Image-guided adaptive brachytherapy in cervical cancer:

Learning curve assessment

for the delineation of the clinical target volumes

Renaud Mazeron

Département de Radiothérapie, Gustave Roussy, Villejuif

Université Paris Saclay, Orsay, France
ABSTRACT

Background and objective

The implementation of image-guided adaptive brachytherapy in locally advanced cervical cancer requires skills in imaging and mastering of the GEC-ESTRO recommendations. The aim was to establish learning curves for the delineation of the intermediate and high-risk clinical target volume (CTV_{IR} and CTV_{HR}).

Material and methods

Three residents, naive in brachytherapy, who priory attended a teaching course on the GEC-ESTRO recommendations, prospectively contoured consecutive cases. The accuracy in the delineation of the CTV_{HR} and CTV_{IR} was evaluated by experienced seniors, and scored as acceptable or unacceptable. The LC-CUSUM (learning curve-cumulative summation) test was used to determine if an adequate level of performance level could be reached. This level was defined as a maximum 15% rate of failures.

Results

The trainees were aged from 29 to 31 years with an experience in radiation oncology ranging from 3.5 to 4.5 years. A total of 66 cases were delineated during a 4-month observation period. For the CTV_{HR}, two trainees reached the acceptable performance threshold after 13 and 17 cases respectively. The last one, did not reach the threshold after 18 procedures. For the CTV_{IR}, only one trainee could have reached the threshold after 17 performances.
Conclusion

The LC-CUSUM appeared applicable to the evaluation of the delineating competency. Preliminary results demonstrates a fast learning curve in residents, especially for the CTV_{HR}. The study needs to be continued and expanded to additional trainees and refine these findings.

KEY WORDS

Learning curve, competency, LC-CUSUM, image-guided adaptive brachytherapy, cervix cancer, CTV, delineation.
INTRODUCTION

The treatment of locally advanced cervical cancer relies on the combination of external beam radiotherapy (EBRT) and brachytherapy (BT) [1, 2]. The crucial role of this last component has been highlighted in a recent epidemiologic database analyses showing that the decline in the use of brachytherapy in the United States of America was in mirror to an increase in the use of high precision modern EBRT techniques and associated with a significant decline in overall survival rates. The treatment has undergone dramatic changes during the last two decades [3]. First, concomitant chemoradiation became a standard in 1999, with a 6% overall-survival rate gain at 5 years [4]. The second great advance has been the advent of image-guided adapted brachytherapy (IGABT), a high-precision radiation technique relying on the ability to control the displacement of a miniature stepping source, under the guidance of 3D imaging [5]. For years, the prescription has mainly relied on point geometrically constructed from the implant such as point A and thus was not adapted to patient’s anatomy, tumor shape or tumor response. Another system, used in France, consisted in prescribing on an isodose envelope. The coexistence of these different systems impaired comparisons.

In 2005, in a will of harmonizing the reporting of brachytherapy, the GEC-ESTRO (Groupe Européen de Curiethérapie – European Society for Radiation Oncology) published recommendations on target definition and dosimetric parameters [6]. They proposed 3 volumes:

- A low-risk clinical target volume (CTV_{LR}), corresponding to EBRT targets, comprising all potentially involved tissues (cervix, uterus, parametria, part of the vagina, pelvis±paraaortic lymph nodes regions) [7]
- A high-risk CTV (CTV_{HR}) including the whole cervix and residual disease at the time of brachytherapy, as well as so-called T2-MRI “grey zones”, considered to be disease in regression
An intermediate-risk CTV (CTV_{IR}) encompassing the CTV_{HR} with directional margins, and at least the initial volume at diagnosis (Figure 1). This volume was particularly innovative as it is adapted to tumor response to chemoradiation.

Since these recommendations were rapidly and word widely adopted, the ICRU (International Commission for Radiation Units and measurements) proposed a new report, incorporating these recommendations [8]. Moreover recent studies reported clear dose-volume effect correlations between the D_{90} CTV_{HR} (minimal dose prescribed to 90% of the CTV_{HR}) and D_{90} CTV_{IR} and local control as well as correlations between the D_{2cm^3} (minimal dose calculated in the maximally exposed 2cm^3) of the bladder and rectum and the occurrence of late morbidity [9-12]. Therefore these volumes are not only pertinent for reporting but crucial for prescribing. Dose-volume parameters depend on the accuracy of the contouring. A Study reporting on interobserver variations has been published suggesting good agreements among experts [13]. However, its appliance in less experience observers have not been studied. Delineation of the two CTVs requires expertise in imaging, in addition to the appropriation of the CTV concepts and clinical skills.

The aims of this study were to establish trainees’ learning curve and to evaluate their performance using a statistical model, the LC-CUSUM (Learning Curve-Cumulative Summation).

MATERIAL AND METHODS

Treatments

The study was initiated at Gustave Roussy and focused on consecutive patients treated for locally advanced cervical cancer with IGABT. All patients were treated with a combination of pelvic±paraortic EBRT to 45-50.4 Gy with a common fractionation. Some of these patients received this first sequence of
treatment in other centers and were secondary referred for brachytherapy. Then they were treated with pulsed-dose rate image-guided adaptive brachytherapy. The procedure has been detailed in previous publications [14, 15]. Briefly, after insertion of the personalized vaginal mould applicator, patients underwent a 1.5 or 3 T MRI with T2 axial, sagittal, and coronal acquisitions with dummy sources inserted into the applicator catheters to facilitate its reconstruction [16, 17]. Images were transferred to Oncentra® (Nucletron, an Elekta Company, Stockholm, Sweden) or Brachyvision® (Varian Medical Systems, Palo-Alto, USA) platforms for delineation of the structures of interest (CTVIR and CTVIR, bladder, rectum, and sigmoid colon, Figure 1) and planning. Institutional planning aims were 85 Gy to the D90 of the CTVIR (in 2 Gy equivalent, summing EBRT and BT doses, and applying the linear quadratic model with an α/β of 10 Gy and a half-time repair of 1.5 h), 60 Gy to the D90 of the CTVIR, 85 Gy to the D2cm3 of the bladder, and 75 Gy to the D2cm3 of the sigmoid and rectum (using the same model and an α/β of 3 Gy). Dwell positions and times were adapted to optimize the plans. At the end of the process, and in order to respect a maximal physical dose per pulse of 0.5-0.6 Gy to organs at risk (OAR), the number of pulses was adjusted. Therefore, for a physical prescription of 15 Gy to the D90 of the CTVIR, the number of pulses ranged from 30 pulses of 50 cGy to 60 pulses of 25 cGy, thus respecting the total physical dose.

Trainees

The trainees were the residents of the brachytherapy unit. They were three during the study period, all of them naive in the field of brachytherapy. They all followed a Brachytherapy theoretical 2-day teaching course held in Nice, France, in March 2016 and had access to the GEC-ESTRO publication [6]. During their stay, the cases were distributed to the trainees according to their availability. They had to delineate alone. Then, the contours were reviewed and evaluated by one of the three seniors of the service. If the
contours were suitable for treatment the case was considered as a success. On the opposite, if significant corrections were required, the case was considered as a failure. The objective was to detect unacceptable deviation from the GEC-ESTRO recommendation that might impair the chance of achieving local control or lead to over treat the patient and thus increase the risk of morbidity. Prior to the initiation of the study, a list of in unacceptable deviations was defined, aiming to homogenize the seniors’ judgments (Table 1). Cases were successively and prospectively evaluated. After each failure, the decision was explained to the trainee, and the contours corrected conjointly (“buddy system”). Trainees gave their consent to be evaluated such as described above.

Statistics and hypothesis

The LC-CUSUM test was selected to assess trainees’ performances. The LC-CUSUM sequentially tests the null hypothesis that performance is unacceptable against the alternative one that performance is adequate. A score is computed from the successive trainees’ attempts. Each success increases in the score whereas each failure yields a decrease. Once the score reaches a predefined value, the null hypothesis is rejected. The performance level is then considered acceptable and the trainee competent. Conversely, the performance remains unacceptable as long as the LC-CUSUM score remains below this decision threshold. For the purposes of the study, an adequate performance was defined by a maximal failure rate of 15%, and an inadequate performance as a rate of more than 30%, with an equivalent zone of 7.5%. The decision limit was set at 1 after simulations so that with 7 trainees and 50 procedures per trainee, the probability to rightly detect a trainee’s competency was 91% and conversely to miss a genuine competency of 16%. Statistical analyses were performed using R statistical package 3.1 (The R Project for Statistical Computing, https://www.r-project.org/, Vienna, Austria).
RESULTS

Three trainees participated. Their characteristics, as well as an overview of the cases, are summarized in Table 2. On the overall, they were well-advanced in their residency. The observation period started at the beginning of their stay in May 2016 until the end of August 2016. During this 4-month period 66 patients were treated for locally advanced cervical cancer in the brachytherapy unit. The number of cases delineated per trainee was 20, 18, and 26. In all cases but 2, the brachytherapy was guided by MRI, and in 2 cases by CT due to the unavailability of the MRI scanner (one breakdown and one maintenance).

On the overall, 55 out of 66 cases (83.3%) were considered as acceptable for the CTV_{HR}. This rate was 72.7% (48 out of 66) for the CTV_{IR}. Focusing on the first 10 procedures per trainee, successes were observed in 24 cases out of 30 (80%) and 17 out of 30 (56.7%) for the CTV_{HR} and CTV_{IR} respectively. The LC-CUSUM test showed that trainee #1 and 3 reached the competency threshold after 13 and 17 cases respectively. The last trainee, #2, did not reach the threshold after having performed 18 procedures (Figure 2).

Concerning the CTV_{IR}, trainee #3 was the only one who reached the threshold during the observation period, after 17 cases (Figure 3). The two remaining trainees did not reach the decision limit after 18 ad 20 attempts respectively.

DISCUSSION

These preliminary results show the feasibility of the appliance of the LC-CUSUM in the assessment of the competency in delineating target volumes. This is the first time that this method is applied to radiation oncology. Previously this method has shown its efficiency in assessing the performances of trainees in
technical procedures in orthopedics or gynecology [18-20]. Radiotherapy which relies on successive and repetitive technical tasks seems perfectly suitable for such evaluation. The LC-CUSUM test could be used either in the initial formation or later in the professional life to evaluate the implementation of new techniques. Image-guided adaptive brachytherapy is a recent progress in the treatment of cervical cancer that relies on imaging and thus beyond the appropriation of the target risk-volume concepts requires skills in pelvic MRI/CT interpretation. The GEC-ESTRO recommendations have been rapidly and widely adopted worldwide, and number of teaching courses have be proposed to spread the technique. However, the question of defining the competency remains open. Our study showed that two trainees reached the threshold for the CTV\textsubscript{HR} within the first 17 procedures. The third trainee failed to demonstrate his competency. However, the amount of 50 procedures warranting a 91% probability of rightly detecting its competency was not reached. The study will be pursued during 2 months until the next residents’ rotation. Concerning the CTV\textsubscript{IR}, the learning seems lengthier. On one hand, this volume is linked to CTV\textsubscript{HR} with the addition of directional margins, which implies that a failure in the genesis of the CTV\textsubscript{HR} will inevitably lead to an inadequate CTV\textsubscript{IR}. On the other hand, the CTV\textsubscript{IR} is less intuitive as it comprises adaptive margins which are not only based on the per-brachytherapy MRI/CT. Petric et al. showed in an inter-observer study on CTV delineation that CTV\textsubscript{IR} was more subject to uncertainties among experts than the CTV\textsubscript{HR} [13]. Therefore it is not surprising that its learning is more difficult than the one of CTV\textsubscript{HR} in a sample of trainees.

Our findings suggest that the acquisition of the competency to define the CTV in IGABT can be acquired within a short time. However it should be keep in mind that these trainees were volunteers and chose to stay in the brachytherapy unit, indicating a real interest in the field and a will to acquire this skill. Moreover they were well advanced in their residency, and thus experienced in radiotherapy in contouring. Two of them had even spent 6 months in a radiology department. Another specificity is the number of cases treated within a short time per trainee. During the four-month period 66 patients with
locally advanced cervical cancer were referred for brachytherapy in which is a consequent amount for an occidental center. This accrual probably contributes to a quick learning. However, in smaller center a less important recruitment can be circumvented by de archiving old cases, which is indeed more time consuming than treating real patients. In our study, trainees were in a realistic situation, performing clinical examination, the implantation, and then contouring with full access to clinical cartoons, reports, and imaging.

There are still pending questions. First, there is a need to complete the study and refine the results. The number of observations required based on statistical hypothesis was not reached during the 4-month observation, nor the number of observers. The evaluation of the first 3 trainees will be pursued until the end of their stay, and the study relaunched in November with the coming residents’ rotation. The project aims not only to include the residents from Gustave Roussy but also those training at University Hospitals Saint Louis, Pitié Salpêtrière and Tenon, in Paris. This will allow reaching a more significant number of observers but also addressing additional questions such as the impacts of the experience in radiotherapy with the inclusion of younger residents or of the imaging-guidance modality on the learning curve. It has been showed that MRI is superior to CT in the definition of the CTVs. The use of CT leads to an overestimation of the target volumes, which might be overcome by experience [21]. However, the modality is more accessible than MRI and the opening of the study to other centers will allow establishing a learning curve for CT.

The optimal goal would be to define a number of cases to delineate to be considered competent. The LC-CUSUM test could be used in other areas of radiation oncology to define thresholds, depending on the complexity of each task, to ensure that the competence is acquired by most of the trainees. The LC-CUSUM test can be used to define these learning objectives. Another issue is also the continuous assessment of performance. For instance, it would be of interest to determine the number of cases required to maintain the competency. In this field, the LC-CUSUM test has been used [22]. For instance,
Biau et al. monitored the performances of a sonographers assessing the nuchal translucency of fetus during the first trimester of pregnancy. They observed, based on the LC-CUSUM test, a decline in the performance after 250 procedures, becoming unacceptable.

CONCLUSION

The LC-CUSUM appeared applicable to the evaluation of the delineating competency. Preliminary results demonstrates a fast learning curve in residents, especially for the CTV_{HR}. The study needs to be continued and expanded to additional trainees and refine these findings.
REFERENCES


ACKNOWLEDGEMENTS

David Biau, Department of Orthopaedic Surgery, Hospital Cochin - University Paris-Descartes and INSERM U1153: study design and statistics

Cyrus Chargari, Christine haie-Meder, Tony Felefly, Charles Canova, and Andres Huertas: participation and data collecting.
FIGURE LEGENDS

Table 1: Failure criteria

Table 2: Trainees’ and cases’ characteristics

Figure 1: Delineation of clinical target volumes for a stage IIB lesion (T2 MRI axial slice)

Figure 2: Individual LC-CUSUM scores for CTV\textsubscript{HR}

Figure 3: Individual LC-CUSUM scores for CTV\textsubscript{IR}
Table 1: Failure criteria

<table>
<thead>
<tr>
<th>Failure Criteria</th>
<th>High-risk CTV (CTV&lt;sub&gt;HR&lt;/sub&gt;)</th>
<th>Intermediate-risk CTV (CTV&lt;sub&gt;IR&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CTV encompassing organs at risk (except adequately in IVA disease)</td>
<td>CTV encompassing organs at risk (except adequately in IVA disease)</td>
</tr>
<tr>
<td></td>
<td>More than one 3mm thick slice missing</td>
<td>Insufficient margin around CTV&lt;sub&gt;HR&lt;/sub&gt;: &lt; 1cm in all directions</td>
</tr>
<tr>
<td></td>
<td>Obvious residual disease not included</td>
<td>Overestimation of margins around CTV&lt;sub&gt;HR&lt;/sub&gt; &gt; 15 mm</td>
</tr>
<tr>
<td></td>
<td>Grey-zone not included</td>
<td>CTV&lt;sub&gt;IR&lt;/sub&gt; not encompassing the disease at diagnosis:</td>
</tr>
<tr>
<td></td>
<td>Overpassing of the CTV&lt;sub&gt;HR&lt;/sub&gt; limits</td>
<td>- Vaginal or uterine involvement at diagnosis not taken into account</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Initial distal parametrium involvement not taken into account</td>
</tr>
</tbody>
</table>
Table 2: Trainees’ and cases’ characteristics

<table>
<thead>
<tr>
<th></th>
<th>Trainee 1</th>
<th>Trainee 2</th>
<th>Trainee 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>29 years</td>
<td>31 years</td>
<td>30 years</td>
</tr>
<tr>
<td><strong>Experience in radiation therapy</strong></td>
<td>4.5 years</td>
<td>4.5 years</td>
<td>3.5 years</td>
</tr>
<tr>
<td><strong>Procedures #</strong></td>
<td>20</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td><strong>MRI/CT guidance</strong></td>
<td>90% / 10%</td>
<td>100% / 0%</td>
<td>100% / 0%</td>
</tr>
<tr>
<td><strong>FIGO stages</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II B</td>
<td>20.0%</td>
<td>27.8%</td>
<td>37.5%</td>
</tr>
<tr>
<td>II A</td>
<td>15.0%</td>
<td>5.6%</td>
<td>4.2%</td>
</tr>
<tr>
<td>II B</td>
<td>50.0%</td>
<td>50.0%</td>
<td>37.5%</td>
</tr>
<tr>
<td>III A</td>
<td>0.0%</td>
<td>5.6%</td>
<td>4.2%</td>
</tr>
<tr>
<td>III B</td>
<td>15.0%</td>
<td>11.1%</td>
<td>16.7%</td>
</tr>
<tr>
<td>IV A</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

FIGO: Fédération Interationale de Gynécologie Ostétrique classification.
Figure 1: Delineation of clinical target volumes for a stage IIB lesion (T2 MRI axial slice)

CTV_{HR}: high risk clinical target volume, CTV_{IR}: intermediate risk clinical target volume, SBow: small bowel, Blad: bladder, Rect: rectum

The HR encompasses the whole cervix as well as any residual disease. The CTV_{IR} comprised the CTV_{HR} with margins of 10 to 15 mm taking into account the initial disease extension.
Figure 2: Individual LC-CUSUM scores for $CTV_{HR}$

Horizontal dotted line: decision threshold defining the competency ($LC$-CUSUM score = 1)
Figure 3: Individual LC-CUSUM scores for $CTV_{ir}$

Horizontal dotted line: decision threshold defining the competency (LC-CUSUM score = 1)