Radiotherapy and Oncology 107 (2013) 13–19

Contents lists available at SciVerse ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Image guided brachytherapy

Dosimetric impact of interobserver variability in MRI-based delineation for cervical cancer brachytherapy

Taran Paulsen Hellebust [a,b,c,*], Kari Tanderup [d,e], Christoffer Lervåg [f], Elena Fidarova [g], Daniel Berger [h], Eirik Malinen [a,c], Richard Pötter [h,i], Primož Petrič [j]

[a] Department of Medical Physics, Oslo University Hospital, The Radium Hospital, Oslo; [b] Department for Radiation Protection and Nuclear Safety, Norwegian Radiation Protection Authority, Østerås; [c] Department of Physics, University of Oslo, Norway; [d] Department of Oncology, Aarhus University Hospital; [e] Institute of Clinical Medicine, Aarhus University, Denmark; [f] Department of Oncology, Ålesund Hospital, Norway; [g] Division of Human Health, International Atomic Energy Agency, Vienna; [h] Department of Radiotherapy, Medical University of Vienna, Vienna; [i] Christian Doppler Laboratory for Medical Radiation Research for Radiation Oncology, Medical University of Vienna, Austria; [j] Department of Brachytherapy, Institute of Oncology Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 12 September 2012
Received in revised form 22 December 2012
Accepted 29 December 2012
Available online 22 February 2013

Keywords:
Brachytherapy
Cervical cancer
Interobserver variability

ABSTRACT

Purpose and background: To study the dosimetric impact of interobserver delineation variability (IODV) in MRI-based cervical cancer brachytherapy.
Materials and methods: MR images of six patients were distributed to 10 experienced observers worldwide. They were asked to delineate the target volumes and the organs at risk (OARs) for each patient. Two types of reference contours were created (Expert Consensus – EC and Simultaneous Truth and Performance Level Estimation – STAPLE). Optimised plans based on both EC- and STAPLE-contours were prepared. These plans were transferred to each of the observer contour sets and the resulting DVH parameters (D_{90} and D_{2cc}) were calculated. For each patient the standard deviation (SD) for the 10 observers was calculated.
Results: A mean relative SD of 8–10% was found for GTV and High Risk CTV (HR-CTV) D_{90} analysing one single fraction. For rectum and bladder the mean relative SD for D_{2cc} was 5–8% while sigmoid was at 11%. For the whole treatment the IODV in HR-CTV caused an uncertainty of ±5 $Gy_{\alpha/\beta=10}$ (1SD). The corresponding figure for OARs was ±2–3 $Gy_{\alpha/\beta=3}$. The results were not sensitive as to which structure set was used for the optimisation.
Conclusions: For the target volumes the dosimetric impact of IODV was smallest for the GTV and HR-CTV, while IODV had an even smaller impact on the bladder and rectum.
 Radiotherapy and Oncology 107 (2013) 13–19

MRI is the recommended imaging modality for 3D image based cervical cancer brachytherapy [1–3]. MR-based image guided adaptive brachytherapy (IGABT), offers the ability to individualise the treatment according to the size and location of the target volumes and the organs at risk (OARs) for each patient [4–7]. A reduction in pelvic recurrence and major morbidity using IGABT compared to historical patient series has been demonstrated [8]. In these historical series less individualised treatment planning was used.

One fundamental principle of 3D treatment planning is reliable delineation of target volumes and OARs. Even if an optimal image modality is used, interobserver variability is seen when many observers are asked to delineate target volumes and/or OARs on the same patient case. It is important to quantify such variability and consider this uncertainty during the treatment planning process. Several studies have addressed this topic for various tumour sites, both for external beam radiotherapy (EBRT) [9–15] and brachytherapy [16–19].

In interobserver studies, the contour data may be analysed using different volumetric indices like the concordance/conformity index [9,11,19], the Kappa index [13] or the DICE [14]. However, such indices do not directly reflect the consequences of interobserver contouring variability on the resulting variability in dose delivered to the patient. This has been addressed for EBRT [10,15], while two studies for low dose rate prostate brachytherapy have been performed [16,17]. The impact of different delineations performed by two observers in cervical cancer patients has been investigated previously [18,19]. However, these studies are not expected to provide representative estimates of interobserver variability, as only two observers were included. In the present study we aim to quantify the dosimetric impact of interobserver variability on MRI-based delineation for cervical cancer brachytherapy by analysing target and OAR delineations from ten observers on six different cervical cancer patients.

* Corresponding author at: Department of Medical Physics, Oslo University Hospital, The Radium Hospital, Oslo, Norway.
E-mail address: taran.paulsen.hellebust@nrpa.no (T.P. Hellebust).

0167-8140/$ - see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.radonc.2012.12.017

Materials and methods

Patients, imaging and delineation procedure

Three patients treated at Medical University of Vienna (MUV) and three patients treated at Institute of Oncology Ljubljana (IOL), all with FIGO stage IIb–IIIb cervical cancer, were selected for this study. In the following, these patients will be referred to as V1–V3 and L1–L3. MR images at time of diagnosis and at brachytherapy for all the six cases were electronically distributed to several institutions worldwide together with a description of clinical findings for each case. In MUV the image acquisitions were performed with a 0.2 Tesla system with a pelvic surface coil and a fast spin echo T2-weighted sequence [20]. For the patients from IOL a 1.5 Tesla system was used, also with pelvic surface coil [21]. A more thorough description of the cases and the imaging is given in Petric et al. [22]. The data were made anonymous prior to distribution.

Twelve centres were invited to participate in the study and ten of these accepted. All the participating observers were experienced in MRI-based brachytherapy of cervical cancer and they had all successfully passed the dummy run (QA procedure) in the International study on MRI-guided brachytherapy in locally advanced cervical cancer (EMBRACE, www.embracestudy.dk). They were asked to delineate the target volumes and OARs in every case according to the GEC-ESTRO recommendations [1,2]. Dicom RT structure files from each observer were electronically collected and imported into the treatment planning system.

Expert group delineation

An expert group was established, consisting of four radiation oncologists with considerable experience in the field of MRI-based gynaecological brachytherapy, two of them being the key contributors to the GEC-ESTRO recommendations [1,2]. For all patients, the expert group created an Expert Consensus (EC) contour set by delineating the Gross Tumour Volume (GTV), High Risk Clinical Target Volume (HR-CTV) and Intermediate Risk Clinical Target Volume (IR-CTV), as well as the bladder, rectum and sigmoid colon.

Algorithm based contours

Algorithm based contours were also generated for all target volumes and OARs using an expectation–maximisation algorithm for Simultaneous Truth and Performance Level Estimation (STAPLE) [23] and RTKIT (Radiotherapy DICOM Toolkit, http://github.com/dicom/rtkit). The algorithm computes a probabilistic estimate of the true delineation for all the different structures in each MR slice, with the implicit assumption that the "true" contour exists within the collection of the submitted contours [24,25].

Treatment planning

The treatment planning was performed in OncentraBrachy® 4.2 (Nucletron, an Elekta company, Veenendaal). Since an applicator library was not available the applicators were reconstructed manually according to the GEC-ESTRO recommendation [26]. All the cases were treated with a combination of tandem/ring applicator and parametrial needles, inserted through the ring template. For V1 and L1 nine needles were used, while five needles were used for the other four cases.

The starting point for the treatment planning was a standardised source configuration with a prescription dose of 7 Gy to point A without any loading in the needles. Dose–volume-histograms (DVH) were used to evaluate the plans, looking at the D_{90} (dose to 90% of the volume) for the target volumes and D_{2cc} (minimum dose in the most exposed 2 cm^3 volume) for the OARs. To optimise the dose distribution, graphical optimisation and manual adjustment of the dwell times were used without re-normalisation. During the optimisation the weight of the dwell positions inside the needles was restricted to 20% of the dwell weight used for the sources inside the tandem/ring (30% for single positions in special cases) [27]. The treatment plans were optimised to achieve HR-CTV D_{90} of at least 7.4 Gy, IR-CTV D_{90} of at least 3.5 Gy, D_{2cc} of maximum 6 Gy, 4.7 Gy and 4.7 Gy for bladder, rectum and sigmoid, respectively. These target constraints correspond to a total 2 Gy-equivalent-dose (EQD2) of 87 and 60 $Gy_{\alpha/\beta=10}$ for HR-CTV and IR-CTV, respectively, when a treatment schedule of 4 brachytherapy treatments plus 45 Gy in 25 fractions of external beam radiotherapy is considered. For the same schedule, the OAR constraints correspond to EQD2 of 86 and 72 $Gy_{\alpha/\beta=3}$ for the bladder and for the rectum and sigmoid, respectively. Treatment plan optimisation was performed separately on the EC-contours and the STAPLE-contours.

Using the source configuration (including dwell weights) from the optimised plans, the D_{90} and the D_{100} for the target volumes and D_{2cc} for OARs were calculated for the observer contour sets for all the six cases. For each DVH-parameter the mean value and standard deviation (SD) for the 10 observers were calculated. In order to compare the variability of the different DVH-parameters, the relative SD (in %) was also calculated. These data will show the impact of interobserver delineation variability on one single fraction. The impact of contouring uncertainties on the total dose delivered will depend on the treatment schedule used. In this study the total dose was calculated for a brachytherapy schedule of four fractions and 45 Gy EBRT delivered in 25 fractions, as described above. For each observer and DVH-parameter the total EQD2 was calculated using the EC- or the STAPLE-contours-optimised plan for all four fractions. From these figures the mean value and the absolute SDs were calculated for each case and DVH parameter expressed in $Gy_{\alpha/\beta=10}$ and $Gy_{\alpha/\beta=3}$ for the target volumes and OARs, respectively.

Results

Volumes of submitted structures

Fig. 1a shows the HR-CTV delineations from 10 observers for the L3 case in a para-transversal MR slice 1 cm cranial from the surface of the ring applicator. The average, the SD and the range of the volumes for the submitted structures are shown in Table 1. The smallest mean GTV was seen for the L2 case, while the largest was seen for L1. For the target volumes the largest interobserver variation in terms of absolute SD was found for the submitted IR-CTVs, with a variation of ±26 cm^3 for the V1 case.

Treatment plans and DVH parameters

The treatment planning constraints were respected for all the cases for the EC-contours. However, for the STAPLE-contours for V1 and V3 the HR-CTV D_{90} constraints had to be compromised in order to fulfil the OAR constraints. Fig. 1b and c shows the EC-contours- and the STAPLE-contours-optimised plan for the L1 case. The calculated DVH-parameters for HR-CTV, bladder, rectum and sigmoid delineations using the source configuration from the EC-contours-optimised treatment plan are shown in Fig. 2 for all observers. The figure shows that the dosimetric variation is different for target and OARs and that the variation is also depending on the patient case. For instance, the rectum in the L1 and the L3 show a much higher relative SD compared to the other cases. This is due to two outliers, as can be seen in Fig. 2, lower, left panel. In both of these cases two observers have defined the superior border of the