standard with Dice similarity coefficient (DSC) and Hausdorff distances calculated; and the pre- and post-confidence scores and timing were compared. **Results:** The improvement in DSC were significant in prostate, rectum and seminal vesicles, with a post-training median DSC of 0.87 ± 0.06, 0.92 ± 0.04 and 0.80 ± 0.14, respectively. The median Hausdorff improved with a median of 1.46 ± 0.78 mm, 0.52 ± 0.32 mm and 1.11 ± 0.86 mm for prostate, rectum and seminal vesicles, respectively. Bladder concordance was high both pre- and post-training. Urethra contours improved post-training, however, remained difficult to contour with a median post-DSC of 0.51 ± 0.24. Overall, confidence scoring improved ($P < 0.001$) and timing decreased by an average of 4.4 ± 16.4 min post-training. **Conclusion:** Radiologist-delivered training improved concordance of male pelvis contouring on MR datasets. Further work is required in the identification of urethra on MRs. These findings are of importance in the MRL adaptive workflow.

Introduction

Magnetic resonance imaging (MRI) is rapidly increasing in application in radiation oncology, particularly with the adoption of magnetic resonance (MR) simulators and MR linear accelerators (MRL). With the introduction of these MR technologies to radiation oncology departments, upskilling is required for the incorporation of MR images into daily clinical practice. Radiation oncologists (RO) and radiation therapists (RT) are well versed in contouring targets and organs at risk on computed tomography (CT) planning scans, and these base anatomical and technical skills can be transferred to the MR setting. However, in the MRL workflow, the daily adaptation of contours is completed online while the patient is on the treatment couch, adding time pressures particularly in the setting of prostate cancer where intrafraction motion can occur.[1–3]

In reporting their institutional failure mode and effect analysis for the implementation of online adaptive MRL

treatments, Klüter et al.[4] identified incorrect daily recontouring as one of the most critical risks in the MRL daily adaptive process. The implementation of daily adaptive MRL treatment requires careful consideration around these complex workflows, with daily replanning having a significant impact on workflow.[3] The MRL workflows can be guided in part by MR simulator implementations, such as that reported by Rai et al.,[5] which details the collaboration, training and upskilling required by radiation therapists, radiographers and medical physicists.

A recent study has suggested that inter-observer variability of prostate clinical target volumes (CTV) can be reduced when contouring with MRI fusion to the CT scan.[6] Male pelvic contours completed by ROs as well as medical physicists, radiographers and RTs were compared. Smaller contours, such as urethra, penile bulb and neurovascular bundles did show significant inter-observer variation. There was no statistical significance between RO and non-RO contours observed in this study.[6] Further, there was no training provided, only guidelines and a sample dataset for reference. The question as to whether training reduces inter-observer variability was not addressed. Consequently, our study aimed to assess the impact of radiologist-delivered training in MR contouring of relevant structures within the male pelvis.

Another United Kingdom study has highlighted the importance of training before a different profession takes on the contouring role from the RO. Smith et al.[7] showed that without training, radiographer prostate and SVs contours are 60–70% within clinical range of the gold standard. Therapeutic Radiographers (which are equivalent to Australian RTs) involved in contouring had 1–3 years' experience on the MR Linac, and therefore, it is assumed they would have some insight into MR pelvic anatomy, however, up to 40% of contours are outside of clinical threshold.[7] This highlights the importance of a training program directed at MR contouring of the anatomical site being treated.

There is a worldwide movement towards allocating contour adaptation to RTs in order to increase productivity and workflow for patients treated on the MR Linac. However, there is a strong agreement that specific training is required, and only trained and credentialled RTs should have the responsibility.[8,9]

Methods

This study was conducted prior to clinical implementation of our department's MRL, before the first patient was treated in December 2019. The MRL had been installed, with MRI training undertaken; however, this training focused on the acquisition of MRIs, rather than the incorporation of the MRI in the RT workflow. Prior to the implementation of the MRL, diagnostic MRIs were requested for fusion in the RT planning process at the treating RO's discretion. The fused diagnostic MRIs were used primarily by the RO in defining the target in conjunction with the CT, with both ROs and RTs not routinely having experience directly contouring utilising only the MRI.

An interdisciplinary participant group consisting of ROs radiation oncology registrars (Reg), and RTs completed contouring on 10 male pelvis MR datasets both pre- and post-radiologist training (Fig. 1). The MRI datasets from 10 sequential patients referred to our department for radiation therapy to the prostate between September and October 2019 were de-identified for use. Datasets 1–7 were from a 3T diagnostic MR scanner (Siemens Healthcare, Erlangen, Germany) and datasets 8–10 from the newly installed 1.5T Unity MRL (Elekta, Stockholm, Sweden). The diagnostic datasets were 1–3 mm slices, and the MRL datasets were 1 mm slices. Both T1 and T2 sequences were provided for reference; however, contouring was completed on the T2.

A single radiologist completed the contouring on all datasets, with this considered the 'gold standard', as opposed to the RO contours, as the ROs were not accustomed to direct MR contouring, with standard practice at our centre requiring contouring on the planning CT scan, with MRI, when available, as a secondary dataset to complement CT contouring. Additionally, due to the more limited MR scanner access within our regional and rural catchment, not all patients with prostate cancer will undergo a diagnostic MRI as part of their diagnosis workup. The radiologist was an MR specialist with a genitourinary special interest and over 30 years of clinical experience. The MR anatomy tutorial was delivered by the same radiologist over 2 h to the group of participants. The tutorial focussed solely on male MR pelvic anatomy, as this was the planned 'go-live' clinical site and was provided in lecture-style presentation, with the opportunity for participants to ask questions. The presentation highlighted anatomy on separate MRI datasets that were not related to the study datasets. The post-training contours completed at least 2 weeks following this tutorial, on the same datasets as the pre-training. Previous and other participants contours were not available to individual participants.

During both pre- and post-tutorial contouring, the participants assessed image quality of each dataset through a Visual Grading Assessment (VGA) with participants scoring of 1: Excellent image quality – no limitations for clinical use to 5: Poor image quality – image not usable, loss of information, image must be

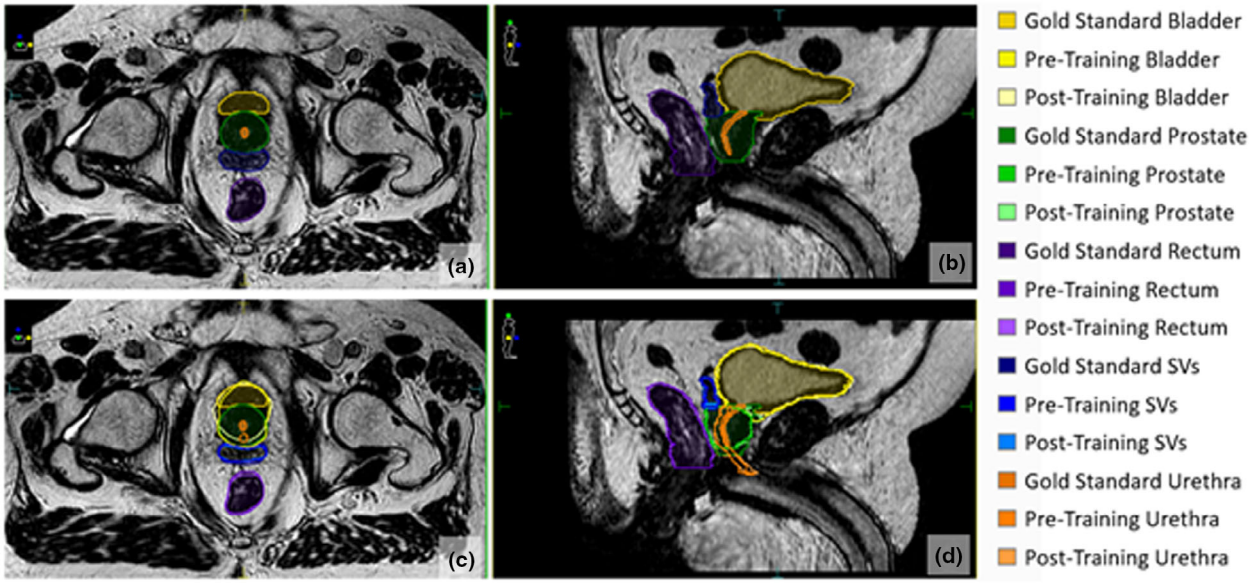

Figure 1. Example of contouring completed on an MRL-acquired scan, with (a and b) demonstrating gold standard contouring on transverse and sagittal planes, respectively, and (c and d) demonstrating one participant's pre- and post-training contours against the gold standard contours, with variance evident.

repeated,[10] and confidence in contouring was scored for each structure of each dataset with an investigator-developed Likert-type scale with 1: Not confident at all, to 5: Very confident. Post-training scoring was completed without reference to pre-training scoring. Participants also self-reported the overall time taken to contour all structures on each data set, both pre- and post-training. Where scan limits included the whole pelvis, only the clinically relevant length of the rectum (i.e. rectosigmoid junction superiorly (approximately S2–S3) to bottom of ischial tuberosities) was analysed. All contouring was completed in the Monaco treatment planning system (Elekta).

Institutional ethics approval was granted (HREC/2019/QTHS/56853), including a waiver of consent for the use of de-identified patient images. All staff participants provided written informed consent.

Analysis

Descriptive statistics as well as summarised measures including VGA score and time to complete contours were determined. Chi-square test compared the scores between pre- and post-training and Wilcoxon rank sums tests for timing differences (with paired data).

Metrics including Dice similarity coefficient (DSC), average Hausdorff distance (AVG_HAUS) and 95th percentile Hausdorff distance (95_HAUS) were calculated in a pairwise assessment, comparing the pre-training to the gold standard structures and the post-training to the gold standard structures. A DSC $\geq$0.7 was considered a good overlap.[11] DSC and Hausdorff distances were calculated in SlicerRT.[12]

Wilcoxon rank sums tests were performed to evaluate differences between structures contoured pre- and post-training ($P < 0.05$ significance level). All analysis was undertaken in R statistical software version 3.6.1.[13]

Results

Ten participants completed the contouring and training. One participant completed the pre-training contours but was unable to attend the training, and instead referred to the training materials and then completed the post-training contours. Two ROs (experience of 14–17 years), two radiation oncology registrars (experience of 1–2 years) and 7 RTs (experience of 2–35 years) completed the training.

Overall, the VGA score generally increased (however not significantly, $P = 0.32$) and timing decreased ($P = 0.01$) when comparing pre- and post- contouring (Table 1) Confidence in the contouring of individual structures also increased across the professional discipline groups (Table 2). There was a general trend for lower quality VGA and decreased confidence in the 1.5T datasets (Table 2); however, due to the smaller sample size (3 datasets), this was not stratified further for analysis.

Table 3 and Figure 2 summarises the DSC measures for each structure pre- and post-training, with an overall

Table 1. Visual Grading Assessment (VGA) scores and timing for each data set pre- and post-training (mean ± SD).

Dataset	MR field strength	VGA score pre† Mean ± SD	VGA score post† Mean ± SD	Timing pre (min) Mean ± SD	Timing post (min) Mean ± SD	Time Difference (min) Post–Pre Mean ± SD
1	3T	2.1 ± 1.0	2.5 ± 1.0	18.2 ± 8.2	18.5 ± 11.6	−0.2 ± 10.5
2	3T	2.3 ± 0.9	2.2 ± 1.1	15.2 ± 9.2	13.5 ± 7.0	−2.2 ± 9.3
3	3T	1.6 ± 0.8	1.8 ± 1.3	28.3 ± 23.0	19.7 ± 9.0	−7.6 ± 24.8
4	3T	2.5 ± 0.7	2.1 ± 1.0	21.9 ± 9.6	20.8 ± 9.3	−1.7 ± 10.3
5	3T	2.5 ± 0.7	2.5 ± 0.8	23.2 ± 14.0	19.4 ± 7.7	−3.7 ± 17.7
6	3T	2.1 ± 0.7	2.6 ± 1.4	23.4 ± 11.3	17.5 ± 7.3	−5.5 ± 10.3
7	1.5T	3.5 ± 0.9	2.5 ± 0.5	48.5 ± 25.7	38.9 ± 14.5	−8.5 ± 24.0
8	1.5T	2.6 ± 0.9	2.5 ± 1.2	47.1 ± 19.5	37.0 ± 23.6	−23.5 ± 19.8
9	1.5T	2.7 ± 0.9	2.4 ± 1.0	38.7 ± 7.3	50.8 ± 18.5	8.0 ± 13.2
10	3T	2.6 ± 1.1	2.2 ± 1.0	22.9 ± 11.7	16.9 ± 11.1	−6.3 ± 14.3
Total Average	–	2.4 ± 1.0	2.3 ± 1.0	27.9 ± 18.0	24.2 ± 16.4	−4.4 ± 16.4

†VGA: 1: Excellent image quality: no limitations for clinical use; 2: Good image quality: minimal limitations for clinical use; 3: Sufficient image quality: moderate limitations for clinical use but no substantial loss of information; 4: Restricted image quality: relevant limitations for clinical use, clear loss of information; 5: Poor image quality: image not usable, loss of information, image must be repeated.

Table 2. Confidence in contouring each structure pre- and post-training (mean, range).

	Overall			Radiation oncologist/Registrar			Radiation therapist		
Structure	Pre-confidence score mean ± SD	Post-confidence score mean ± SD	*P*-value	Pre-confidence score mean ± SD	Post-confidence score mean ± SD	*P*-value	Pre-confidence score mean ± SD	Post-confidence score mean ± SD	*P*-value
Prostate	2.7 ± 0.9	3.2 ± 0.8	**<0.001**	3.3 ± 0.8	3.9 ± 0.8	**0.009**	2.3 ± 0.7	2.9 ± 0.7	**<0.001**
Bladder	3.6 ± 1.0	4.2 ± 0.8	**<0.001**	4.2 ± 0.9	4.3 ± 0.8	0.404	3.3 ± 0.9	4.2 ± 0.9	**<0.001**
Rectum	3.3 ± 0.9	3.9 ± 0.9	**<0.001**	3.8 ± 0.8	4.3 ± 0.8	**0.017**	2.9 ± 0.8	3.7 ± 0.9	**<0.001**
SVs	2.8 ± 1.2	3.5 ± 0.9	**<0.001**	4.0 ± 0.9	3.9 ± 0.8	0.423	2.2 ± 0.8	3.2 ± 0.9	**<0.001**
Urethra	1.9 ± 1.0	2.3 ± 0.9	**<0.001**	2.5 ± 1.2	2.6 ± 1.1	0.807	1.6 ± 0.7	2.1 ± 0.8	**0.001**

Confidence scores: 1, Not at all confident; 5, Very confident; SD, standard deviation; SVs, seminal vesicles. Bold indicates statistical significance of $p < 0.05$.

significant improvement noted in prostate, rectum and seminal vesicle contours ($P < 0.05$). While not significant, improvements were also noted in the bladder and urethra. When considering a DSC ≥ 0.7 as a good overlap, all structures except the urethra achieved this, regardless of professional discipline. As Figure 3 demonstrates, there was considerable variation in urethra DSC across the datasets, and between pre- and post-training, indicating variation in visibility.

Similarly, there was an overall improvement in average Hausdorff distances of prostate, rectum and seminal vesicle structures (Table 4). There was a significant improvement in prostate and rectum Hausdorff distances for both RTs and RO/Regs, and a significant improvement in seminal vesicles by RTs.

Discussion

This study found that the Radiologist-delivered MR training increased confidence overall, decreased timing and improved both DSC and Hausdorff distances in contouring the prostate, bladder, rectum and seminal vesicles. This was particularly evident in the RT participants. It should be noted that it was departmental protocol in the standard planning process prior to MRL implementation that the RTs contour the bladder and rectum, and the ROs contour prostate and seminal vesicles, in addition to verifying the RT volumes in a standard pelvis case – however, this is generally on CT, with fused MRI as a secondary dataset if available. Urethras were not routinely contoured. In the online adaptive MRL workflow, the OAR contours are adapted from the planning reference scans by the treating radiation therapists and checked by the treating RO, with the RO adjusting the target contour as required.

The general trend noted of decreased VGA quality and confidence scores for the three datasets acquired on the 1.5T MRL makes clinical sense, as we would expect 3T MRs to be of higher quality. While the same datasets were being scored, we suggest that the increase in both the VGA scores and confidence were due to the training increasing the participant's knowledge of the MR-specifics

Table 3. Dice similarity coefficient median pre- and post-training for each structure, compared to the Radiologist reference contour.

	Overall			Radiation therapist			Radiation oncologist/Registrar		
Structure	Pre-DSC median ± SD (IQR)	Post-DSC median ± SD (IQR)	*P*-value	Pre-DSC median ± SD (IQR)	Post-DSC median ± SD (IQR)	*P*-value	Pre-DSC median ± SD (IQR)	Post-DSC median ± SD (IQR)	*P*-value
Prostate	0.84 ± 0.06 (0.80–0.89)	0.87 ± 0.06 (0.83–0.91)	**<0.001**	0.82 ± 0.06 (0.78–0.86)	0.85 ± 0.06 (0.82–0.89)	**<0.001**	0.88 ± 0.04 (0.85–0.91)	0.90 ± 0.05 (0.87–0.92)	**0.004**
Bladder	0.92 ± 0.07 (0.92–0.96)	0.93 ± 0.05 (0.91–0.96)	0.209	0.91 ± 0.08 (0.90–0.95)	0.92 ± 0.05 (0.91–0.95)	0.811	0.94 ± 0.03 (0.93–0.96)	0.94 ± 0.03 (0.94–0.96)	0.647
Rectum	0.90 ± 0.05 (0.88–0.94)	0.92 ± 0.04 (0.91–0.95)	**<0.001**	0.90 ± 0.05 (0.89–0.93)	0.91 ± 0.04 (0.90–0.94)	**0.007**	0.90 ± 0.05 (0.87–0.94)	0.94 ± 0.03 (0.92–0.95)	**<0.001**
SVs	0.78 ± 0.12 (0.72–0.87)	0.80 ± 0.14 (0.74–0.89)	**0.026**	0.77 ± 0.13 (0.73–0.87)	0.79 ± 0.13 (0.75–0.89)	**<0.001**	0.79 ± 0.12 (0.72–0.88)	0.80 ± 0.17 (0.73–0.90)	0.376
Urethra	0.50 ± 0.24 (0.34–0.87)	0.51 ± 0.24 (0.36–0.69)	0.070	0.46 ± 0.26 (0.30–0.66)	0.47 ± 0.25 (0.30–0.68)	0.293	0.58 ± 0.17 (0.43–0.70)	0.57 ± 0.21 (0.48–0.75)	0.658

DSC, dice similarity coefficient; IQR, inter quartile range; SD, standard deviation; SVs, seminal vesicles. Bold indicates statistical significance of p <0.05.

and, therefore, scored more positively. The impact of poorer quality MRIs in radiation oncology contouring has been reported, particularly in the context of specified radiation oncology set-ups including immobilisation, which can create a challenge for optimal coil placement and is an area for further exploration in the MRL setting.[5,14]

Inter-observer variability in radiation oncology contouring is well documented, with training/education initiatives aimed at reducing this. Khoo et al.[15] reported on the reduction of both inter- and intra-observer contour variation in prostate contouring following an education program. A radiologist-led workshop in contouring of lung, breast and cervix datasets noted some improvement.[16] Following one training session with a clinical oncologist, good agreement was found with RT prostate contours against a gold standard on CT, T2W and T2*W datasets.[17]

The increased soft tissue definition on MRIs compared to CTs of the male pelvis is also well documented, particularly in delineating the base and apex of the prostate.[18–23] Roach et al.[6] found excellent agreement amongst 13 clinicians undertaking contouring of 5 prostate cancer datasets with CT and MRI, with the aim of quantifying inter-observer variation. Their prostate DSC of 0.83 ± 0.05 was comparable to this study's finding of 0.84 ± 0.06 prior to training. Likewise, in evaluating the prostate contour adaptation performed by RTs in an MRL online setting, excellent DSCs for the prostate of 0.99–1.00 were reported when compared to the RTs contours against ROs.[24]

The prostate contour adaptations by the RTs were acceptable for clinical use in 94.2% of MRL fractions, with a mean contour adaption time of 12.6 min ± 3.8, which included adapting the CTV and OARs within a 2 cm ring around the CTV.[24] Conversely, our timing of 24.23 ± 16.4 min post-training included prostate and OARs in their entirety, except for the rectum for which anatomical limits were set (noting that some participants contoured beyond these limits, which would have also increased contouring time as reflected in the standard deviation ranges). It is recognised that in this offline setting, there was no deforming of previous contours, which would increase contouring efficiencies in the online setting. Willigenburg et al.[24] noted that online adaptation times decreased in latter datasets, indicating a potential learning effect. We report an overall mean decrease in time pre- to post-training of 4.4 ± 16.4 min and hypothesise that with increased clinical experience, contouring efficiencies will continue to improve.

Similar agreement to this study is also reported in bladder, rectum and seminal vesicle DSC of 0.88 ± 0.05,

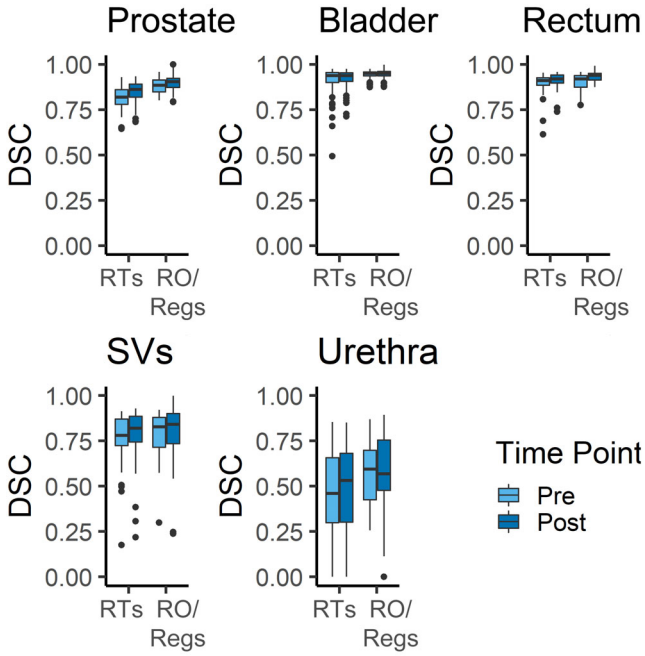

Figure 2. Dice similarity coefficient mean pre–post-training for each structure, by professional discipline. SVs, Seminal Vesicles.

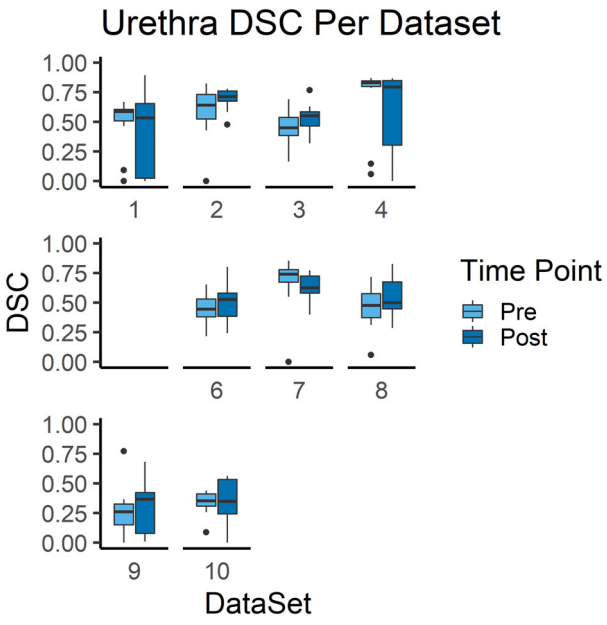

Figure 3. Boxplot of Urethra Dice similarity coefficient, pre- and post-training per dataset. No radiologist gold standard urethra contour was completed for Dataset 5.

0.93 ± 0.03, 0.73 ± 0.14, respectively.[6] As bladder and rectum are potential OARs for pelvic nodal SABR treatments, the good agreement of these structures is reassuring in other pelvic clinical sites than prostate alone treatments on the MRL.

It was found both in this study and in subsequent clinical experience, the urethra can be difficult to visualise with the standard MRL 3D T2 TSE Axial sequence used most frequently on the MRL. This is reflected in both the low DSCs, and low confidence scoring both pre- and post-training, although an improvement in both was noted in the post-training. Our results are in keeping with Blitzer et al.[25] who report a DSC average of 0.303 (range 0 to 0.704) when investigating whether an MRI voiding scan (that is, the patient is instructed to empty their bladder during MR acquisition) could improve urethra identification. Roach et al.[6] similarly reported a lower DSC of 0.41 ± 0.21 when contoured urethra on CT with MR fusion. Richardson et al.[26] found an improvement in urethra contouring when a 3D T2 SPACE MRI was utilised for delineation, with a mean DSC of 0.78 compared to 0.62 on T2 axial MRI scans. Currently, the MRL does not support such sequences within the standard online workflow. Based on our results and those within the literature, utilising a urethra contour in the MRL planning and adaptive process would require more investigation, including further optimising of scan parameters to improve accuracy of contouring.

There are some limitations to this study. The use of the same MRI datasets may have influenced results as participants had contoured on them previously when completing the post-workshop contours. This, however, was minimised by firstly ensuring that at least 2 weeks had elapsed to minimise any recall bias, and secondly,

Table 4. Hausdorff distance (in mm) pre- and post-training for each structure, compared to the Radiologist reference structure.

	Overall			Radiation therapist			Radiation oncologist/Registrar		
Structure	Pre-median ± SD (IQR)	Post-median ± SD (IQR)	*P*-value	Pre-median ± SD (IQR)	Post-median ± SD (IQR)	*P*-value	Pre-median ± SD (IQR)	Post-median ± SD (IQR)	*P*-value
Prostate	1.76 ± 0.83 (1.22–2.26)	1.46 ± 0.78 (1.00–1.88)	**<0.001**	2.06 ± 0.84 (1.54–2.53)	1.70 ± 0.78 (1.15–2.13)	**<0.001**	1.24 ± 0.49 (0.88–1.57)	1.05 ± 0.61 (0.70–1.37)	**0.019**
Bladder	0.99 ± 1.07 (0.55–1.01)	0.83 ± 0.60 (0.50–0.98)	0.422	1.18 ± 1.29 (0.59–1.09)	0.93 ± 0.69 (0.55–1.10)	0.348	0.66 ± 0.30 (0.46–0.83)	0.65 ± 0.35 (0.32–0.88)	0.948
Rectum	0.64 ± 0.42 (0.39–0.76)	0.52 ± 0.32 (0.29–0.68)	**<0.001**	0.66 ± 0.45 (0.37–0.80)	0.58 ± 0.35 (0.36–0.69)	**0.033**	0.61 ± 0.36 (0.39–0.72)	0.42 ± 0.24 (0.25–0.61)	**<0.001**
SVs	1.17 ± 0.63 (0.74–1.51)	1.11 ± 0.86 (0.62–1.35)	**0.018**	1.19 ± 0.61 (0.80–1.47)	1.09 ± 0.80 (0.69–1.30)	**0.003**	1.12 ± 0.68 (0.64–1.52)	1.15 ± 0.96 (0.57–1.49)	0.980
Urethra	1.95 ± 2.31 (0.62–2.42)	1.90 ± 2.08 (0.67–2.41)	0.887	2.38 ± 2.69 (0.75–2.75)	2.30 ± 2.47 (0.72–2.98)	0.919	1.15 ± 0.88 (0.58–1.55)	1.20 ± 0.78 (0.56–1.75)	0.902

IQR, inter quartile range; SD, standard deviation; SVs, seminal vesicles. Bold indicates statistical significance of p <0.05.

labelling the datasets differently, which was blinded to the participants (i.e. Patient 1 randomly assigned another number between 2 and 10). Additionally, while it would have been ideal to have multiple radiologists complete the contouring for the gold standard to create a Simultaneous truth and performance level estimation (STAPLE) contour, this was beyond the resourcing of this study and is recommended for future studies. Only three MRL-acquired datasets were available at the time of the study. In addition, three participants (1 RO, 1 Reg and 1 RT) completed their post-training contours >1 month after the training, which may have influenced results. However, particularly for RTs, this could reflect clinical practice where rostered rotations throughout the department may mean time away from the MRL. The grouping of the RO/Reg disciplines for analysis may have also impacted the results, due to the wide range of clinical experience; however, this was due to the small sample size of both groups.

This retrospective work is of importance in the adaptive workflow setting, as the increased time of staff in the MRL adaptation is well recognised.[24] RT-led processes are becoming increasingly implemented worldwide, particularly in the MRL setting.[8,9] This study validates that with training, the RTs can confidently contour most pelvic structures with concordance to gold standard contours. Additionally, it is recognised that in a clinical online setting, the MRL staff would have reference to the original planned contours, with these contours then undergoing contour deformation to the daily scan, which may aide in defining anatomical borders. This study supports the first step in the MRL workflow, which would then aid in determining the most appropriate online adaptive workflow.

Studies from the United Kingdom have made recommendations for not only a training program but also patient-specific instructions to be able to achieve clinical acceptable contours. Smith et al.[27] recommend a DSC of >0.9 for the prostate and SVs post-patient-specific contouring training. The therapeutic radiographer's prostate and SV contours were comparable to clinician contours during online workflow where there is a reference contour to start with.[27] This resulted in the production of clinically acceptable daily adapted radiation therapy plans. It should be noted that Australian-trained radiation therapists undergo a different level of training and are responsible for the optimisation of the radiation therapy plan and in most centres also contribute to OAR contouring. Adequate monitoring and auditing are required to maintain a high clinical standard; however, this study provides evidence for future implementation of RT-led daily adaption radiation therapy in Australia with consideration to appropriate training and credentialing.

Conclusion

Radiologist-delivered training improved concordance of contouring on male pelvis MR datasets amongst radiation oncologists, radiation oncology registrars and particularly radiation therapists. Further work is required in the identification of urethra on MRIs, and further validation of the training on a larger MRI dataset sample, ideally from the MRL is warranted. These findings are of importance in the MRL adaptive workflow, and a vital consideration in the implementation of MRL daily adaptation of differing clinical sites.

Acknowledgements

The authors thank the participants in this study and Scott Cooper and Georgia Barjaktarovic for their contributions to the study development. It is acknowledged that both Christopher N Rumley and Alex Tan were affliated with Townsville University Hospital at the time of the study and at first mansucript drafting.

Conflict of Interest

Nothing to declare.

Funding Statement

This project was funded by the Townsville Hospital and Health Service Study, Education and Research Trust Account.

Ethics Statement

Ethical approval was granted by the Townsville Hospital and Health Service Human Ethics Committee (HREC/2019/QTHS/56853) and all participants provided written informed consent.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request, and subject to ethical approval.

References

1. Tocco BR, Kishan AU, Ma TM, Kerkmeijer LGW, Tree AC. MR-guided radiotherapy for prostate cancer. *Front Oncol* 2020; **10**: 1–11.
2. Pathmanathan AU, van As NJ, Kerkmeijer LGW, et al. MRI-guided adaptive radiotherapy; a "game changer" for prostate radiotherapy. *Int J Radiat Oncol Biol Phys* 2017; **100**: 361–73.
3. Murray J, Tree AC. Prostate cancer – Advantages and disadvantages of MR-guided RT. *Clin Transl Radiat Oncol* 2019; **18**: 68–73.
4. Klüter S, Schrenk O, Renkamp CK, et al. A practical implementation of risk management for the clinical introduction of online adaptive magnetic resonance-guided radiotherapy. *Phys Imaging Radiat Oncol* 2021; **17**: 53–7.
5. Rai R, Kumar S, Batumalai V, et al. The integration of MRI in radiation therapy: collaboration of radiographers and radiation therapists. *J Med Radiat Sci* 2017; **64**: 61–8.
6. Roach D, Holloway LC, Jameson MG, et al. Multi-observer contouring of male pelvic anatomy: Highly variable agreement across conventional and emerging structures of interest. *J Med Imaging Radiat Oncol* 2019; **63**: 264–71.
7. Adair Smith G, Dunlop A, Alexander SE, et al. Interobserver variation of clinical oncologists compared to therapeutic radiographers (RTT) prostate contours on T2 weighted MRI. *Tech Innov Patient Support Radiat Oncol* 2023; **25**: 100200.
8. McNair HA, Joyce E, O'Gara G, et al. Radiographer-led online image guided adaptive radiotherapy: A qualitative investigation of the therapeutic radiographer role. *Radiography* 2021; **27**: 1085–93.
9. Hales RB, Rodgers J, Whiteside L, et al. Therapeutic radiographers at the helm: Moving towards radiographer-led MR-guided radiotherapy. *J Med Imaging Radiat Sci* 2020; **51**: 364–72.
10. Ludewig E, Richter A, Frame M. Diagnostic imaging – Evaluating image quality using visual grading characteristic (VGC) analysis. *Vet Res Commun* 2010; **34**: 473–9.
11. Vinod SK, Jameson MG, Min M, Holloway LC. Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies. *Radiother Oncol* 2016; **121**: 169–79.
12. Pinter C, Lasso A, Wang A, Jaffray D, Fichtinger G. SlicerRT: Radiation therapy research toolkit for 3D Slicer. *Med Phys* 2012; **39**: 6332–8.
13. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2019.
14. Zukauskaite R, Rumley CN, Hansen CR, et al. Delineation uncertainties of tumour volumes on MRI of head and neck cancer patients. *Clin Transl Radiat Oncol* 2022; **36**: 121–6.
15. Khoo ELH, Schick K, Plank AW, et al. Prostate contouring variation: Can it be fixed? *Int J Radiat Oncol Biol Phys* 2012; **82**: 1923–9.
16. Kumar S, Holloway L, Roach D, et al. The impact of a radiologist-led workshop on MRI target volume delineation for radiotherapy. *J Med Radiat Sci* 2018; **65**: 300–10.

17. Pathmanathan AU, Mcnair HA, Schmidt MA, et al. Comparison of prostate delineation on multimodality imaging for MR-guided radiotherapy. *Br J Radiol* 2019; **92**: 20180948.
18. Rasch C, Barillot I, Remeijer P, Touw A, Van Herk M, Lebesque JV. Definition of the prostate in CT and MRI: A multi-observer study. *Int J Radiat Oncol Biol Phys* 1999; **43**: 57–66.
19. Wachter S, Wachter-Gerstner N, Bock T, et al. Interobserver comparison of CT and MRI-based prostate apex definition. Clinical relevance for conformal radiotherapy treatment planning. *Strahlenther Onkol* 2002; **178**: 263–8.
20. Dinh CV, Steenbergen P, Ghobadi G, et al. Magnetic resonance imaging for prostate cancer radiotherapy. *Phys Med* 2016; **32**: 446–51.
21. Nyholm T, Jonsson J, Söderström K, et al. Variability in prostate and seminal vesicle delineations defined on magnetic resonance images, a multi-observer, −center and -sequence study. *Radiat Oncol* 2013; **8**: 126.
22. Hentschel B, Oehler W, Strauß D, Ulrich A, Malich A. Definition of the CTV prostate in CT and MRI by using CT–MRI image fusion in IMRT planning for prostate cancer. *Strahlenther Onkol* 2011; **187**: 183–90.
23. Khoo VS, Padhani AR, Tanner SF, Finnigan DJ, Leach MO, Dearnaley DP. Comparison of MRI with CT for the radiotherapy planning of prostate cancer: a feasibility study. *Br J Radiol* 1999; **72**: 590–7.
24. Willigenburg T, de Muinck Keizer DM, Peters M, et al. Evaluation of daily online contour adaptation by radiation therapists for prostate cancer treatment on an MRI-guided linear accelerator. *Clin Transl Radiat Oncol* 2021; **27**: 50–6.
25. Blitzer GC, Yadav P, Ko HC, et al. Visualizing the prostatic urethra by MRI voiding scan: Results of a prospective clinical trial evaluating a novel approach to radiotherapy simulation for prostate cancer. *Int J Radiat Oncol Biol Phys* 2020; **108**: e264–5.
26. Richardson M, Skehan K, Wilton L, et al. Visualising the urethra for prostate radiotherapy planning. *J Med Radiat Sci* 2021; **1**: 1–7.
27. Adair Smith G, Dunlop A, Alexander SE, et al. Evaluation of therapeutic radiographer contouring for magnetic resonance image guided online adaptive prostate radiotherapy. *Radiother Oncol* 2023; **180**: 109457.