(Un-)certainties in radiotherapy of rectal cancer

Nijkamp, J.A.

Citation for published version (APA):
Target Volume Delineation Variation in Radiotherapy for Early Stage Rectal Cancer in the Netherlands

Jasper Nijkamp, Danielle F.M. de Haas-Kock, Jannet C. Beukema, Karen J. Neelis, Dankert Woutersen, Heleen Ceha, Tom Rozema, Annerie Slot, Hanneke Vos-Westerman, Martijn Intven, Patty H Spruit, Yvonne van der Linden, Debby Geijsen, Karijn Verschueren, Marcel van Herk and Corrie A.M. Marijnen on behalf of the National Platform for Radiotherapy of Gastroenterology Tumors

1Department of Radiation Oncology, Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam,
2Department of Radiation Oncology, Maastricht University Hospital, Maastricht,
3Department of Radiation Oncology, University Medical Center Groningen, Groningen,
4Department of Clinical Oncology, Leiden University Medical Center, Leiden,
5Department of Radiation Oncology, Medical Spectrum Twente, Enschede,
6Department of Radiation Oncology, Radiotherapeutic Institute Friesland, Leeuwarden,
7Department of Radiation Oncology, Dr. Bernard Verbeeten Institute, Tilburg,
8Department of Radiation Oncology, Isala Clinics, Zwolle,
9Department of Radiation Oncology, University Medical Center Utrecht, Utrecht,
10Department of Radiation Oncology, VU University Medical Center, Amsterdam,
11Department of Radiation Oncology, Amsterdam Medical Center, Amsterdam
Chapter 4

Abstract

Purpose
The aim of this study was to measure and improve the quality of target volume delineation by means of national consensus on target volume definition in early-stage rectal cancer.

Material and methods
The CTV’s for eight patients were delineated by 11 radiation oncologists in 10 institutes according to local guidelines (phase 1). After observer variation analysis a workshop was organized to establish delineation guidelines and a digital atlas, with which the same observers re-delineated the dataset (phase 2). Variation in volume, most caudal and cranial slice and local surface distance variation were analyzed.

Results
The average delineated CTV volume decreased from 620 to 460 cc (p < 0.001) in phase 2. Variation in the caudal CTV border was reduced significantly from 1.8 to 1.2 cm SD (p = 0.01), while it remained 0.7 cm SD for the cranial border. The local surface distance variation (cm SD) reduced from 1.02 to 0.74 for anterior, 0.63 to 0.54 for lateral, 0.33 to 0.25 for posterior and 1.22 to 0.46 for the sphincter region, respectively.

Conclusions
The large variation in target volume delineation could significantly be reduced by use of consensus guidelines and a digital delineation atlas. Despite the significant reduction there is still a need for further improvement.
**Introduction**

Pre-operative radiotherapy of rectal cancer patients can be improved by using intensity-modulated radiotherapy (IMRT), with a significant reduction in dose to the organs at risk [1-3]. As more conformity to the target volume is obtained with IMRT, knowledge and application of geometrical uncertainties in the construction of a planning target volume (PTV) is important to prevent under-dosage of the target volume.

For construction of the PTV, definition of the clinical target volume (CTV) is essential. In rectal cancer radiotherapy two guidelines are available (4, 5). In both the tumor, involved lymph nodes, mesorectum and the perirectal, presacral and internal iliac lymph node regions are included. Roels et al. [4] developed an atlas by selecting required sub-regions based on patterns of local recurrences and subsequently defined the anatomical borders of these CTV sub-regions. The atlas of Myerson et al. [5] was developed by a RTOG consensus panel.

Despite the existence of the first atlas since 2006 [4], little publications are available on the effect on inter-observer delineation variation. For the RTOG consensus atlas [5] one pilot study [6] is available where 13 observers delineated the CTV for 1 patient according to the protocol guidelines, and subsequently 7 re-delineated using the delineation atlas. In this study initial inter-observer variation of over 1.5 cm standard deviation (1SD) was found. This observer variation was much larger than the current known major source of geometric uncertainty, the target volume shape variation, which is in the order of 0.7 cm for systematic and random errors [7, 8]. Re-delineation with the atlas resulted in significantly better agreement in target delineation, but still variation up to 1.0 cm (1SD) was observed. Since this pilot study contains only one patient and the atlas was tested by only 7 observers, more elaborated studies are needed to confirm the beneficial effect of a delineation atlas and rule out the possible selection bias of a single patient. It is, for example, unknown whether rectal cancer CTV delineation variation is influenced by anatomic differences between male and female patients or by tumor location.

In the Netherlands, several centers are planning to replace 3- or 4-field conformal RT by IMRT. National consensus guidelines in order to reduce delineation variation are needed [6] for safe IMRT introduction. The Dutch National Platform for Radiotherapy of Gastroenterology Tumors initiated a 2-phase delineation study. The purpose of the study was

- to examine the delineation variation between radiation oncologists (phase1);
- to establish consensus guidelines and produce a delineation atlas;
- to evaluate the effect of use of these by the same observers (phase 2).
Material and methods

Patients and observers

Eight patients with early stage rectal cancer (cT1-3, N0-1, no circumferential resection margin (CRM) involvement) treated with pre-operative RT were retrospectively selected. In the selection process patient orientation (prone/supine on a flat table), gender (male/female) and tumor location (low/high seated tumor) were taken into account (Table 4.1). Low seated tumors were defined as tumors within 5 cm from the anal verge to be treated with an abdominoperineal resection (APR). High seated tumors were defined as more than 5 cm from the anal verge and to be treated with a low-anterior resection (LAR).

A total of 24 observers representing all 21 Dutch radiotherapy institutes were asked to participate.

Scans and delineation software

For each patient a planning CT scan ranging from the L2-L3 junction to below the perineum with 5 mm slice spacing was available. Five patients were scanned with intravenous contrast. A T2 weighted MR scan, reconstructed in orthogonal planes, was available for all patients.

The target volumes were delineated using the “Big Brother” software [9]. The CTV was delineated on the transversal slices. Sagittal and coronal views of the CT scan were shown to the observer simultaneously, also at the same time MR reconstructions were available in a separate window, without co-registration, because of large anatomic differences between the CT and MRI scans. All human-computer interaction was recorded during the delineation sessions. These were used to register the delineation time per observer per patient, as well as the frequency of the atlas use in the second phase.

Study layout

In the first phase observers were asked to delineate the CTV based on the provided clinical data and according to the local hospital guidelines. After analyzing the data of phase 1 a meeting was organized with the observers. Inter observer variation results were discussed and consensus guidelines were developed based on the delineation guidelines of Roels et al. [4]. The consensus aim was to describe which regions should be delineated for which patients, and which anatomic borders were appropriate to define these regions. The guidelines were subsequently translated into a digital delineation atlas by three experts (D.F.M. de H.-K., J.C.B. and C.A.M.M.) who were not part of the observer group. The digital atlas consisted of an in house developed application in which all delineation regions according to the guidelines were provided for a rectal cancer case. Difference with existing paper based atlases is that observers could approach the digital atlas similar to CT scans in delineating target volumes.

In the second phase, starting 7 months after the consensus meeting, observers re-delineated the 8 patients using the consensus guidelines and the digital atlas. Observers were free to either delineate the separately described regions of the CTV, or to compose one CTV.
Data analysis

The following analyses were separately performed for phases 1 and 2 and compared subsequently.

For volumetric analysis the mean and standard deviation of the delineated volume was calculated per patient. The generalized conformity index (CI), sum of the common volumes divided by the sum of the encompassing volumes between each pair of observers, was calculated for each patient [10]. The cranial-caudal delineation variation was calculated as the standard deviation in the position of the most cranial and caudal delineated slice over the observers.

For each patient, a median 3D surface of the CTV’s of all observers was computed as reference structure [11]. The median surface represents 50% coverage of the delineations of all observers. The median surface was subsequently sampled using 8000 equally distributed points. For each point the perpendicular distance to each delineated CTV surface was calculated. The standard deviation over the distances for each point

Table 4.1: Volumetric and cranial caudal slice delineation variation for each patient in phase 1 and phase 2 for 11 observers. For the p-values a 2-sided paired T-Test was used to compare phase 1 and 2

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Average volume in cc (1SD)</th>
<th>Caudal slice variation 1SD (cm)</th>
<th>Cranial slice variation 1SD (cm)</th>
<th>Conformity Index (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1, local hospital guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: female, prone, APR</td>
<td>638 (147)</td>
<td>1.5</td>
<td>0.7</td>
<td>0.56</td>
</tr>
<tr>
<td>2: female, supine, APR</td>
<td>687 (121)</td>
<td>1.6</td>
<td>0.7</td>
<td>0.60</td>
</tr>
<tr>
<td>3: male, prone, APR</td>
<td>761 (154)</td>
<td>2.4</td>
<td>0.7</td>
<td>0.62</td>
</tr>
<tr>
<td>4: male, supine, APR</td>
<td>623 (132)</td>
<td>2.1</td>
<td>0.8</td>
<td>0.64</td>
</tr>
<tr>
<td>5: female, prone, LAR</td>
<td>520 (95)</td>
<td>1.6</td>
<td>0.7</td>
<td>0.60</td>
</tr>
<tr>
<td>6: female, supine, LAR</td>
<td>516 (95)</td>
<td>1.7</td>
<td>0.5</td>
<td>0.65</td>
</tr>
<tr>
<td>7: male, prone, LAR</td>
<td>609 (82)</td>
<td>1.7</td>
<td>0.7</td>
<td>0.70</td>
</tr>
<tr>
<td>8: male, supine, LAR</td>
<td>602 (86)</td>
<td>1.3</td>
<td>1.0</td>
<td>0.65</td>
</tr>
<tr>
<td>RMS/Average</td>
<td>620</td>
<td>1.8 cm</td>
<td>0.7</td>
<td>0.63</td>
</tr>
</tbody>
</table>

| Phase 2, delineation guidelines and digital atlas |
|--------------------------|---------------------------|---------------------------------|---------------------------------|----------------------|
| 1: female, prone, APR | 455 (88) | 0.6 | 0.5 | 0.61 |
| 2: female, supine, APR | 472 (75) | 0.6 | 0.8 | 0.66 |
| 3: male, prone, APR | 573 (92) | 2.0 | 0.5 | 0.68 |
| 4: male, supine, APR | 499 (52) | 0.5 | 0.7 | 0.73 |
| 5: female, prone, LAR | 423 (77) | 1.7 | 0.8 | 0.62 |
| 6: female, supine, LAR | 384 (57) | 0.8 | 0.4 | 0.70 |
| 7: male, prone, LAR | 444 (72) | 0.6 | 0.9 | 0.68 |
| 8: male, supine, LAR | 430 (79) | 1.3 | 0.9 | 0.61 |
| RMS/Average | 460 | 1.2 | 0.7 | 0.66 |

Different from phase 1  

| p < 0.001 | p = 0.01 | p = 0.46 | p = 0.06 |
**Fig. 4.1:** Example images of patient 2 showing the difference in delineation variation between the 1st (left) and 2nd (right) delineation phase. The top image is the most caudal common delineated slice. The middle image is a mid-rectum slice just cranial of the bladder. The bottom image is the sagittal view. Each observer has a corresponding color on all images.
Target volume delineation variation for early stage rectal cancer in the Netherlands

was calculated as a measure of local observer variation (local SD). The overall observer variation (overall SD) for each patient was calculated as the quadratic mean of the local SD weighted for surface. For sub-region analysis the median surface was divided into 4 regions for the high seated tumors, being posterior, the left and right lymph node region, and anterior. For the distal tumors an additional sphincter region was defined caudal from the caudal edge of the mesorectum. For each sub-region the overall observer variation was calculated.

To analyze the impact of the guidelines and atlas in clinical practice, the time spent delineating the patients in both rounds was evaluated, as well as the frequency of viewing the atlas in the second round.

Average volumes and the conformity index within and between phases 1 and 2 were compared using a paired 2-sided Student’s t-test. To compare standard deviations a 2-sided F-test was used. p-Values lower than 0.05 were considered to be statistically significant.

Results

A total of 17 observers submitted the delineations in phase 1. Of these 17 observers, 11 also delineated phase 2. In the 1st phase, 1 of the 11 observers did not delineate patient number 3. In the 2nd phase, 1 observer was missing for patient 8. Only delineations of the 11 observers who were present in both delineation sets are considered in the remainder of this article, resulting in a total of 172 evaluated CTV delineations.

Phase 1

The observers were not confident with the delineation software, whilst the first delineation lasted on average 36 min per observer (1SD 16.3), the remainder took about 17 min (1SD 9.4). All observers made use of the coronal and sagittal representation of the CT scan and the MR scans, as was recorded by changes in level, window, zoom and slice in these viewers.

Large differences between observers were present, as can be appreciated from the example images of patients 2 (Fig. 4.1) and 8 (Fig. 4.2). The average delineated CTV volume of the 8 patients ranged from 516 cc to 761 cc, with standard deviations ranging from 82 cc to 154 cc (Table 4.1). With an average volume of 675 cc for the APR patients and 562 cc for the LAR patients, there was a clear difference in CTV definition for both patient groups (p < 0.0001, 2-sided 2-sample t-test). Part of the volumetric variation was due to differences in definition of the caudal and cranial border of the CTV (Table 4.1). For the caudal border differences up to 14 slices between observers were found. The cranial border differences were smaller, but still substantial. Delineation variation was also demonstrated by the conformity index, with an average CI of 0.63.

From the local surface SD between observers projected on the median surface of each patient we learn that the pattern of delineation variation is similar between patients (Fig. 4.3). With a color scale up to 2.2 cm it is clear that large differences exist between observers, especially for the anterior CTV part (average 1.02 cm SD) and the sphincter region (average 1.22 cm SD) (Table 4.2). For the lateral lymph node regions the observer variation is lower, but still substantial with an average 0.6 cm SD.
Fig. 4.2: Example images of patient 8 showing the difference in delineation variation between the 1st (left) and 2nd (right) delineation phase. The top image is the most caudal common delineated slice. The middle image is a mid-rectum slice just cranial of the bladder. The bottom image is the sagittal view. Each observer has a corresponding color on all images.
Consensus guidelines and digital atlas
At the consensus meeting in October 2009, the delineation guidelines were established (Appendix 4.1). Four major CTV regions were described, being the mesorectum, the iliac lymph node regions (left and right), the obturator lymph node region (left and right) and the presacral region. For the mesorectal region, differences were made between patients undergoing a LAR, where the caudal border of the CTV was set to 4 cm caudal of the tumor, and patients undergoing an APR, where the whole sphincter complex was included in the CTV. The obturator region was only included if the primary tumor was <10 cm from the anal verge [4]. A 1 cm margin around the gross tumor volume (GTV) and 0.5 cm margin around pathologically enlarged lymph nodes was applied.

For the digital delineation atlas a planning CT scan with 2 mm slice distance was available with a patient in supine position, together with a T2 weighted MR scan (reconstructed in orthogonal planes). The MR scans were co-registered to the planning CT, since anatomical differences were small. The regions defined in the guidelines were delineated by the 3 observers and subsequently the consensus of these 3 observers was used in the digital atlas. The digital atlas can be downloaded from: ftp://ftp-rt.nki.nl/RectumAtlas/RectumAtlas.zip

**Table 4.2:** Surface distance variation for phase 1 and phase 2 for 11 observers.

<table>
<thead>
<tr>
<th>Phase 1, local hospital guidelines</th>
<th>Total CTV</th>
<th>Posterior</th>
<th>Left</th>
<th>Anterior</th>
<th>Right</th>
<th>Sphincter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>0.79</td>
<td>0.22</td>
<td>0.62</td>
<td>1.35</td>
<td>0.65</td>
<td>1.07</td>
</tr>
<tr>
<td>Patient 2</td>
<td>0.73</td>
<td>0.28</td>
<td>0.59</td>
<td>1.02</td>
<td>0.68</td>
<td>1.01</td>
</tr>
<tr>
<td>Patient 3</td>
<td>0.84</td>
<td>0.31</td>
<td>0.62</td>
<td>0.93</td>
<td>0.54</td>
<td>1.44</td>
</tr>
<tr>
<td>Patient 4</td>
<td>0.76</td>
<td>0.29</td>
<td>0.61</td>
<td>0.91</td>
<td>0.60</td>
<td>1.30</td>
</tr>
<tr>
<td>Patient 5</td>
<td>0.79</td>
<td>0.34</td>
<td>0.69</td>
<td>1.21</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Patient 6</td>
<td>0.57</td>
<td>0.35</td>
<td>0.59</td>
<td>0.91</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>Patient 7</td>
<td>0.53</td>
<td>0.33</td>
<td>0.58</td>
<td>0.74</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Patient 8</td>
<td>0.68</td>
<td>0.49</td>
<td>0.69</td>
<td>0.97</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>RMS</td>
<td>0.72</td>
<td>0.33</td>
<td>0.62</td>
<td>1.02</td>
<td>0.63</td>
<td>1.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phase 2, delineation guidelines and digital atlas</th>
<th>Total CTV</th>
<th>Posterior</th>
<th>Left</th>
<th>Anterior</th>
<th>Right</th>
<th>Sphincter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>0.49</td>
<td>0.22</td>
<td>0.44</td>
<td>0.93</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>Patient 2</td>
<td>0.44</td>
<td>0.29</td>
<td>0.46</td>
<td>0.47</td>
<td>0.49</td>
<td>0.50</td>
</tr>
<tr>
<td>Patient 3</td>
<td>0.42</td>
<td>0.18</td>
<td>0.48</td>
<td>0.56</td>
<td>0.42</td>
<td>0.51</td>
</tr>
<tr>
<td>Patient 4</td>
<td>0.38</td>
<td>0.22</td>
<td>0.43</td>
<td>0.57</td>
<td>0.43</td>
<td>0.33</td>
</tr>
<tr>
<td>Patient 5</td>
<td>0.58</td>
<td>0.31</td>
<td>0.64</td>
<td>0.91</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Patient 6</td>
<td>0.40</td>
<td>0.20</td>
<td>0.46</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Patient 7</td>
<td>0.54</td>
<td>0.26</td>
<td>0.67</td>
<td>0.77</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>Patient 8</td>
<td>0.64</td>
<td>0.32</td>
<td>0.71</td>
<td>0.99</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>RMS</td>
<td>0.49</td>
<td>0.25</td>
<td>0.55</td>
<td>0.74</td>
<td>0.53</td>
<td>0.46</td>
</tr>
</tbody>
</table>
Phase 2

Using the guidelines and the atlas the 1st patient delineation lasted on average 63 min per observer (1SD 34.9). Delineation of patients 2 to 8 lasted 27 min (1SD 16.9) per observer. The digital atlas was consulted during 31 of the 86 delineations with a wide variation: 2 observers did not open the atlas at all contrarily to 2 observers using the atlas for 7 out of 8 patients. Usage of the coronal and sagittal representation of CT scan and MR scans was comparable to phase 1.

We found a statistically significant reduction in the delineated volumes (Table 4.1), with on average 160 cc. In APR patients, variation between observers reduced from 139 cc (1SD) to 90 cc SD (p = 0.01), comparable to the volume in the LAR patient group. The large variation in the most caudal slice in phase 1 was significantly reduced from 1.8 cm to 1.2 cm SD (p = 0.01), while for the cranial slice there was no difference. The CI increased on average from 0.63 to 0.66, which was borderline significant (Table 4.1).

![Fig. 4.3: Left-coronal view of the local surface distance variation (1SD) over the 11 observers projected on the median surface of the 8 patients in phase 1 (left) and phase 2 (right) of the study. The white contours indicate the bladder for reference.](image-url)
Target volume delineation variation for early stage rectal cancer in the Netherlands

Improvement in delineation variation can best be appreciated using the local surface distance variation (Fig. 4.3). Note the differences in CTV shape between the median surfaces from phase 1 and phase 2. Delineation variation especially reduced in the anterior region (reduction of 0.28 cm SD) and the sphincter region (reduction of 0.76 cm SD) (Table 4.2, phase 2).

Improvement in consensus between observers is well illustrated by the images of patient 2 (Fig. 4.1). Distances between observer delineations reduced for some regions from centimeters in phase 1 to the order of millimeters in phase 2. For patients 7 and 8, the observer variation did not change globally. For both patients the anterior part adjacent to the prostate/bladder was delineated more consistently in phase 2, but at the anterior border of the iliac lymph node regions and the anterior border adjacent to the small bowel a variation increase was found (Fig. 4.2). Both patients were male, designated to undergo a LAR.

Discussion

In this study we have shown that target volume delineation variation is a major geometric uncertainty for early-stage rectal cancer radiotherapy. National consensus guidelines and a digital delineation atlas were established and resulted in a significant reduction of the delineation variation.

In the analysis of CTV delineations, generally no golden standard exists. Reduction of delineation variation is therefore the goal, since it has an impact on the evaluation precision of trial results, tumor control and side effects.

With the large delineation variation known from both the literature [6] and the first phase of the current study, the question arises what size of delineation variation reduction should be reached for consensus. In this study we show that by using a delineation atlas and organizing a consensus meeting a large step forward could be made. Increased consensus was proven by the significant reduction in volume variation, caudal slice variation and local surface distance variation. Besides smaller differences between the observers within patients, the median surfaces between patients assigned to similar surgery were also more comparable in shape (Fig. 4.3). As an additional effect of the consensus guidelines, the average delineated CTV volume significantly reduced by 160 cc. This reduction was mainly achieved because in the first round some observers included parts of the cervix, uterus, bladder, prostate or seminal vesicles in the CTV (Figs. 4.1 and 4.2), while these were explicitly excluded in the developed guidelines. Apparently delineation guidelines and an atlas are needed to exclude the normal tissues from the CTV, as was already shown in the example images of the pilot study by Fuller et al. [6].

Target volume delineation is generally only performed once per patient and has impact on every treatment fraction. Delineation errors can therefore be considered as a systematic error. For clinical implications of the found and reduced delineation variation the consequences for the small bowel should be evaluated, being the major organ at risk [12, 13]. When using IMRT, the dose to the small bowel is mainly dependent on the required PTV margin at the anterior region. Delineation variation at the anterior region reduced from 1.02 cm SD in phase 1 to 0.74 cm SD in phase 2. Other sources of
systematic geometric uncertainties are target volume shape variation in the order of 0.7 cm SD [7, 8] and setup errors in the order of 0.2 cm SD [7]. Delineation variation in phase 1 was larger than any other error, and therefore unacceptably large. The reduction in delineation variation resulted in errors comparable to shape variation errors, making it a relevant reduction, which is open for further improvement.

There were no apparent differences in delineation variation when comparing male and female patients. The patients with a low seated tumor benefitted more from the delineation guidelines, since only in these patients the sphincter region is part of the CTV. In general we can conclude that delineation variation in early-stage rectal cancer is mainly dependent on the complexity of defining the CTV borders in the pelvic area, and not by gender or tumor location.

When evaluating the delineation variation in sub-regions, for the majority of patients substantial residual variation remains at the anterior border of the CTV cranial to the bladder and at the anterior border of the lateral lymph node regions (Fig. 4.3). For the anterior border of the lateral lymph node regions it is hard to define anatomic landmarks which can be easily recognized by all observers at all slices for all patients. For the anterior-cranial border of the CTV the delineation guidelines were interpreted differently amongst the observers, especially when small bowel loops were found adjacently (Fig. 4.2). Further improvement of the delineation variation might be reached if the delineation guidelines are adapted to be more explicit. Observer could, for example, first delineate both ureters and use them as anterior border for the lateral lymph node regions.

Recent publications on patterns of local recurrence and its reflection on CTV definition have indicated that the cranial border of the CTV might be lowered to the level of the S2-S3 interspace if a patient has an expected negative CRM and clinical negative lymph nodes [14, 15]. This adaptation of the CTV definition would result in omitting a large part of the CTV region adjacent to small bowel, consequently reducing residual delineation variation in that specific region. However, this adaptation was not taken into account in the current study, since the publication [15] was not yet available at the consensus meeting.

The guidelines in the current study were mainly based on the study of Roels et al. [4], which was chosen because these guidelines were based on local recurrences, distinctive anatomical borders and were widely accepted (more than 39 citations). Although most observers in the current study were familiar with the study of Roels et al. [4], large delineation variation in the first phase of the current study indicated that this was not enough for an acceptable level of delineation variation. Therefore, we found it necessary to make a point-by-point description of the delineation guidelines adjusted for early stage rectal cancer only (Appendix 4.1). Main differences with the proposed guidelines of Roels et al. can be found in the division in the different regions, but in general the same target volume is described. Small differences can be found in the amount of included fat around the sphincter region, the ventral border of the obturatorial lymph node regions and the caudal border of the internal iliac lymph node regions.
Larger differences exist when comparing the current delineation guidelines to the RTOG guidelines [5]. In the RTOG atlas the entire mesorectum down to the pelvic floor is always included, where in the current study 4 cm below the primary tumor is demanded. The sphincter region is, however, only included in the CTV when there is radiological proof of ischiorectal fossa invasion, while the current study includes this region for all patients designated for an abdominoperineal resection. The posterior obturator lymph node regions are always included, where the current study they are only included when the tumor is <10 cm from the anal verge. Finally, day-to-day variation in bladder volume and small bowel loops close to the CTV was recommended to be taken into account in the CTV by the RTOG, while in the current study the PTV should cover the variations.

A novelty in the current study is the use of a digital delineation atlas. Most published delineation guidelines, also for other cancer sites, are paper based [4, 5, 16, 17], preventing the user to approach the atlas in a similar way as during delineation of target volumes. Ideally the efficacy of a digital atlas should have been compared to a paper based atlas to establish the additional value. The current study took, however, already 16 months for 2 rounds and only 11 of the 17 participating radiation oncologists completed the whole assignment.

The digital atlas was viewed only during 31 of the total 86 delineations of the 2nd round. Usage was varying among the observers from never to almost always. For observers who viewed the atlas during 2 or more patients the average conformity index with the median surface significantly increased (0.74 vs. 0.79; p = 0.01), indicating that these observers delineated closer to each other in phase 2. Observers who viewed the atlas never or just once did not improve in conformity index with the median surface (0.72 vs. 0.73; p = 0.94). Further investigation might therefore be needed to improve the atlas use frequency and efficacy.

The only other study describing CTV delineation variation in rectal cancer is the pilot study of Fuller et al. [6]. In their study a single T3N0M0 case was delineated according to the RTOG atlas [5]. A direct comparison to our study is complex, since the RTOG atlas requires a different CTV, as described above. Although different delineation guidelines were used, their study also suggested that delineation variation is one of the major uncertainties compared to other geometric uncertainties in rectal cancer radiotherapy. In addition, they also suggested that delineation guidelines and an atlas can be used to reduce delineation variation.

Since the delineation variation in phase 2 was still a major source of geometric uncertainty, a further reduction of this variation is preferred. Besides re-evaluation of the defined guidelines, addition of multiple modalities might reduce delineation variation, as has been proven for other treatment sites [18, 19]. There is one study available in rectal cancer target volume delineation showing FDG-PET use resulted in reduced variation for delineation of the GTV [20]. However, for the CTV, mainly an elective target volume, the FDG-PET has no additional value.
Magnetic resonance images are currently used for rectal cancer staging, determining the surgical procedure, assessing the circumferential resection margin and the nodal status [21, 22]. For target volume delineation MR scans are used less often, while it is assumed to be superior for localization of the mesorectal fascia interface and the presacral space [4]. For delineation of the GTV in low rectal cancer MRI has been shown to be superior to CT delineation [23]. To enable CTV delineation on MR for treatment planning, image fusion with the planning CT is needed, since most treatment planning systems still require a CT scan. Problems could occur when large anatomic differences exist between both scans. In the current study no co-registration between the MR and CT scans was provided due to the anatomical differences. Benefit and disadvantages of MR scan usage was therefore limited in the current study. With the superior soft-tissue contrast in MR images we estimate a large potential benefit for addition of co-registered images during target volume delineation.

**Conclusions**

Target volume delineation is generally only performed once per patient and has therefore impact on every treatment fraction. Large variation was found in clinical target volume delineation between radiation oncologists for pre-operative radiotherapy of rectal cancer patients. In this study a reduction of delineation variation was achieved by establishing national consensus guidelines and a digital delineation atlas. This resulted in delineation variation of 0.74, 0.54, 0.25 and 0.46 cm SD for, respectively, the anterior, lateral, posterior and the sphincter region. Resulting errors were comparable to the other major systematic geometric uncertainty, shape variation of the clinical target volume. With resulting delineation variation still being one of the major geometric uncertainties there is room for further improvement.
References


Chapter 4


Appendix 4.1
Guidelines for clinical target volume (CTV) delineation in early stage rectal cancer (T1-3, N0-N1) to be treated with pre-operative 5 x 5 Gy RT

Besides the tumor and suspected/involved lymph nodes plus a margin, the CTV can be divided into 4 different areas: the lymph node area’s around the iliac vessels, obturator regions, the mesorectum and the pre-sacral region.

**Mesorectum:**
- caudal border: LAR: at least 4 cm caudal of the primary tumor  
  APR: anus/sphincter complex, with 1 cm circular margin
- cranial border: at the level where the rectum curves ventrally and forms the sigmoid
- ventral en dorsal border: mesorectal fascia (excluding small bowel)
- lateral border: mesorectal fascia

**Iliac lymph node regions (lateral left and right):**
- caudal border: the level where the internal iliac vessel crosses the piriform muscle
- cranial border: caudal side of the division of the vena iliaca communis
- ventral border: 7 mm ventral of the internal iliac vessel
- lateral border: pelvic wall muscles
- dorso-medial border: mesorectal border

**Obturator regions (lateral left and right):** only included when the primary tumor <10 cm from the anus
- caudal border: at the level where the ureter inserts the bladder
- cranial border: up until where the obturator muscle is visible
- ventral border: ureter, unless there is a large concentration of small vessels ventrally of the ureter, then these small vessels also need to be included
- lateral border: pelvic wall muscles
- dorso-medial border: mesorectal border, avoid seminal vesicles, uterus and vagina
- dorsal border: ventral border of the iliac lymph node region

**Pre-sacral region:** area between the iliac vessels and also drain area of the a. rectalis superior
- caudal border: cranial border of the mesorectum
- cranial border: cranial border of the iliac lymph node region
- ventral border: 2 cm ventral van het sacrum, measured perpendicular to the sacrum (superior rectal artery has to be included in this region)
- lateral border: iliac lymph node regions
- dorsal: pelvic wall muscles, neuro foramen not included

**GTV to CTV:**
- for tumor take a uniform margin of 1.0 cm
- for suspected or involved lymph nodes take a uniform margin of 0.5 cm