Physics

Dosimetric impact of contouring and image registration variability on dynamic $^{125}$I prostate brachytherapy

Hendrik Westendorp$^{1,3,4}$, Kathrin Surmann$^{1,2,4}$, Sandrine M.G. van de Pol$^1$, Carel J. Hoekstra$^1$, Robert A.J. Kattevilder$^1$, Tonnis T. Nuver$^1$, Marinus A. Moerland$^3$, Cornelis H. Slump$^2$, André W. Minken$^1$

$^1$Radiotherapiegroep behandellocatie Deventer, Deventer, The Netherlands
$^2$MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, The Netherlands
$^3$Department of Radiation Oncology, University Medical Center Utrecht, Utrecht, The Netherlands

ABSTRACT

PURPOSE: The quality of permanent prostate brachytherapy can be increased by addition of imaging modalities in the intraoperative procedure. This addition involves image registration, which inherently has inter- and intraobserver variabilities. We sought to quantify the inter- and intraobserver variabilities in geometry and dosimetry for contouring and image registration and analyze the results for our dynamic $^{125}$I brachytherapy procedure.

METHODS AND MATERIALS: Five observers contoured 11 transrectal ultrasound (TRUS) data sets three times and 11 CT data sets one time. The observers registered 11 TRUS and MRI data sets to cone beam CT (CBCT) using fiducial gold markers. Geometrical and dosimetrical inter- and intraobserver variabilities were assessed. For the contouring study, structures were subdivided into three parts along the craniocaudal axis.

RESULTS: We analyzed 165 observations. Interobserver geometrical variability for prostate was 1.1 mm, resulting in a dosimetric variability of 1.6% for $V_{100}$ and 9.3% for $D_{90}$. The geometric intraobserver variability was 0.6 mm with a $V_{100}$ of 0.7% and $D_{90}$ of 1.1%. TRUS–CBCT registration showed an interobserver variability in $V_{100}$ of 2.0% and $D_{90}$ of 3.1%. Intraobserver variabilities were 0.9% and 1.6%, respectively. For MRI–CBCT registration, $V_{100}$ and $D_{90}$ were 1.3% and 2.1%. Intraobserver variabilities were 0.7% and 1.1% for the same.

CONCLUSIONS: Prostate dosimetry is affected by interobserver contouring and registration variability. The observed variability is smaller than underdosages that are adapted during our dynamic brachytherapy procedure. © 2017 American Brachytherapy Society. Published by Elsevier Inc. All rights reserved.

Keywords: Low-dose rate; Prostate; Brachytherapy; Registration; Contouring; Variability

Introduction

$^{125}$I prostate brachytherapy depends heavily on imaging. Intraoperative visual feedback for contouring and seed deposition is commonly provided by transrectal ultrasound (TRUS) (1). The quality of permanent prostate implants can be improved using additional imaging modalities, like MRI and cone beam CT (CBCT) (2).

With MRI, intraprostatic structures and lesions can be identified (3–5), allowing for boosting of subvolumes (6) or focal treatments (7–9). Intraoperative CBCT enables more accurate localization of the deposited seeds (10–13). When imaging modalities are added to the implantation procedure, a registration to the primary (TRUS) data set needs to be performed. This registration requires manual interaction, resulting in additional inaccuracies because of inter- and intraobserver variabilities. This variability directly affects dosimetry.

In addition to registration, contouring is another major source of variability (14). Interobserver variability depends

Received 30 November 2016; received in revised form 23 January 2017; accepted 24 January 2017.

Conflicts of interest: None declared.

* Corresponding author. Radiotherapiegroep behandellocatie Deventer, Nico Bolkesteinlaan 85, 7416 SE, Deventer, The Netherlands. Tel.: +31-570-646900; fax: +31-570-646901.

E-mail address: r.westendorp@radiotherapiegroep.nl (H. Westendorp).

4 Both authors contributed equally to this article.

1538-4721/S - see front matter © 2017 American Brachytherapy Society. Published by Elsevier Inc. All rights reserved.

http://dx.doi.org/10.1016/j.brachy.2017.01.010
on the imaging modality that is used for contouring (15). Contouring on CT results in more variability than contouring on TRUS and MRI (15). Prostate outer contours show similar variability on TRUS and MRI (15, 16).

De Brabandere et al. (14) studied the dosimetric impact of interobserver variability in seed localization, contouring, and image registration in a multicenter setting. Eight observers contoured and registered postimplant CT and MRI for 3 patients. \(D_{90}\) showed a large dosimetric variability in contouring (17–23%) and image registration (6–16%). In this multicenter study, the authors suggested that personal or institutional habits could have played a major role in the large contouring variability. Training would probably help to lower contouring variability, whereas registration variability could possibly be reduced using automated tools.

When adding imaging modalities to an implantation procedure, the benefit of improved accuracy of contouring should outweigh the additional uncertainties in dosimetry that are caused by the registration itself. Ideally, the registration variability should be small compared with other sources of uncertainty, like contouring and seed localization.

Since 2006, we routinely apply a dynamic dosimetry technique. After implantation, we acquire a C-arm CBCT scan and register it to a contoured TRUS image data set (12, 17). In a previous study, we concluded that TRUS-CBCT-based dosimetry enables identification of underdosed regions that need adaptation by placing remedial seeds (17).

In the present study, we assessed the inter- and intraobserver variabilities in contouring and registration for our dynamic implantation procedure. Anticipating the implementation of pretreatment MRI in our procedure, to enable focal treatments, we additionally incorporated the registration of MRI with CBCT. The purpose of this study was to quantify the dosimetric variability caused by contouring and registration and to compare this variability with the dosimetric improvements made by our dynamic dosimetry technique.

Methods and materials

Patients

Data from 11 prostate cancer patients, treated with a boost with stranded \(^{125}\)I seeds after external beam radiation therapy (EBRT), were used. Each patient had received 47 Gy (20 × 2.35 Gy) with EBRT and 110 Gy as a brachytherapy boost.

Treatment procedure

Before EBRT treatment, four gold fiducial markers (Heraeus GmbH, Hanau, Germany) were implanted for EBRT position verification and image registration. For EBRT, contouring of regions of interest was performed on a 2-mm-thick sliced CT data set (Brilliance Big Bore 16 Slice; Philips, Best, The Netherlands) registered with T1- and T2-weighted MRI data sets with a slice spacing of 2 mm (Signa HDxt; GE Medical, Milwaukie, WI). The MRI data set was acquired ~7 weeks before the implantation procedure.

Approximately 2 weeks after finishing EBRT, the brachytherapy procedure was performed. We used a dynamic dosimetry implantation procedure (12, 13, 17). The implantation procedure started with the acquisition of a TRUS scan (FlexFocus 400; BK Medical, Herlev, Denmark) with 5-mm spaced slices that were contoured. Immediately after finishing implantation, a second TRUS scan was acquired with 2.5-mm spaced slices on which the regions of interest were contoured. This TRUS data set was registered to a 2.5-mm-thick sliced CBCT data set (Siemens Arcadis Orbic 3D; Siemens Medical Systems, Erlangen, Germany), on which implanted seeds were identified. The CBCT data set was obtained directly after the TRUS scan. The registration was started using the least-squares method in the treatment planning system (TPS; Variseed 8.0.2; Varian Medical Systems Inc, Palo Alto, CA) using the fiducial markers and manually adjusted if deemed necessary. The hemispherical shape of the ends and the larger diameter (1.0 mm) improve visibility of fiducial markers considerably compared with \(^{125}\)I seeds. A month after implantation, a CT data set with 2-mm-thick slices (CT 30) was acquired to assess the Day 30 dose distribution. More details of the procedure have been published previously (12, 17).

Multiobserver study

Five observers from our institute participated in the inter- and intraobserver variability studies. The group of observers consisted of two experienced radiation oncologists who routinely perform the implantation procedure (CH and SP), a medical physicist with long-term expertise (HW), a dedicated research brachytherapy technologist (RK), and the primary investigator (KS, trained by SP).

The observers contoured the prostate, urethra, and rectum three times on the pre-implant TRUS. For comparison, the prostate and rectum were contoured once on CT 30. During the acquisition of CT 30, there was no catheter present to allow contouring of the urethra. To minimize bias, observers could not review their previous sessions. The sessions were at least 1 week apart.

Observers used the TPS to contour and register the data sets for all sessions. The data were exported in Dicom format and processed using MATLAB (version 8.1.0.604; The MathWorks Inc., Natick, MA).

TRUS–CBCT registrations were performed three times by each observer. To assess the impact of anticipated implementation of MRI, we additionally performed three MRI–CBCT registrations. The observers were instructed to start the registration based on all fiducial markers. The least-squares fit (points)
registration algorithm of the TPS was used. After that, if needed, the registration was manually adjusted, guided by the visible seeds. After each registration, dosimetric parameters were recorded. For the prostate, \( V_{100} \) and \( D_{250} \) were reported, for the urethra, we recorded \( D_{30} \), and for the rectum \( V_{100} \).

Fiducial markers were always implanted at four fixed locations: cranial left, cranial right, caudal left, and caudal right, using two needles. If a fiducial marker could not be identified on the image data set, the anatomical location (e.g. cranial left) of the missing fiducial marker was reported.

Two experienced radiation oncologists had reviewed the clinically used (TRUS) contours and both approved the result during the implantation procedure. These clinically used contours were regarded as the reference data set for the interobserver variability studies. Clinical intraoperative (TRUS—CBCT) and Day 30 (TRUS—CT) dosimetry were set as reference, both for the contouring and the registration study. CT contours were compared with the TRUS reference contours.

We contoured TRUS multiple times and assessed geometrical and dosimetric variabilities. One session of CT contouring showed that interobserver variability in CT contouring was much higher than in TRUS. Therefore, we decided to restrict the intraobserver variability study to TRUS contouring.

Contours were analyzed with a Matlab script. The prostate surface was sampled with 10° increments of the polar and azimuthal angles (15). The center of mass of the reference prostate was set to the origin (0, 0, 0), and all observations were relative to this center. For the urethra, the in-plane center was calculated for each slice. Urethra contours were analyzed only on slices with prostate contours. The rectum was resampled around the center of the TRUS probe for each slice. Rectum contours were assessed over the length of the prostate under an angle of 25° left and right of the midline with an increment of 10°. A schematic drawing of the acquisition of the sample points is depicted in Fig. 1. Each structure from the reference data set was divided into three equal parts along the craniocaudal axis: superior, central, and inferior.

The distance between the actual and the reference contour determined the geometrical interobserver variability. The distance between the actual contour and the mean of the three contouring sessions determined the intraobserver variability. The distance between the reference contours (inter) or mean of the three contouring sessions (intra) was calculated for each sample point. For each structure or part of a structure (superior, central, and inferior), these sample points were averaged. Next, variabilities were calculated (Supplementary Data).

The dosimetric contouring parameters were determined by the actual contours in the reference dose distribution. Dosimetric parameters were normalized to the reference (clinical situation for interobserver and average for intraobserver). The clinically obtained parameters served as references for the interobserver variability. For the intraobserver variability, the mean of three sessions served as reference. The variability was reported as the SDs of the observations (\( \text{SD}_{\text{inter}} \) and \( \text{SD}_{\text{intra}} \), Supplementary Data).

The registration study was conducted similarly as the contouring study. Reference contours and dose distributions were used, leaving the registration the only variable. The TPS did not report registration parameters (i.e. rotations and translations). We recorded dosimetric parameters only.

**Results**

**Contouring variability**

The interobserver variability in contouring (1 SD) for the whole prostate on TRUS was 1.1 mm and splitting the prostate in three equal parts along the craniocaudal axis, 2.1 mm for the superior, 0.4 mm for the central, and 2.0 mm for the inferior part. The intraobserver contouring variability (1 SD) was 0.6 mm for the whole structure and 0.9, 0.4, and 1.0 mm for superior, central, and inferior parts, respectively. Regarding the urethra, an interobserver contouring variability of 1.1 mm and an intraobserver variability of 0.5 mm were observed. For the rectum, we found an interobserver contouring variability of 0.6 mm and an intraobserver variability of 0.4 mm. All contouring variabilities are visualized as box plots in Fig. 2. The observers contoured the prostate smaller than the reference, especially in the superior and inferior
parts (Fig. 2). The superior and inferior parts had larger inter- and intraobserver variabilities than the central part.

The dosimetrical consequences of contouring variability are presented in Fig. 3 and Table 1. Compared with the reference contours, the prostate was contoured smaller on TRUS, resulting in a higher $D_{90}$. Contouring on CT resulted in an interobserver variability of 2.0 mm for the whole structure and 3.1, 2.4, and 3.8 mm for the superior, central, and inferior parts, respectively. Large differences with the reference (TRUS) contours, of up to 5.1 mm, were observed. Prostate contouring on CT showed larger interobserver variabilities for $V_{100}$ (5.9%) and for $D_{90}$ (11.1%) than on TRUS: 1.6% for $V_{100}$ and 9.3% for $D_{90}$ (Table 1).

**Registration variability**

Observers localized 91.3% of the fiducial markers on TRUS, 100% on CBCT, and 99.3% on MRI. After the automatic registration, the TRUS–CBCT registrations were manually adjusted in 78.2% of the observations. The observers reported that the seeds and urethra trajectory were used to manually improve the registration. For MRI–CBCT, 17.6% of the registrations were manually adjusted.

Dosimetric results of the registrations are shown in Fig. 3. Table 1 lists the corresponding inter- and intraobserver variabilities, together with the corresponding dosimetric parameters. MRI–CBCT registrations had smaller interobserver variabilities than TRUS–CBCT registrations for the prostate $V_{100}$ and $D_{90}$. The intraobserver TRUS–CBCT and MRI–CBCT registrations showed little difference in $V_{100}$ and $D_{90}$.

**Discussion**

**Contouring variability**

We performed a multiobserver study, quantifying inter- and intraobserver contouring and registration variabilities for our dynamic planning technique (12, 17).

The superior and inferior parts of the prostate showed the greatest contouring variabilities (Fig. 2). This is in agreement with other studies (15, 16). The contouring variability of the whole prostate on CT is less than the contouring variability of its parts S, C, and I (Fig. 2). This is a result of averaging of sample points (Eqs. A.3 and A.4). Larger contouring of the superior part and smaller contouring of the inferior part average out resulting in a smaller variability for the whole structure. Compared with CT, we found that TRUS contouring showed a smaller variation in geometrical differences (Fig. 2) and substantially smaller variations in dosimetric parameters (Fig. 3 and Table 1). Observers mentioned that prostate contouring on CT scans was affected by the visible $^{125}$I seeds as the prostate boundary could not be identified accurately on CT. Poor prostate visualization by CT results in large contouring variability compared with TRUS-based prostate contouring variability. This observation agrees with previously published work (15, 18).
Table 1
Inter- and intraobserver variability of the dosimetric parameters based on TRUS and CT contouring

<table>
<thead>
<tr>
<th>Contours</th>
<th>Registrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUS</td>
<td>CT</td>
</tr>
<tr>
<td>Number</td>
<td>165</td>
</tr>
<tr>
<td>Prostate $V_{100}$ (%)</td>
<td></td>
</tr>
<tr>
<td>Average reference</td>
<td>98.8</td>
</tr>
<tr>
<td>Average observer</td>
<td>98.8</td>
</tr>
<tr>
<td>SD$_{inter}$</td>
<td>1.6</td>
</tr>
<tr>
<td>SD$_{intra}$</td>
<td>0.7</td>
</tr>
<tr>
<td>Prostate $D_{90}$ (%)</td>
<td></td>
</tr>
<tr>
<td>Average reference</td>
<td>122.1</td>
</tr>
<tr>
<td>Average observer</td>
<td>130.6</td>
</tr>
<tr>
<td>SD$_{inter}$</td>
<td>9.3</td>
</tr>
<tr>
<td>SD$_{intra}$</td>
<td>1.1</td>
</tr>
<tr>
<td>Urethra $D_{90}$ (%)</td>
<td></td>
</tr>
<tr>
<td>Average reference</td>
<td>136.4</td>
</tr>
<tr>
<td>Average observer</td>
<td>154.1</td>
</tr>
<tr>
<td>SD$_{inter}$</td>
<td>14.6</td>
</tr>
<tr>
<td>SD$_{intra}$</td>
<td>1.5</td>
</tr>
<tr>
<td>Rectum $V_{100}$ (cm$^3$)</td>
<td></td>
</tr>
<tr>
<td>Average reference</td>
<td>0.83</td>
</tr>
<tr>
<td>Average observer</td>
<td>2.33</td>
</tr>
<tr>
<td>SD$_{inter}$</td>
<td>0.43</td>
</tr>
<tr>
<td>SD$_{intra}$</td>
<td>0.23</td>
</tr>
</tbody>
</table>

CBCT = cone beam CT; “—” = not available; SD$_{inter}$ = (1 SD with respect to the reference average) the interobserver variability; SD$_{intra}$ = (1 SD with respect to the observer average) the intraobserver variability; TRUS = transrectal ultrasound.

The variability of the rectum contours increased slightly from superior to inferior as the distance between the rectal wall and TRUS probe increased and the rectal wall contrasted less distinctly from the surrounding tissues (Fig. 2). The volume of the rectum that received the prescribed dose was small, and the variability in $V_{100}$ was within 0.5 cm$^3$.

Registration variability

On TRUS, $^{125}$I seeds cause bright reflections similar to those of fiducial markers; consequently, seeds can be mistaken for fiducial markers. During registration, however, the observer had an indication of the positions of the markers on the CBCT or MRI data set, providing guidance to identify the fiducial markers on TRUS. With this guidance, observers were able to identify 91% of the fiducial markers on TRUS compared with 99.3% on MRI and 100% on CT.

The identification of gold markers on MRI (voids) and CBCT (bright spots) was straightforward and lead to an accurate automatic registration by the TPS. As no seeds were present in the MRI data set, observers had limited feedback to adjust the registration. The observers reported that they made less manual interactions during MRI–CBCT registrations (18%) than during TRUS–CBCT registrations (78%). The TPS did not allow for quantification of the translation and rotation resulting from the registration.

Registration of MRI with CBCT resulted in interobserver variabilities of 1.3% for the $V_{100}$ and 2.1% for the $D_{90}$. De Brabandere et al. (14) reported a $V_{100}$ variability of 2.9% based on CT–T1–T2 registration together with a $D_{90}$ variability of 7%. These variabilities based on multimodality image registration are larger than the variabilities found in the present study. In line with the observations of the contouring study, the smaller variabilities of our study likely originate from the fact that the present study is performed with the assistance of well-visible fiducial gold markers, in a single institute, with one set of instructions and training for all observers. Furthermore, the registration in our study started with an automated step, defining a good starting point for the registration and leaving the amount of manual interaction limited.

The mean difference between the normalized dosimetric parameters achieved with TRUS–CBCT and MRI–CBCT registration was <1% for $V_{100}$ and $D_{90}$. We do not consider these differences to be clinically relevant.

MRI provides superior soft-tissue contrast compared with TRUS and CT (14, 15). It is possible to visualize suspect lesions in the prostate and define a volume for focal (9) or differential dose prescription brachytherapy (6). The registration of TRUS and MRI data sets to CBCT introduces additional uncertainties. Table 1 lists that the intra- and interobserver variabilities is in the order of 1–3% for the prostate $V_{100}$ and $D_{90}$ and the urethral $D_{90}$. These values are expected to be higher for smaller volumes, for example, focal or intraprostatic boost volumes (21). Therefore, when incorporating a pretreatment MRI in the implantation procedure, treatment margins should be applied to correct for the registration uncertainties (6).

General

Contouring and registration can be affected by the slice distance and slice thickness of the studied image data sets.
Particularly, using TRUS, the variability in contouring of the base and the apex may be lower with closer spaced slices. Also the registration variability may reduce with closer spaced slices. Furthermore, the TPS affects the results. The automatic point-based image registration, volume determination, and dosimetric parameter calculation may depend on the software that is used.

In addition to VariSeed 8.0.2, the authors reviewed 3 cases in VariSeed 9.0. Small differences were observed (<0.5% for \(D_{90}\) and \(V_{100}\)). The observed differences cannot be attributed to the registration only. The dose calculation, volume determination, and DVH calculation may be different too. In the present study, we think that the results are little affected by the software that performs the point-based registration.

The dosimetric inter- and intraobserver variabilities of TRUS contouring are higher than the variability in the registrations. The quality of the procedure can be improved most by reducing the greatest source of uncertainty: contouring variability.

Our study showed that within our institute, having the same training and protocols, the interobserver variability was considerably smaller than the values reported by De Brabandere et al. (14). This demonstrates that training and guidelines can lead to better consensus in prostate contouring. This conclusion is in agreement with a study from Khoo et al. (18), who reported reduced inter- and intraobserver variabilities after three training sessions, even for experienced radiation oncologists.

In our dynamic dosimetry procedure, we strive to solve underdosages during the implantation procedure (12, 13). Adaptations should not be made if the underdosage is an undesirable side effect of the registration of two data sets or a result of contouring variability. The present study points out that the variabilities in \(V_{100}\) of 2.0% and 1.6% caused by registration and contouring, respectively (Table 1), were smaller than the magnitude of underdosages we adapt (9 ± 6%) (13). The adaptation resulted in an increase of 15 ± 9% in \(D_{90}\), whereas variabilities in \(D_{90}\) caused by registration and contouring were 3% and 9%, respectively (Table 1). The underdosages that we observed in the implantation procedure are predominantly caused by other factors than registration or contouring variability. In our procedure, implant dynamics (e.g. edema and seed displacement) are, therefore, the main cause of underdosages that are adapted during implantation (13, 17).

Both, contouring and registration can be improved. Consistency of contouring can be increased using guidelines and training (18). Furthermore, accuracy can be gained in the registration procedure. For instance, the speed and accuracy of TRUS–CBCT registration could increase by applying a registration that uses all seeds that are localized on TRUS and CBCT (22). Ideally the implantation is performed under live MRI guidance (23), obviating the registration step.

Conclusion

Prostate dosimetry is affected by interobserver contouring (2% for \(V_{90}\), 9% for \(D_{90}\)) and registration variability (\(D_{90}\), 3% for TRUS–CBCT and 2% for MRI–CBCT). The base and apex of the prostate show more geometric contouring variability than the central part. The observed dosimetric variability is smaller than underdosages that are adapted during our dynamic brachytherapy procedure.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.brachy.2017.01.010.

References


