FDG-PET-CT reduces the interobserver variability in rectal tumor delineation

Citation for published version (APA):

Document status and date:
Published: 01/03/2012

DOI:

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:
repository@maastrichtuniversity.nl
providing details and we will investigate your claim.

Download date: 17 Sep. 2023
PET in dose planning
FDG–PET–CT reduces the interobserver variability in rectal tumor delineation

Jeroen Buijsen a,⁎, Jørgen van den Boogaard a, Hiska van der Weide a, Stephanie Engelsman b, Ruud van Stiphout a, Marco Janssen a, Geerard Beets c, Regina Beets-Tan b, Philippe Lambin a, Guido Lammering a

a Department of Radiation Oncology (MAASTRO), GROW - School for Oncology and Developmental Biology; b Department of Radiology; and c Department of Surgery, Maastricht University Medical Centre, The Netherlands

ARTICLE INFO

Article history:
Received 6 June 2011
Received in revised form 9 November 2011
Accepted 23 December 2011
Available online 24 January 2012

Keywords:
Rectal cancer
PET–CT
Contouring
Interobserver variation
Automatic contouring

ABSTRACT

Background and purpose: Previously, we showed a good correlation between pathology and an automatically generated PET-contour in rectal cancer. This study analyzed the effect of the use of PET–CT scan on the interobserver variation in GTV definition in rectal cancer and the influence of PET–CT on treatment volumes.

Materials and methods: Forty-two patients diagnosed with rectal cancer underwent an FDG–PET–CT for radiotherapy planning. An automatic contour was created on PET-scan using the source-to-background ratio. The GTV was delineated by 5 observers in 3 rounds: using CT and MRI, using CT, MRI and PET and using CT, MRI and PET auto-contour. GTV volumes were compared and concordance indices (CI) were calculated. Since the GTV is only a small portion of the treatment volume in rectal cancer, a separate analysis was performed to evaluate the influence of PET on the definition of the CTV used in daily clinical practice and the caudal extension of the treatment volumes.

Results: GTV volumes based on PET were significantly smaller. CIs increased significantly using PET and the best interobserver agreement was observed using PET auto-contours. Furthermore, we found that in up to 29% of patients the CTV based on PET extended outside the CTV used in clinical practice. The caudal border of the treatment volume can be tailored using PET-scan in low seated tumors. Influence of PET on the position of the caudal border was most pronounced in high seated tumors.

Conclusion: PET–CT increases the interobserver agreement in the GTV definition in rectal cancer, helps to avoid geographical misses and allows tailoring the caudal border of the treatment volume.

© 2012 Elsevier Ireland Ltd. All rights reserved. Radiotherapy and Oncology 102 (2012) 371–376

Pre-operative radiotherapy has become an essential part of the treatment of most patients with rectal cancer, since it is very effective in reducing the risk of a locoregional recurrence [1]. Combining radiotherapy and chemotherapy results in downstaging and in up to 10–33% of patients pathological complete responses (pCR) have been reported [2–5]. Patients with a good clinical response may benefit from less invasive surgery, like sphincter-saving surgery or transanal endoscopic microsurgery (TEM). Even, in selected cases, a wait-and-see policy might be safe [6]. These innovative modified surgical approaches may lead to a better quality of life. Thus, it would be attractive to further increase the probability of a good tumor response, e.g. by increasing the dose to the tumor [5,7,8]. To achieve higher doses to the tumor, a simultaneous integrated boost technique has been shown to be feasible [9–11]. In order to identify the boost volume easily and reliably, high quality imaging is important. MRI is considered the most accurate staging method for rectal cancer [12–14], but its role in a precise determination of the boost volume is unknown [15]. For PET-imaging it has been reported that it is reliable in defining tumor size in rectal cancer [16] and has the additional advantage that it can easily be acquired in treatment position simultaneously with a CT-scan, which is needed for treatment planning.

Another reason why it is important to define more precisely the primary tumor in rectal cancer is that it may help to reduce long term toxicity that is observed after radiotherapy for rectal cancer [17,18], through a further reduction of treatment fields. As our group has shown before, 3-D conformal planning results in a better PTV coverage and dose homogeneity as compared to standard 3- or 4-field techniques based on bony anatomy [19]. However, the use of more conformal techniques, like IMRT, poses the risk of geographical misses. Furthermore individual delineation makes it possible to better spare normal tissues. A better identification of the primary tumor allows avoiding irradiation of the sphincter in selected cases and reducing the volume of small bowel in high seated tumors, resulting in less toxicity. It has been shown that positive lymph nodes are most frequently located at the level of the tumor.
PET–CT for rectal cancer delineation

The proximal spread of lymph nodes is limited to 5 cm from the distal margin of the tumor and the distance to the most distal nodes is 4 cm at maximum [20,21]. Furthermore, an analysis of the Dutch TME trial showed that in primary resectable rectal cancer short course radiotherapy is especially effective in the prevention of anastomotic recurrences [22,23]. This finding confirms that it is important to know the exact location of the tumor in order to safely reduce radiotherapy treatment fields.

For these reasons we hypothesized that the use PET–CT can help to define the GTV in rectal cancer more accurate leading to a better tailored definition of the treatment volume, that it would diminish interobserver variability and diminish the time needed to define the GTV. Furthermore we hypothesized that the influence of the use of PET–CT on treatment volume would be larger in low seated tumors as compared to high rectal tumors, because of the low soft tissue contrast on CT in the lower part of the pelvis.

Methods and materials

For this study 42 patients diagnosed with rectal cancer (cT2-4N0-2M0) were selected. Patients were scheduled to undergo a neo-adjuvant treatment consisting of chemoradiotherapy (28 × 1.8 Gy with concurrent capecitabine 825 mg/m² bid). All patients underwent an FDG–PET–CT scan for radiotherapy planning on an integrated PET–CT scanner (Truepoint Biograph 40, Siemens Erlangen, Germany). The PET-scan protocol has been described in detail earlier [24]. PET–CT images were fused and an automatic contour around the primary tumor was created, using the Signal-to-Background-Ratio (SBR)-method as described earlier [25,26] using dedicated software (Esoft 5.0, Siemens MI, Erlangen, Germany). This contouring method has been shown to have a good correlation with pathology in rectal cancer [16].

We balanced the number of patients with low- and high-seated tumors (22 high, 20 low). For this study high seated tumors had a caudal border ≥7 cm from the anal verge.

For delineation, fixed window/level settings were used (400/50 for CT and 30000/15000 for PET). In PET-scans not showing enough contrast using these settings, an adjusted W/L setting was used, identical for all 5 observers.

GTV delineation and interobserver variability

The GTV was delineated by 5 observers: 2 radiation-oncologists sub-specialized in gastro-intestinal (GI) tumors, 1 senior-resident, 1 radiation-technologist and 1 radiologist. At the time of delineation clinical details were available and presented in a standardized format to each observer including the findings on digital rectal examination and the endoscopy and pelvic MR-imaging reports.

Each study set was delineated 3 times by each observer in 3 consecutive rounds. During each delineation round observers had, in addition to the standardized clinical information, access to different imaging information, creating 3 sets of GTV contours per observer. In each round an MR-scan was available. MR and CT were not fused and MR was projected on a second screen. Round 1: MR- and CT-images only (CT–GTV (GTVCT)), round 2: MR-, CT- and PET-images (PET–GTV (GTVPET)) and in round 3 in addition to the MR-, CT- and PET-images the automatic generated contour on PET was provided (automatic GTV (GTVauto)). In the third round observers were asked to edit the provided contour, in such a way to obtain a clinically acceptable GTV. For each observer, the delineation rounds were spaced with a minimum interval of 4 weeks, to prevent bias from a preceding delineation round. Observers were blinded to each other’s delineations.

Time needed for each delineation was registered by the observers. The volumes of the different GTVs were collected from the planning system and compared. Pairs of contours from the 3 different delineation methods were compared by calculating the concordance index (CI), defined as the ratio of the intersection and the union of the two volumes [27,28].

\[ CI = \frac{(A \cap B)}{(A \cup B)} \]

Differences in GTV delineation were analyzed for the total patient group as well as for high- and low-seated tumors separately.

CTV delineation and interobserver variability

Since the GTV is only a small part of the clinical target volume in the current treatment of rectal cancer, 3 observers also delineated the complete CTV as used in daily clinical practice (CTVcompl), including regional lymph nodes, according to our local protocol as described earlier [19]. In brief, the CTVcompl included at least 3 cm of the rectal wall in the oral and aboral directions, to cover possible intramural tumor spread, the mesorectal subsite, posterior pelvic subsite, and the regional lymph nodes at risk, which were defined by contouring the internal iliac vessels with a margin of 5 mm and the obturator region for low seated tumors (<7 cm from the anal verge in this protocol). The obturator region was delineated as proposed by Roels et al. [29]. The CTV of the primary tumor was obtained by circumferential expansion of the GTVs with 0.5 cm, resulting in CTVCT, CTVPET and CTVauto. The percentages of the different CTVs located outside the CTVcompl were analyzed.

We were particularly interested in the caudal tumor extension. There is evidence that the risk of microscopic disease in lymph nodes >4 cm caudal from the caudal border of the primary tumor is very limited [20,21]. Reducing the caudal CTV extension could result in a reduced radiation dose delivered to the anal sphincter, decreasing the risk of sphincter dysfunction as described in patients undergoing pre-operative radiotherapy followed by sphincter sparing surgery [18,30]. Therefore we also analyzed if the caudal border intramural margin differed between the CT- and PET-based delineations.

Statistics

SPSS 17.0 (SPSS, Chicago, IL) was used to perform statistical analysis. For the comparison of the differences in time needed to perform the delineations a paired-samples t-test was used. The volumes of the GTVs, CIs of both methods, as well as the percentages of CTV lying outside the CTVcompl and differences in caudal borders were compared using the Wilcoxon signed rank test, because data did not follow a normal distribution. Two-sided p-values are provided; p-values <0.05 were considered significant.

Results

The availability of PET images resulted in about 40% decrease in the time needed to complete GTV delineation (mean time GTVCT 4.1 min, GTVPET 2.5 min (p < 0.001) and GTVauto 1.6 min (p < 0.001)). The volumes for each observer are shown in Table 1. GTV volumes were significantly smaller using PET-scan (mean GTVCT 46.8 cm³ vs. mean GTVPET 28.8 cm³ (p < 0.001)). Editing automatically created contours resulted in the smallest volumes (mean GTVauto 18.2 cm³ (p < 0.001)).

An example of the delineation of 2 patients is depicted in Fig. 1. Conformity indices increased when PET information was added, reflecting a better agreement between observers. The mean conformity index (mean ± SD) for the 5 observers was 0.79 ± 0.17 (range:
0.82 ± 0.16 (range: 0.10–1.00, \(p = 0.103\)) using PET-data without automatically created contours and 0.93% ± 10.5 (range: 0.28–1.00, \(p < 0.001\)) using PET auto-contours (Fig. 2). Using CT-scans only, in 2 cases a complete disagreement between observers occurred (reflected by a CI of 0).

No differences were found between low- and high-seated tumors: 0.78 vs 0.79 (\(p = 0.31\)) for CT-only, 0.82 vs. 0.82 (\(p = 0.50\)) for PET-manual and 0.93 vs 0.92 (\(p = 0.94\)) for PET with auto-contours (Fig. 2).

The analysis of the CTVs showed that, with the addition of PET, in some patients a part of the tumor CTV was not covered by the CTV_{compl}. The mean CI for \(CTV_{PET}\) was 0.98 (range: 0.27–1.00) and for \(CTV_{auto}\) 0.98 (0.29–1.00). For \(CTV_{PET}\) the percentage of volume lying outside \(CTV_{compl}\) exceeded 5% in 8 cases (19%) (4 times in observer 1, 2 times in observer 2 and 2 times in observer 3). For \(CTV_{auto}\) this was the case in 12 delineations (29%) (6 times in observer 1, 3 times in observer 2 and 2 times in observer 3). In 1 patient more than 75% of the \(CTV_{PET}\) and \(CTV_{auto}\) was lying outside the \(CTV_{compl}\) in 2 observers. This was a high seated tumor which was not correctly delineated based on CT and MR only, but was correctly identified on PET.

On average, the caudal border of the intramural margin was located 0.6 cm more cranial if based on automatic PET-contours as compared to CT. In 7 patients (17%), the caudal border of the PET-based margin was located 1 cm or more caudal than the CT-based margin for at least one observer. Four of them had a high seated tumor, 3 had a low seated tumor. In 3 of these patients the intramural margin extended \(\geq 1\) cm caudally based on PET.

### Table 1

Mean GTV volumes for the 5 different observers using three different delineation methods.

<table>
<thead>
<tr>
<th>Observer</th>
<th>CT-based Mean</th>
<th>Range</th>
<th>PET-based Mean</th>
<th>Range</th>
<th>PET-auto Mean</th>
<th>Range</th>
<th>(p^a)</th>
<th>(p^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>46.8</td>
<td>6.3</td>
<td>185.7</td>
<td>28.8</td>
<td>113.9</td>
<td></td>
<td>(&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>Observer 1</td>
<td>41.2</td>
<td>6.3</td>
<td>159.3</td>
<td>26.0</td>
<td>107.3</td>
<td></td>
<td>(&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>Observer 2</td>
<td>52.7</td>
<td>11.6</td>
<td>168.5</td>
<td>26.2</td>
<td>115.3</td>
<td></td>
<td>(&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>Observer 3</td>
<td>48.8</td>
<td>9.6</td>
<td>185.7</td>
<td>37.2</td>
<td>131.9</td>
<td></td>
<td>(&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>Observer 4</td>
<td>43.6</td>
<td>7.3</td>
<td>166.6</td>
<td>29.6</td>
<td>119.7</td>
<td></td>
<td>(&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>Observer 5</td>
<td>47.4</td>
<td>7.9</td>
<td>167.5</td>
<td>25.2</td>
<td>128.7</td>
<td></td>
<td>(&lt;0.001)</td>
<td></td>
</tr>
</tbody>
</table>

\( \text{CT-based PET-based PET-auto} \)

\( \text{Mean Range Mean Range Mean Range} \)

\( p^a \) Compared with CT-based delineation.

\( p^b \) Compared with PET-based delineation.

### Fig. 1

Example of the delineations by 5 observers in a patient with a high-seated (1) and a low-seated tumor (2), based on CT-only (a), PET–CT (b) and PET–CT with auto-contour (c).
for all 3 observers. In 19 (45%) patients the caudal border of the PET-margin was located \( \geq 1 \) cm higher than the CT-margin. In 6 patients this difference of \( \geq 1 \) cm was observed in all 3 observers, in another 6 patients in 2 observers and in the remaining 7 patients in 1 observer. Sixteen of these tumors were located high and 3 were located low. On average the caudal border of PET-margin of high-seated tumors was located 1.1 cm more cranial than CT-margin. For low seated tumors the difference was negligible (Fig. 3). The difference between low- and high-seated tumors was statistically significant.

**Discussion**

To the best of our knowledge, this is the first study looking at the influence of the use of automatically created PET-contours on GTV delineation and interobserver variability in rectal cancer. As has been shown earlier, the SBR method results in contours with a very good correlation with pathology [16]. The smallest GTV volumes are created using PET-based autocontours as compared to CT-based contours and manual PET-based contours (23.6 cc vs. 28.8 cc vs. 46.9 cc). This study confirms that the use of PET-based autocontours leads to a very good interobserver agreement, reflected by a mean CI of 0.93 for automatic PET-contours as compared to 0.82 (manual PET-contours) and 0.79 (CT-based contours). Furthermore, we found that PET-scan may help to avoid geographical misses in selected cases, especially in very low and very high located tumors and therefore may be helpful to define an adequate boost volume. In addition it could help to reduce treatment fields leading to a reduction of the amount of sphincter in the radiation volume, possibly leading to less late sphincter related toxicity and it allows for GTV-boosting in dose escalation trials.

A better interobserver agreement using PET-scan has been shown for the delineation of lung-, brain- and head and neck tumors [31–33]. Furthermore it has been demonstrated that the use of automatically created contours results in better agreement than visual interpretation of PET [34]. Obviously, visual interpretation depends on many factors, like window/level settings and the display mode of the PET information (for example grayscale or colormap) and experience of the observers. In this study 5 observers with different experience and background were asked to delineate the GTVs. Although it is difficult to draw firm conclusions, we observed that the differences in volumes differed statistically significant for all but 2 observer combinations when using CT data only. Differences in volume became smaller using PET-data and CIs decreased for all observers, indicating that PET is helpful for observers with different levels of experience in delineation of rectal tumors.

Other groups have looked at the influence of PET on target volume delineation before. Ciernik et al. [35] was the first group that published about the use of PET-CT in delineation for rectal cancer using a growing region algorithm. They found a good correlation with GTVs created manually on CT, but did not analyze the influence of PET information on delineation decisions by the physician. No comparison was made between different observers. In a study of Bassi et al. [36] tumor delineation was done by 2 radiation-oncologists together. They concluded that GTVs based on PET were significantly smaller than CT-based GTVs, which is in line with our findings. Patel et al. compared the delineations of tumor and lymph nodes in 6 rectal cancer patients who underwent an FDG–PET as well as FLT–PET [37]. In contrast to our findings, they did not observe clear differences in GTV volumes, but the interobserver agreement was better using PET. No differences between FDG–PET and FLT–PET were seen and no SUV-based auto-contour was used.

Results of rectal cancer treatment have improved markedly in the last decades, due to better surgical techniques and the widespread use of radiotherapy [1]. However, the use of radiotherapy results in long-term toxicity in a substantial part of patients. Therefore, it is important to make a better patient selection on one hand and to tailor treatment fields as much as possible on the other hand. Irradiation of the sphincter may result in problems with fecal continence [38,39]. In the Dutch TME trial 62% of patients without a stoma reported fecal incontinence or soiling in the radiotherapy and surgery arm versus 38% in the surgery only arm [18]. This study shows that the use of PET-scan can help to better tailor treatment fields. Especially the caudal border of the radiation fields can be limited in an important proportion of patients, resulting in a lower sphincter dose and a lower dose to the distal rectal wall in higher seated tumors, without the risk of geographical misses. Theoretically this could lead to less late toxicity. We hypothesized that the influence of PET on GTV volume and caudal extension of the treatment volume would be largest in lower seated tumors. This study showed that influence of PET was not different between low- and high-seated tumors. Therefore, the use of PET may be beneficial to all rectal cancer patients. When looking specifically at the influence of imaging modality on the position of the caudal border of the treatment volume (Fig. 3), it can be concluded that PET is helpful to individualize this border. In some patients this border will be located more cranial when based on PET, possibly resulting in a lower dose to the sphincter and distal rectal wall, while in others it will be located more distal. For the analysis of the caudal border we included a 3 cm margin in the course of the rectal wall in both directions. In our protocol we do not stop this margin at the border of the sphincter. For example: if the caudal border of the tumor is located 2 cm cranial to the sphincter, 1 cm of the anal canal was included in the intramural margin. It can be argued whether this is really necessary or whether one could see the anorectal junction as an anatomical barrier. Of course this can influence the results of our analysis. Recent literature suggests that distal surgical margins as close as 1 cm may be safe, but this is based on the surgical data of patients who have been treated with pre-operative radiotherapy in majority [40,41]. Therefore, we do not feel comfortable at the moment to leave the complete sphincter out of the treatment volume in very low lying tumors and use the intramural margin based on PET-scan.

As stated in the introduction, adequate identification of the tumor is essential to create a reliable boost volume. This allows to study if boosting of the primary tumor results in more pathological

---

**Fig. 3.** Mean difference ± SD in cm of the 3 observers who delineated the complete treatment volume between the most caudal extension of the intramural margin based on CT and the caudal boundary based on the automatic PET-based contour. A negative value means that the caudal border based on PET was located more cranial than the CT-based border (i.e. in that case the volume based on PET was shorter in caudal direction).
complete responses and if this can lead to the use of less invasive surgery. This study shows that use of PET results in a good agreement between observers and our pathology validation study showed a very strong agreement between tumor length defined by automatic PET-contours and measured by the pathologist in the surgical specimen. If we assume that the representation of the position of the tumor and tumor edges is accurate, we can conclude that PET-CT makes it possible to define a reliable GTV in rectal cancer. Although we think that this assumption is very plausible, no analysis of the position of the tumor on PET–CT and in vivo has been performed. Therefore, in clinical practice this method should be used with caution and a clinical prospective evaluation is necessary.

A second problem that has to be solved to define an adequate boost volume is the internal organ motion, which can be quite substantial in the case of rectal cancer. It has been shown that especially in the cranial part of the mesorectum deformations can be quite substantial and that these deformations are caused mainly by differences in bowel filling [42]. Although a PET-scan adequately images the primary tumor and can be used for tumor delineation, it is not reliable for the distinction between benign and pathological lymph nodes [43]. The specificity is acceptable, but the sensitivity is rather low [44–46]. Therefore, additional imaging is strongly needed to adequately identify positive nodes. MR in combination with special contrast agents seems to be a promising method [47]. For this study we did not compare different PET segmentation algorithms, because we found a good correlation between pathology and SBR-based PET-contours. However, the SBR-method has several disadvantages [48]. It is dependent on many parameters, so that each modification in the process makes it necessary to perform a new calibration and each scanner has to be calibrated separately. In addition this method does not perform well if the source-to-background ratio is low. In future projects we will compare the performance of other segmentation methods with the SBR method in rectal cancer. Apart from a useful tool in delineation, FDG–PET can also be helpful to get insight in tumor heterogeneity. Our group has shown that in NSCLC residual metabolic active areas after radiotherapy are the areas with the highest uptake before treatment [49,50]. Recently we showed that this is also the case for rectal cancer [51]. Another step for the future could therefore be the development of sub-boosting techniques.

In conclusion, PET–CT reduces interobserver variation and volumes in GTV definition in rectal cancer, enables tailoring treatment fields, especially in cranio-caudal direction, and makes it possible to define volumes for boosting.

References

PET–CT for rectal cancer delineation


