Original Article

Pilot study: rectal contouring variability in patients treated with radical radiotherapy for prostate cancer: impact on rectum dose—volume histograms

Jacqueline E. Martin, P. Kirkbride

Clinical Oncology Department Weston Park Hospital, Sheffield, UK

Abstract

Purpose: The organ at risk in prostate radiotherapy is the anterior rectal wall. This pilot study was conducted to quantify the inter-observer variability of rectal contouring in conformal prostate radiotherapy planning using four contouring methods and to determine a standard rectal contouring definition.

Methods and materials: Ten patients with T1/T2 disease and six clinical oncologists were recruited. Two cross-sectional and two length limits generated four rectal volumes. Each clinician contoured the four volumes for each patient and the dose—volume histograms (DVHs) were analysed. The percentage rectal volume receiving 20%, 50%, 80%, 90% and 95% of the total delivered dose and the mean and median rectal doses were calculated. Data were presented as mean ± 2 standard deviations.

Results: The Sh Rec method (contouring the rectum including its contents extending from 2 cm above the upper limit of the prostate to 2 cm below the prostatic apex) was the least variable in the 80%, 90% and 95% percentage ranges. The mean difference in Sh Rec-contoured volume was 18.7 cm³ (± 22.3 cm³).

Conclusions: The Sh Rec-contouring method showed the least inter-observer variability. The results are informative and will help define a standard rectal contouring method.

Keywords
contouring; prostate; radiotherapy; rectum; variability

INTRODUCTION

The aim of any kind of treatment is ‘first do no harm’; following this concept in prostate radiotherapy, a variety of techniques can be used to achieve high-radiation doses within the tumour itself while endeavouring to keep radiation doses to the neighbouring tissues to a minimum. Techniques include conventional external beam radiotherapy, three-dimensional conformal radiotherapy (3DCRT), intensity-modulated radiotherapy (IMRT) and ultrasound-guided trans-perineal radio-implant. Planning and delivery of conformal radiotherapy for prostate cancer is a multi-step process, from the localisation of the target volume within the patient to the consistently reproducible positioning of the patient for treatment.
The anterior rectal wall is the tissue considered as the organ at risk (OAR) in prostate radiotherapy. Although, it is intuitively obvious that precision in radiation delivery and avoidance of OARs would be beneficial in reducing side effects of treatment, reducing the physical dose to the anterior rectal wall without a similar reduction in the posterior peripheral zone of the prostate is difficult because of the proximity of the prostate to the anterior rectal wall.1

Literature evidence has already established that several ways of contouring the rectum and generating dose–volume distributions over the organ are possible.2,3 Acknowledging that the anterior rectal wall is the tissue considered as the OAR and that the volume defined is influenced by both the cross-sectional contour as well as the length of the contoured organ, the aims of this study were to create a robust rectal contouring definition influenced only by a modest inter-observer variability error factor, to quantify the inter-observer variability of rectal contouring in the CT planning of radical prostate radiotherapy and to determine a standard rectal contouring definition to be used at Weston Park Hospital, Sheffield, in future external beam prostate radiotherapy planning.4

**METHODS AND MATERIALS**

**Patients**

Ten patients with T1 or T2 biopsy-confirmed prostate cancer whose treatment plans with planning target volume (PTV) prostate clinical target volume (CTV) and beam arrangement were available. Patients were excluded from participation if they had previous pelvic surgery or hip replacement surgery. All patients gave written informed consent to participate and the study was approved by the Local Research Ethics Committee.

**Clinical oncologists**

Two consultant clinical oncologists with a special interest in urological malignancies and four clinical oncology specialist registrars, who had received no less than 6 months of formal training in urological malignancies, were recruited.

**Image acquisition**

Varian Medical Systems radiotherapy planning and delivery system were used. It included the ‘Eclipse’ CT and virtual simulation/planning system, ‘Acuity’TM and ‘Portal Vision’ treatment verification systems and ‘Clinac®’, the treatment delivery system. The treating clinical oncologist delivered the radical radiotherapy treatment to the patient as per their normal practice, with their own contouring parameters and a hypo-fractionated dosing schedule of 55 Gray (Gy) in 20 fractions, delivered Monday to Friday, in a three-field arrangement. Copies of the planning scans were used in the study. All contours were ‘hidden’ on the CT images from the participating clinician (i.e., rectal contour, prostate contour and PTV contour).

**Rectal delineation**

The six participating oncologists were given written instructions defining four rectal volume parameters. The planning software was adapted to allow the participating clinicians to delineate the rectal contours anonymously. Throughout the delineation process, the clinician was unable to view any previously delineated structures, the CTV or the PTV.

The rectal cross-sectional limits were the anterior half of rectal wall only, as anterior rectal wall is the tissue considered as the OAR, i.e., anterior rectal wall (arw) and the whole rectum including contents, i.e., rectum (rec). The rectal length limits were defined as: Long (Lg), whereby, the cranial border started at level where the rectum turned horizontally into the sigmoid and the caudal border finished at 2 cm below the prostatic apex; Short (Sh), whereby, the cranial border started at 2 cm above the upper limit of the prostate and the caudal border finished at 2 cm below the prostatic apex. These cross-sectional and length limits generate four possible rectal volumes, i.e., Sh arw, Lg arw, Sh rec and Lg rec (Figures 1–6).

Each clinician contoured these four rectal volumes on each of the 10 patients, yielding 240 rectal contours for analysis.
Data collection
All rectal volume contours were evaluated using the CT planning computer system and dose–volume histograms (DVHs). The volume in cm$^3$ of each rectal volume was computed by the planning system and recorded. The uppermost (cranial) and lowermost (caudal) CT axial slice containing a rectal contour was recorded. DVHs of all volumes were produced using the ‘Eclipse’ planning system and the original three-field treatment-planning algorithm. From these DVHs, the percentage rectal volume receiving 20%, 50%, 80%, 90% and 95% of the total delivered dose was recorded, namely v20, v50, v80, v90 and v95, respectively (Figure 7).

The mean and median rectal dose as calculated by the planning system was recorded for each plan.
Statistical analysis
The data set consisted of four dependent samples which posed restrictions to statistical analysis in terms of yielding a $p$-value for statistical significance. Ninety-five percent confidence intervals serve as a meaningful representation of the degree
of variance. Moreover, as the populations are small, means and standard deviations of the mean compared with medians and inter-quartile ranges are equally affected by outliers; hence, all data are presented as mean ± 2 standard deviations.

RESULTS

Ten patients provided written informed consent. All ten patients had histologically and radiologically proven T1 or T2 prostate cancer. All
were planned to receive radical radiotherapy for their prostate cancer at a dose schedule of 55 Gy in 20 fractions. The mean age of the study population was 61.4 years (50–74 years).

The Gleason sum score is a microscopic grading of prostate cancer cells, graded on a scale of 1–5. The Gleason’s score is the sum of the dominant and secondary patterns. Higher scores indicate a greater potential for aggressive growth. The Gleason sum score was 6 (3 + 3) in four of the patients and 7 in the remainder (3 + 4 in three patients and 4 + 3 in three patients). The mean prostate-specific antigen (PSA) value was 8.1 (3.5–22.3) and median PSA was 9.9.

These cross-sectional and length limits generated four possible rectal volumes, i.e., Sh arw, Lg arw, Sh rec and Lg rec. First, the cranial and caudal CT slice was recorded for each volume. The range of variability of the caudal CT slice level was 3–7 five-millimetre width axial slices (15–35 mm difference) with a mean difference of 25 mm. The range of long rectal volume cranial levels recorded was 1–9 five-millimetre width axial slices (5–45 mm difference) with a mean difference of 25 mm. Finally, the range of short rectal cranial levels recorded was from 1–6 five-millimetre slices (5–30 mm difference) with a mean variation of 10 mm difference (Table 1).

Table 1. Variation in cranial and caudal limits on CT planning

<table>
<thead>
<tr>
<th>Number of CT images (mm)</th>
<th>Difference (mm)</th>
<th>Mean difference (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caudal limit (Lg and Sh)</td>
<td>3–7</td>
<td>15–35</td>
</tr>
<tr>
<td>(2 cm below the prostate apex)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cranial limit (Lg) (level where the rectum turned horizontally into the sigmoid)</td>
<td>1–9</td>
<td>5–45</td>
</tr>
<tr>
<td>Cranial limit (Sh) (2 cm above the upper limit of the prostate)</td>
<td>1–6</td>
<td>5–30</td>
</tr>
</tbody>
</table>
Second, the rectal volume expressed in cubic centimetres was recorded. The mean difference in volume for the Lg arw was 16.4 (±10.43 cm³), Lg rec was 26.52 (±23.72 cm³), Sh arw was 16.9 (±11.2 cm³) and Sh rec was 18.7 (±22.3 cm³) (Table 2).

The original treatment beam arrangements and dose prescription of 55 Gy in 20 fractions were applied to the plans and DVHs generated for each rectal contour for each patient. The beam arrangement for each plan was an anterior and two lateral beams of 10 MV energy. The percentage volume of each rectal contour that received 20%, 50%, 80%, 90% and 95% of the applied dose was collected from the histograms using the ‘Eclipse’ computer system. The minimum, maximum, mean and median doses expressed as a percentage of the applied total dose schedule of 55 Gy for each rectal contour was recorded. The 95% confidence interval was calculated for each dataset, i.e., ±2 SD (Tables 3 and 4).

When this data is graphically presented, it is appreciated that variability existed between the contouring methods at the 20%, 50%, 80%, 90% and 95% dose levels for both the mean and median values (Figures 8 and 9).

**DISCUSSION**

The need to concentrate on the contouring variability techniques and inter-observer variability is vital to ensure best practice and reduce

---

**Table 2. Summary of the mean and median difference in volume recorded for each rectal volume parameter**

<table>
<thead>
<tr>
<th>Volume label</th>
<th>Mean difference in volume (cm³)</th>
<th>± 2 SD of the mean volume (cm³)</th>
<th>Median difference in volume (cm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lg Arw</td>
<td>16.4</td>
<td>10.43</td>
<td>14.3</td>
</tr>
<tr>
<td>Lg Rec</td>
<td>26.52</td>
<td>23.72</td>
<td>22.6</td>
</tr>
<tr>
<td>Sh Arw</td>
<td>16.9</td>
<td>11.2</td>
<td>17.1</td>
</tr>
<tr>
<td>Sh Rec</td>
<td>18.7</td>
<td>22.3</td>
<td>15.5</td>
</tr>
</tbody>
</table>

**Table 3. Mean and median standard deviations of percentage rectal volumes**

<table>
<thead>
<tr>
<th></th>
<th>Lg arw</th>
<th>Lg rec</th>
<th>Sh arw</th>
<th>Sh rec</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>20%</td>
<td>0.94</td>
<td>0.76</td>
<td>0.97</td>
<td>0.87</td>
</tr>
<tr>
<td>50%</td>
<td>6.11</td>
<td>2.11</td>
<td>15.60</td>
<td>17.58</td>
</tr>
<tr>
<td>80%</td>
<td>19.68</td>
<td>13.83</td>
<td>13.13</td>
<td>9.3</td>
</tr>
<tr>
<td>90%</td>
<td>22.99</td>
<td>21.45</td>
<td>11.98</td>
<td>8.4</td>
</tr>
<tr>
<td>95%</td>
<td>19.85</td>
<td>17.71</td>
<td>13.31</td>
<td>11.51</td>
</tr>
</tbody>
</table>

**Table 4. Degree of difference between the mean and the median 2 SD for each of the contour types**

<table>
<thead>
<tr>
<th></th>
<th>Lg arw</th>
<th>Lg rec</th>
<th>Sh arw</th>
<th>Sh rec</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>0.18</td>
<td>0.10</td>
<td>0.20</td>
<td>0.40</td>
</tr>
<tr>
<td>50%</td>
<td>4.00</td>
<td>-1.98</td>
<td>4.63</td>
<td>5.49</td>
</tr>
<tr>
<td>80%</td>
<td>5.85</td>
<td>3.83</td>
<td>2.91</td>
<td>4.15</td>
</tr>
<tr>
<td>90%</td>
<td>1.54</td>
<td>3.58</td>
<td>0.51</td>
<td>1.11</td>
</tr>
<tr>
<td>95%</td>
<td>2.14</td>
<td>1.80</td>
<td>3.42</td>
<td>3.15</td>
</tr>
</tbody>
</table>
morbidity in prostate cancer patients treated using our current three-field conformal techniques. Moreover, with the introduction of dose escalation schedules with IMRT techniques, the need is greater. Late rectal toxicity is considered to be the dose-limiting complication of radiotherapy for localised prostate cancer. The volume of the rectum, which displays a dose volume effect, irradiated to high doses, is a crucial predictor of late rectal damage. Dose distributions delivered to the rectal wall over a course of radiotherapy inevitably vary due to daily unavoidable movement of the rectal wall. Conventions for rectal contouring vary. Cross-sectional contours used in previous studies include, rectum and its contents, rectal wall
only or anterior rectal wall. Lengths of the rectum used include anatomic rectum or rectum over the length of the treatment field. This leaves it difficult to extrapolate dose-constraint recommendations from literature into clinical practice. In order to use dose-volume cut-off recommendations from literature, without making over-estimations or under-estimations of incidence of rectal toxicity, the identification of a robust definition of the rectal volume is required.

Fiorino et al. studied the impact of inter- and intra-observer variability of rectal contouring on radical prostate plans. Inter-observer variability being the variability that exists between two individuals and intra- being the variability seen when one individual, with a 6-month time interval, delineates contours. Intra-observer variability was found to be lower than inter-observer variability. They concluded that: 'once a robust definition of rectum is assessed, the intra- and inter-observer variability would be modest'. Other studies have observed similar variations of absolute rectal volumes in DVHs used for the prediction and calculation of late radiation side effects, depending on the definition of the rectal borders used.

This pilot study was conducted to quantify the inter-observer variability of rectal contouring in the CT planning of radical prostate radiotherapy using four dependent populations (four contouring methods applied to the same sample population) and determine a standard rectal contouring definition to be used at Weston Park Hospital, Sheffield, in future prostate radiotherapy planning. The cross-sectional volumes chosen in this study were the rectum including its contents and the anterior rectal wall. Ideally, to delineate part of, or all of the rectal wall, excluding bowel contents, a validated mathematical algorithm would be required to allow for poor discrimination between rectal wall and any retained stool in the rectum post evacuation. Despite the aforementioned limitations with rectal wall only volume, as the anterior rectal wall is considered to be the OAR in radical prostate radiotherapy, delineation of this cross-sectional volume was included in the study, although it was unsurprising that large variability existed in this volume’s delineation as clinical judgement was a confounding factor. The length limits chosen in this study were the Long length (Lg) and the Short length (Sh). The Long length endeavoured to represent the anatomical rectum. The dentate line, i.e., the level within the colon, where the anus changes to rectum would be subject to considerable inter-observer variability; to endeavour to minimise this, 2 cm below the prostate was considered to be an approximation of the start of the rectum. The upper limit of the long rectum limit was based on the anatomical definition of the point at which the rectum turns horizontally into the sigmoid colon. As noted with the anterior rectal wall, it was to be expected that greater clinical judgement would be required than a definite point above the prostate apex. The short rectum (Sh) limits, i.e., 2 cm above and below the prostate were chosen to ensure that all the rectum within the PTV (i.e., prostate (CTV) + 1 cm) was determined to predict toxicity. Due to the limitations brought about by the study design, the use of a statistical significance test would disregard elements of the test assumptions and then a p-value would be quoted in error. The analysis carried out in this study looked at the variability within each subject and then, for each method, calculated a mean standard deviation. These mean standard deviations were compared visually between the different methods to find the contouring method with the least variability. We acknowledge that with this small sample size, the contouring methods are likely to yield larger confidence interval limits than would be expected in a larger population. However, we have demonstrated that in a small study population, fixed measurements, namely 2 cm above or below the prostate, were more consistently discernable on radiological image sets by multiple individuals than points that were influenced by a difference in interpretation or misinterpretation of a definition, namely ‘the point at which the rectum turns horizontally into the sigmoid colon’. It was noted that a maximum difference of 9 five-millimetre CT slices (45 mm) were recorded as the point at which the rectum turned horizontally into the sigmoid colon between individual doctors. The point 2 cm above and below the prostate recorded by clinicians varied
The volume difference between the maximum and minimum volumes was noted to be greatest in the Lg Rec contour method. This was likely to be influenced by the fact that the cranial level was subject to a greater than modest inter-observer variability. The variability was magnified by the fact that the cross-sectional area was greater than the Lg Arw contour method.

Neal et al. evaluated the relative merits of different field arrangements most frequently used for conformal radiotherapy of the prostate using DVHs and normal tissue complication probabilities (NTCPs). Isocentric plans for each patient were devised using three, four, six and eight conformal field plans. The plans were evaluated using DVHs of the OARs namely bladder, rectum and both femoral heads. The authors concluded that none of the techniques studied consistently proved to be superior when applied to prostate cancer plans with respect to sparing all of the OARs.8

Muren et al. reviewed the use of margins to account for geometrical uncertainties around the rectum in radiotherapy planning. By analysing repeated weekly CT scan of 16 bladder cancer patients and 6 prostate cancer patients, and comparing the position of the rectum in each scan with the conformal radiotherapy planning scans, it was noted that 24% of the repeat scan rectal volumes were displaced outside the planning scan contours, and wall movements of up to 30 mm were observed. They concluded, that rectal planning OAR volume margins of 5–6 mm will encompass the systematic component of rectal motions in 89% of cases, while margins up to 16 mm are required to account for most of the random movements observed.9

Research evidence consistently supports the use of rectal DVHs as a means of predicting late rectal toxicity. The type of dose–volume analysis differs between series, some using DVHs and others using biologically equivalent DVHs.10 Greco et al. retrospectively analysed the DVHs of 135 patients with prostate cancer treated with 3DCRT between 1996 and 2000 to a total dose of 76 Gy. The area under the percent volume DVH for the rectum of the bleeding patients was significantly higher than in the patients without late rectal toxicity. Rectal DVHs are an essential tool used to discriminate patients into a low- and a high-risk category of developing complications of late rectal bleeding.11,12 Radiation therapy oncology group (RTOG) rectal toxicity score RTOG 2 where p = 0.042.13

Bauman and Rodrigues concluded that, in order to facilitate comparisons between series, reports should include standard information about percentage of rectal volume receiving greater than the specified dose.10 During radiation treatment planning, doses to surrounding normal tissues/OARs, i.e., V95, D50 and Dm should be kept to a minimum.14 The dose–volume constraints to minimise rectal toxicity recommended at our study institute are such that no more than 50% of the rectal volume should receive greater than 90% of the total applied dose. Hence, in our study, variability of the higher DVH dose parameters was deemed more important than the lower dose parameters. The volume generated by contouring the rectum including the contents 2 cm above to 2 cm below the prostate (Sh Rec volume) on each of the 10 patients by the 6 doctors showed the least variability for the higher dose percentage levels of 80%, 90% and 95% of the total applied radiotherapy dose of 55 Gy in 20 fractions. The mean ± standard deviation of the average percentage volume that received 80%, 90% and 95% of the total applied dose was lowest for the Sh Rec, with values of ±8.12 cm3, ±3.75 cm3 and ±5.79 cm3 of their volume means, respectively. Dale et al. tried to demonstrate some difference in correlation between the histograms generated from contouring the whole rectum and contouring the rectal wall only, but their results were not statistically significant.15 Hornby et al. investigated whether the length of the rectum outlined affected the NTCP. Our results are similar to the conclusions of Hornby et al., i.e., that a standardised delineated length of 2 cm beyond the beam edge for normal structures is recommended.16 It should be noted that our study used the point
of reference of 2 cm above and below the prostate as the study design did not permit the participating clinician to view the planned target volume (PTV), i.e., the 95% isodose, and therefore the ‘beam edge’ which would be represented by the 50% isodose. A further oversight in our study design was that the 2 cm above and below the prostate to ensure that the entire volume was within the treatment field limits did not account for any increase in the upper limit of the PTV generated by the inclusion of seminal vesicles in the higher risk patient. This omission most certainly would lead to an underestimation of toxicity in those patients.

CONCLUSIONS
The findings of this pilot study indicated that a Sh Rec–contouring method was the least variable contouring method in the 80%, 90% and 95% percentage ranges. Despite the low numbers of patients and the statistical analysis restrictions, the results are informative and will help define a standard rectal contouring method for Weston Park Hospital. If this standard is used, potentially rectal toxicity data could then be assessed in those that have their treatments delivered with a robust rectal definition.

A larger study should also be considered, whereby the rectal contouring method of Sh Rec would be relative to the PTV rather than the prostate alone and then would be assessed for variance and a statistical significance test applied. Other factors to be taken into consideration would include CT slice width, i.e., decreasing the CT axial slice width to 2.5 mm from 5 mm and the addition of an ‘at risk margin’ to the contoured ‘short rectum and its contents’, Sh Rec, to allow for rectal wall movement throughout the whole treatment schedule.

ACKNOWLEDGEMENTS
The authors extend their gratitude to Dr. C.J. Ferguson, Dr. O. Din, Dr. S. Garbutt, Dr. M. Fernando, Dr. J. Horsman, Ms. S. Otter, Mr. D. Ramsden and Dr. L. Jones for their help with the study.

References
