Harmonization of practices between radiotherapy centres in the Nord and Pas-de-Calais regions (France): A three-year evaluation

Harmonisation des pratiques entre les centres de radiothérapie des départements du Nord et du Pas-de-Calais : évolution sur 3 ans

D. Pasquier a,b,c,d,*, F. Darloy e, S. Dewas f, L. Gras e, S. Maillard f, H. Rhlouch g, M. Tokarski h, J.P. Wagner i, A.C. Degrendel-Courteguise i, C. Dufour k, M. Fares i, L. Gilbeau m, O. Olszyk n, B. Castelain a, É. Lartigau a,b,c,d

ARTICLE INFO

Article history:
Received 19 December 2017
Received in revised form 18 February 2018
Accepted 5 March 2018

Keywords:
Radiotherapy
Delineation
Interobserver variability

ABSTRACT

Purpose. – The delineation of volumes of interest can be a source of significant interobserver variability. The purpose of this study was to improve the homogeneity of delineation between oncologist-radiotherapists in the territorial departments of Nord and Pas-de-Calais (France) through discussions of clinical cases and the adoption of common published reference documents.

Materials and methods. – All eleven radiotherapy centres in the Nord and Pas-de-Calais departments of France participated. The localizations assessed to date included prostate, head and neck, breast and brain cancers. For each localization, the junior or senior physician(s) in charge of pathology delineated the volumes of interest according to their usual practices. Validated indices, including the Dice similarity coefficient, were used to quantify the delineation differences. The anonymized results were presented at two to three annual meetings. A second delineation of the clinical cases was then carried out to quantify homogenization. An evaluation of dosimetry practices was also conducted for prostate cancer. Wilcoxon assay matched data were used.

Results. – Our work showed either satisfactory delineation concordance after the initial assessment of improved delineation concordance. For prostate cancer, the Dice similarity coefficient values were greater than 0.6 initially in two of the three clinical cases. For head and neck cancers, a statistically significant improvement was observed for only one of the clinical target volumes. More than half of the Dice similarity coefficient values were greater than 0.6 in the first comparison. The study of clinical cases of breast cancer allowed a homogenization of the delineation of five of the six lymph node clinical target volumes. The dosimetry study of prostate cancer allowed for a homogenization of practices.

Conclusion. – This work makes it possible to harmonize the delineation practices around validated standards. An extension to the entire Hauts-de-France region is planned.

© 2018 Société française de radiothérapie oncologique (SFRO). Published by Elsevier Masson SAS. All rights reserved.
**RÉSUMÉ**

**Objectif de l’étude.** – La délinéation des volumes d’intérêt est le siège d’une variabilité entre les observateurs. L’objectif de cette étude était l’harmonisation de la délinéation de tumeurs par les oncologues radiothérapeutes des départements du Nord et du Pas-de-Calais.


**Résultats.** – Nos travaux ont montré soit une concordance de délinéation satisfaisante après l’évaluation initiale, soit une meilleure concordance de délinéation. Dans le cas du cancer de la prostate, les valeurs du coefficient de similarité de Dice du volume cible anatomoclinique étaient supérieures à 0,6 dans deux des trois cas cliniques. Pour les cancers de la tête et du cou, une amélioration statistiquement significative n’a été observée que pour un seul des volumes cibles anatomocliniques ; plus de la moitié des valeurs du coefficient de similarité de Dice étaient supérieures à 0,6 dès la première comparaison. Cinq des six volumes cible anatomocliniques ganglionnaires mammaires ont été améliorés. L’étude dosimétrique du cancer de la prostate a permis une homogénéisation des délinéations.

**Conclusion.** – Ce travail a permis d’harmoniser les pratiques autour de référentiels validés. Une extension à l’ensemble des Hauts-de-France est envisagée.


---

**1. Introduction**

During the period from 2005 to 2009, cancer was the leading cause of death in France for men and the second leading cause of death for women. In total, 385,000 cases were diagnosed in 2015, and cancer is the leading cause of premature death before age 65 regardless of sex [1]. In the Lille proximity area, there is a 20% higher rate of cancer in men and a 16% higher rate in women compared to the national rate. Specifically, cancer rates in the Lille proximity area are higher than national rates for localizations linked to tobacco and alcohol consumption. In women, breast, colorectal, ovarian, cervical, lip-mouth-pharyngeal and bladder cancers are also present at higher rates [2].

Radiation therapy is a treatment for cancer, whether it is for curative or palliative purposes. In 2015, more than 204,000 patients were treated in 172 French centres [3]. The departments of Nord and Pas-de-Calais comprised eleven centres. Treatment techniques have evolved considerably in recent years with the introduction of intensity-modulated radiation therapy and stereotactic radiotherapy. Morphological and metabolic imaging has played a major role in treatment planning. The delineation of interest volumes (target volumes and organs at risk) is a crucial planning step, with inadequate delineation leading to a loss of treatment efficacy due to the lack of target coverage and/or increased risk of acute and late toxicity.

Delineation guidelines have been developed for most clinical situations, and training is provided by scientific societies, including the French society of radiotherapy and oncology (Société française de radiothérapie oncologique, SFRO) [4–7]. Nevertheless, very little work has evaluated the appropriation of these standards by professionals in practice. The objective of our work is to improve the homogeneity of delineation between oncologist-radiotherapists in the regions of Nord and Pas-de-Calais (France) through discussions of clinical cases and the adoption of common published reference documents. The first results on prostate cancer have shown a quantifiable improvement in this harmonization [8], and a three-year follow-up assessment is presented here.

---

**2. Materials and methods**

All eleven centres were involved: eight private centres, two public-private joint centres and one university centre. This work was financed by the regional council, the European Regional Development Fund (ERDF), the general interest mission of the regional health agency, the Cancer League, the centres themselves and was coordinated by the Regional Cancer Network (RCN). The localizations focused on in this project included prostate, head and neck, breast, brain and lung cancers. Lung cancer localization practices are currently being evaluated. A quality engineer (C. Viot) integrated into the RCN team has been in charge of project management, logistics engineering and statistical analysis of regional data.

For each of the locations studied, one or more clinical cases were sent to the centres along with the corresponding clinical, anatomopathological and radiological data. Nuclear medicine (positron emission tomography) data were also included in the clinical lung cancer case. For each clinical case, the physician (5) at each centre in charge of pathology delineated the volumes of interest according to their usual practices. This delineation could also be carried out by residents in the university department or a department participating in the training. The contours created were anonymized and sent via a P2E console (Aquilab SAS) to the RCN for evaluation. Validated indices are used to quantify the delineation differences [9–12] (Table 1). The scope of this article is to present the general approach and key results of the work rather than to discuss all the results in detail. Dice similarity coefficient results for most of the target volumes are reported. Organs at risk results will not be reported here.

The results were presented at two to three annual meetings of the association that initiated the work (http://www.onco-npdc.fr/partenaires/cron). The delineations were discussed in detail, slice by slice. At the end of these meetings, a validated and published delineation reference system was chosen. A second delineation of the clinical cases was then carried out to quantify the homogenization of the practices using the same indices. An evaluation of the dosimetric practices was also conducted for
intermediate risk to prostate cancer and breast cancer using a similar methodology; these evaluations are in progress. The same clinical cases were sent to all centres with a predefined delineation. At least one dosimetry was performed per centre, with a prescribed dose and margin imposed on all. A second dosimetry was performed after the presentation of the results for comparison of the dose–volume histograms. The Wilcoxon assay was used on the matched data. Statistical analyses were performed using JMP software (Version 10; SAS Institute Inc., SAS Campus Drive, Cary, North Carolina). A P-value of less than 0.05 was considered to be statistically significant.

3. Results

The locations studied were the prostate, head and neck, breast, and glial brain tumours. In total, nearly 250 cases were delineated across the 11 centres. Only the Dice similarity coefficient values for the target volumes (clinical target volume) are presented here. Agreement was considered medium, satisfactory or excellent for Dice similarity coefficient values between 0.41–0.60, 0.61–0.80 and 0.81–1.00, respectively.

3.1. Prostate cancer

The prostate cancer results have been previously published [8]. The indices were not improved for case 1 (low-risk recurrence cancer), for which the match was found to be satisfactory in the first comparison (DSC = 0.83 ± 0.06). The second comparison of case two (pelvic lymph node area) resulted in an increase in Dice similarity coefficient value (0.58 ± 0.09 vs. 0.63 ± 0.07; P = 0.0098). The Dice similarity coefficient value was not improved for case 3 (prostatec-tomy bed), but its value was already greater than 0.6 in the first comparison (0.63 ± 0.08).

3.2. Head and neck cancer

Three head and neck cases were studied (cases 1 to 3: oropharynx, cavum and pharyngolarynx). Cases 1 and 3 were classified as T3N2bM0; case 2 was classified as T1N2cM. Computed tomography (CT) scan was realized with biphasic contrast medium injection. Three clinical target volumes (high, intermediate and low risk) were delineated for each of these cases. For the second delineation, the recommendations published by Lapeyre et al. were imposed for the delineation of low-risk clinical target volume [13,14]. The results of the Dice similarity coefficient values are presented in Table 2 and Figs. 1 and 2. A statistically significant improvement was found for the low risk clinical target volume in case 2.

3.3. Breast cancer

Three breast cancer cases were studied:

• case 1: breast with axillary involvement after partial mastectomy and axillary clearance (target volumes to be delineated: breast, surgical bed, lymph node levels I to IV, internal mammary chain, interpectoral node according to [15]);
• case 2: breast with axillary involvement after total mastectomy and axillary clearance (target volumes to be delineated: breast, surgical bed, lymph node levels II to IV, internal mammary chain, interpector-al node);
• case 3: breast without indication of boost or irradiation of the lymph node areas. CT scan carried out with and without cerclage of the mammary gland.

The results of the Dice similarity coefficient for the node clinical target volume in cases 1 and 2 are presented in Table 3 and Fig. 3. The largest discrepancies were observed at the external and caudal boundary of area II, and at the cranial and caudal boundaries of areas III and IV. The Dice similarity coefficient values were low for the internal mammary chain, probably because of their very small volume. Five of the six lymph node volumes were improved.

3.4. High-grade glioma

For the brain cancer case, the Dice similarity coefficient values of the clinical target volumes were 0.84 ± 0.08 and 0.87 ± 0.04 in the first and second comparisons, respectively.

3.5. Prostate cancer dosimetry

All centres sent out calculated treatment plans with intensity modulation for the prostate cancer cases. For case 1, the comparison of the mean dose–volume histograms showed a significant decrease in the area under the curve for the rectum and bladder

| Table 1 |

| Study on the harmonization of practices between radiotherapy centres in the Nord and Pas-de-Calais regions (France): indices used to compare delineations. |

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition</th>
<th>Optimal value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume ratio</td>
<td>V2/V1</td>
<td>1</td>
</tr>
<tr>
<td>Dice similarity coefficient</td>
<td>2 × (C2 ∩ C1)/(C2 ∪ C1)</td>
<td>1</td>
</tr>
<tr>
<td>Volume overlap</td>
<td>C2 ∩ C1</td>
<td>1</td>
</tr>
<tr>
<td>Common volume</td>
<td>C2 ∩ C2</td>
<td>100%</td>
</tr>
<tr>
<td>Additional volume</td>
<td>C2 ∪ C2</td>
<td>0%</td>
</tr>
</tbody>
</table>

V1: Contour volume to be compared; V2: Reference contour volume; C2: contour to be compared; C1: reference contour.

Please cite this article in press as: Pasquier D. et al. Harmonization of practices between radiotherapy centres in the Nord and Pas-de-Calais regions (France): a three-year evaluation. Cancer Radiother (2018). dx.doi.org/10.1016/j.canrad.2018.03.007
Fig. 2. Study on the harmonization of practices between radiotherapy centres in the Nord and Pas-de-Calais regions (France): superposition and concordance (hot colours) of the low-risk clinical target volume of case 3 (pharyngolarynx cancer classified as T3N2bM0). (a) first comparison; (b) second comparison.

between V5 and V55 and V19 and V67, respectively (P<0.05) (Figs. 4, 5 and 6). The results were the same for case 2, with a significant difference of V10 to V48 for the bladder (P<0.05), without an increase in the dose delivered to the small intestine. For case 3, a decrease in the irradiated volumes (bladder and rectum) from V20 to V50 was not significant (P=0.05 for V50). In all three cases, the D2% of the femoral heads remained below 50 Gy and the best sparing of the organs at risk was observed without altering the coverage of the target volume.

4. Discussion

Treatment techniques have evolved with the widespread use of intensity-modulated radiation therapy. In 2013, 68% of centres reported that they practiced this technique compared to 20% in 2009. Stereotactic radiotherapy techniques have also been progressing (+10% in 2015 compared to 2009), whether intra- or extracranial [3]. It has been shown that the quality of treatment planning influences the outcome of the treatment in head and neck cancers. The two-year freedom from locoregional failure and overall survival of patients with inadequate treatment plans were significantly lower (50% vs. 70%; hazard ratio [HR] = 1.99; P=0.001 and 54% vs. 78%; HR = 2.37; P=0.001, respectively). One of the major causes of quality deterioration in these plans was the incorrect definition of target volumes [16]. In a meta-analysis of eight studies, the frequency of quality assurance gaps ranged from 8% to 71%. These differences in treatment quality were associated with a statistically

<table>
<thead>
<tr>
<th>Case 1</th>
<th></th>
<th>Case 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IMC CTV</td>
<td>Area IV CTV</td>
<td>Area I-III + IN CTV</td>
</tr>
<tr>
<td>Comparison 1</td>
<td>0.15 ± 0.06</td>
<td>0.18 ± 0.06</td>
<td>0.62 ± 0.20</td>
</tr>
<tr>
<td>Comparison 2</td>
<td>0.39 ± 0.15</td>
<td>0.38 ± 0.08</td>
<td>0.72 ± 0.22</td>
</tr>
</tbody>
</table>

IMC: internal mammary chain; CTV: clinical target volume; IN: interoperative node. In bold: statistically significant difference (P<0.05).
significant decrease in overall survival (HR = 1.74; 95% confidence interval [CI] = 1.28 to 2.35; \( P < 0.001 \)) [17].

The aim of this work was to improve the homogeneity of the delineation practices in the 11 radiotherapy centres in the Nord and Pas-de-Calais region. We did not seek to further describe interobserver variation, as this topic has already been thoroughly discussed in the literature. Rather, we aimed to highlight the results of this collaborative work across the Nord and Pas-de-Calais region. This work involves both residents and senior physicians, and does not replace the delineation and training workshops provided by scientific societies, such as the SFRO, the Association de formation continue en oncologie radiothérapie and the European Society for Radiotherapy and Oncology [4–6]. Our work showed an improvement in the delineation concordance for most target volumes. For prostate cancer, the Dice similarity coefficient values were greater than 0.6 in cases 1 and 3 in the first comparison and after the second comparison in case 2. The indices were not improved in case 1, where the concordance was judged satisfactory from the outset. The delineation of the target volumes for head and neck cancers is among the most complex. Statistically significant improvement was observed for only one of the clinical target volume. Nevertheless, more than half of the Dice similarity coefficient values were greater than 0.6 from the first comparison. Two clinical cases of breast cancer with node involvement were studied, and five of the six lymph node volumes were improved. The Dice similarity coefficient values are very low for the internal mammary chain volumes, and this coefficient is sensitive to small deviations, especially for smaller volumes. It is interesting to note that the volumetric indices used in our study to compare the clinical target volumes are more sensitive than the metric measures. For example, the volume overlap, which is close to the Dice similarity coefficient, is 0.74, with two volumes overlapping at 85%. The volume overlap of two cubes composed of \( 10 \times 10 \times 10 \) voxels after moving a voxel on the diagonal of the cube was 0.57 (729/1271), while the average distance between the two cubes was about one voxel [12]. This may explain the

---

**Fig. 4.** Study on the harmonization of practices between radiotherapy centres in the Nord and Pas-de-Calais regions (France): rectal dose–volume histograms of the rectum for case 1. (a) first comparison; (b) second comparison.
lower Dice similarity coefficient values observed for breast cancer.

To our knowledge there is no comparable work in France in the field of radiotherapy. Some studies have shown an improvement in delineation after training residents in radiation therapy [18,19]. A homogenization of the delineation of head and neck cancers was observed among eleven residents after teaching. In this study, the evaluation by teachers was subjective with no metric or volumetric index [19]. In another teaching study focused on the radiography of the head and neck, fifteen residents in oncology-radiotherapy reported being more confident in delineation. Nevertheless, an objective improvement in delineation by the Dice similarity coefficient was observed for only three of the 20 structures of interest [20,21]. This multidisciplinary Canadian national training program was repeated, involving 29 residents from ten universities. The studied locations were the head and neck, spine, abdomen and pelvis. Knowledge was assessed by multiple choice questions and was significantly improved (average of correct answers before and after training: 60% vs. 80%, P < 0.001). The median improvement in the Dice similarity coefficient value was 0.2 for all locations (P < 0.001). A statistically significant increase in the Dice similarity coefficient value was observed for ten of the 18 structures of interest [22]. It is important to note that this evaluation concerned the definition of structures of interest (clivus, sphenoid sinuses, cochlea, stomach, bladder, etc.) whereas the results presented here address the more complex delineation of the clinical target volume.

Our work allows us to harmonize the delineation and dosimetry practices around the validated standards. An extension to include the entire Hauts-de-France region is being considered.

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgements

The authors wish to thank C. Viot, Dr A. Auliard, Dr N. Blanchard, Dr D. Carlier, Dr A. Henni, Dr H. Koutou, Dr X. Liem, Dr JF Minne, Dr P. Meyer, Dr T. Mulliez, Dr N. Rezvoy, Dr T. Leroy, Dr F. Vasseur, E. Bouché, Y. Boudam, C. du Faux, R. Hardhuin, T. Lacornerie, L.H. Lelièvre, M. Levert, S. Libert and S. Picot.

References

