

Comparaison de scores géométrique et dosimétrique pour la mise en correspondance inter-patients en radiothérapie de la prostate

Gaël DRÉAN^{1,2}, Oscar ACOSTA^{1,2}, Jean-Jacques BELLANGER^{1,2}, Antoine SIMON^{1,2}, Guillaume CAZOULAT^{1,2}, Pascal HAIGRON^{1,2}, Renaud DE CREVOISIER^{1,2,3}

¹INSERM, U642, Rennes, F-35000, France

²Université de Rennes 1, LTSI, F-35000, France

³Département de Radiothérapie, Centre Eugène Marquis, Rennes F-35000, France

gael.drean, oscar.acosta, jean-jacques.bellanger, antoine.simon, guillaume.cazoulat, pascal.haigron@univ-rennes1.fr, r.de-crevoisier@rennes.fnclcc.fr

Résumé – Comparer les matrices de dose ou les organes segmentés de différents patients traités par radiothérapie implique une mise en correspondance dans un espace commun de référence. Les techniques de recalage sont alors généralement validées par des calculs de métriques de superpositions spatiales d’organes segmentés comme la similarité de Dice. Néanmoins, une telle métrique semble insuffisante pour valider une mise en correspondance de dose. Nous proposons une métrique complémentaire reposant aussi sur les distributions de dose : la *Différence Relative d’Aire* et l’avons testé dans deux cadres expérimentaux (Fantômes numériques et Données réelles). Les résultats en zone de gradient de dose montrent des différences, indétectables par la similarité de Dice, apparaître grâce à la métrique proposée.

Abstract – Comparing dose distributions or delineated organs from different patients treated by radiotherapy implies a mapping in a common template. Registration methods are classically validated by using spatial overlap metrics between delineated organs like the Dice Score. Nevertheless, such a metric stay insufficient to evaluate the dose mapping accuracy. We suggest a metric relying on the dose distributions : the *Relative Difference of Areas* (RDA). We tested it in two experimental frameworks (Numerical Phantoms and Actual Data). The results in heterogeneous dose areas show differences thanks to the RDA, differences not shown by the Dice Score.

1 Introduction

La problématique essentielle dans le cadre du traitement du cancer de la prostate par radiothérapie est double : maximiser la dose reçue par la cible clinique - ici la prostate et les vésicules séminales - tout en minimisant celle touchant les organes sains voisins (organes à risques) - vessie et rectum - et éviter ainsi des effets secondaires indésirables de toxicité. La prédiction de la toxicité peut avoir un impact direct sur le planning du traitement en permettant d’estimer, et donc de comparer, la toxicité en fonction de différents plannings de dose calculés sur des *Treatment Planning System* (TPS) utilisant des méthodes de Monte Carlo. Actuellement, cette prédiction repose sur des modèles de toxicité qui se basent principalement sur les *Histogrammes Dose-Volume* (HDV). Un HDV normalisé représente le pourcentage de volume d’une structure recevant au minimum une certaine quantité de dose, ramenant alors la distribution tridimensionnelle de la dose à une représentation 2D [1] et éliminant ainsi toute l’information spatiale de la distribution de dose. Afin de localiser de manière plus précise des zones à risques dans chaque organe, il est nécessaire de comparer les matrices de dose 3D entre les individus. Pour ce faire, les images des organes et les matrices de dose doivent être préalablement recalées dans un même espace commun de

référence. Une récente étude utilise une analyse statistique de la dose 3D entre deux populations (avec/sans événements de toxicité) pour localiser les zones à risques [2]. Néanmoins, les méthodes utilisées pour recalculer les images des organes dans un même espace restent approximatives en termes de localisations spatiales. Dans une perspective d’analyse statistique fine de la distribution de la dose, des méthodes de validation du recalage relatives à la superposition des organes et à la dose s’avèrent alors nécessaires. La plupart des études valident leurs techniques de recalage en utilisant une métrique de superposition spatiale entre deux organes (organe recalé / organe de référence) comme la *Similarité de Dice*. Cependant, des mesures objectives de déformation et de mise en correspondance de la dose restent encore à explorer, comme par exemple celle présentée par Ebert [3]. Cette problématique est d’autant plus importante lorsqu’il s’agit d’organes à risques localisés dans des zones de gradient de dose. Une telle situation est illustrée dans la figure 1 sur deux scénarios différents ① et ②. Lorsque l’organe d’un individu (O_i) est recalé vers l’organe cible (dans le système commun de référence), l’erreur de recalage après une transformation $T1$ ou $T2$ se traduit par un déplacement dx dans deux directions opposées. Mais bien qu’une mesure de superposition spatiale produise le même résultat, la dose serait mappée différemment dans les deux cas du fait du gra-

dient de dose $\Delta D/dx$. Dans cet article, nous proposons une métrique dosimétrique basée sur l'hypothèse de la conservation de l'HDV après transformation. Nous la comparons ensuite à une métrique de superposition spatiale (la Similarité de Dice) afin de constater l'apport de l'utilisation de la métrique proposée. Pour ce faire, des mesures quantitatives ont été réalisées au moyen de fantômes numériques ainsi que sur des données réelles après recalage élastique.

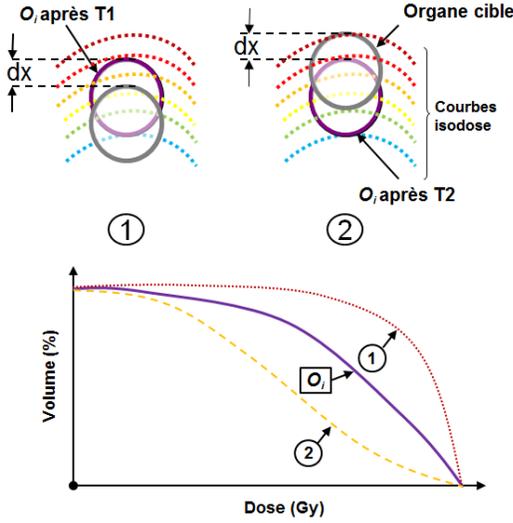


FIGURE 1 – Deux situations de recalage produisant un même décalage dx (même score de superposition spatiale) entre l'organe cible et l'organe (O_i) avec des transformations $T1$ et $T2$ différentes et engendrant des distributions de dose différentes sur l'organe cible.

2 Méthodes

2.1 Métriques

2.1.1 Mesure de recouvrement spatial : Dice Score

Les mesures de recouvrement spatial [4] estiment le recouvrement de deux volumes binaires. Cette coïncidence peut se mesurer par la similarité de Dice, définie ici par le rapport entre le double du nombre de voxels en commun entre les deux structures et le nombre total de voxels, soit : $Dice = \frac{2|A \cap B|}{|A| + |B|}$

, où $|\cdot|$ indique le nombre de voxels du volume considéré, et où A et B représentent l'ensemble des voxels des volumes.

2.1.2 Mesure de conservation de la dose : distance entre deux histogrammes avec la Différence Relative d'Aires

Dans l'espace commun de référence, un recalage parfait engendre une similarité de Dice de 1 et des HDV h_1 (Organe Patient Propagé - Dose Patient Propagée) et h_2 (Organe Cible - Dose Patient Propagée) superposés. En revanche, si le recalage n'est pas parfait, la similarité de Dice sera inférieure à 1 et les HDV précédents seront différents du fait d'une mise en corres-

pondance de dose différente (Figure 1). Cette différence peut être calculée en évaluant la différence des HDV normalisés (Organe Patient Propagé, Dose Patient Propagée) et (Organe Cible, Dose Patient Propagée) dans l'espace de référence. Pour cela, nous proposons alors une mesure d'écart : la *Différence Relative d'Aires* (DRA) entre deux histogrammes Dose-Volume h_1 et h_2 , affines par morceaux, définis tous les deux sur un même intervalle $[0, D_{max}]$, et à valeurs dans $[0, 100]$ (histogrammes normalisés avec un volume total égal à 100) :

$$DRA = \frac{\int_0^{D_{max}} |h_1 - h_2| dx}{\max \left\{ \int_0^{D_{max}} h_1 dx, \int_0^{D_{max}} h_2 dx \right\}} \quad (1)$$

de manière à ce qu'elle soit normalisée entre 0 (quand $h_1 = h_2$) et 1.

2.2 Méthodes

Le principe global de notre étude apparaît sur la figure 2. Disposant des différentes données patient (organes et matrices de dose) transformées vers un espace commun de référence après un processus de recalage et de propagation, deux mesures de la mise en correspondance entre l'organe cible et l'organe recalé (O_i) sont calculées : la Similarité de Dice et la DRA appliquée aux histogrammes dose-volume normalisés. L'étape de mise en correspondance a été soit simulée sur fantômes numériques, soit calculée sur des données réelles après avoir obtenu la transformation avec un recalage élastique dense entre les images de tomodensitométrie (TDM) du patient $_i$ et du template dans l'espace commun de référence et avoir propagé les organes (prostate, rectum, vessie) et la distribution de dose.

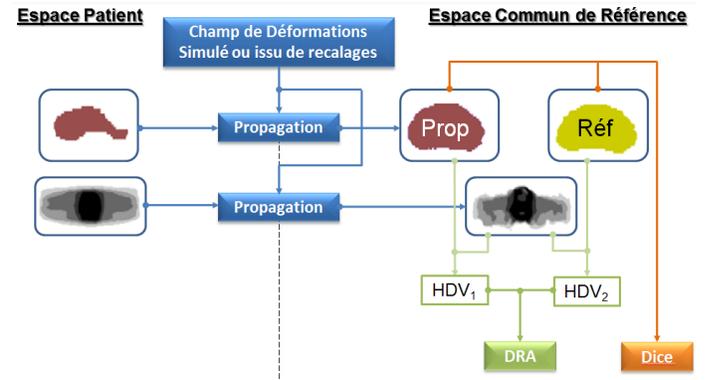


FIGURE 2 – Principe global de validation de la mise en correspondance.

2.3 Expériences

2.3.1 Simulation sur fantôme Numérique

Dans le but de vérifier la pertinence de l'utilisation de la DRA, nous avons procédé à des expériences sur fantômes numériques simplifiés. Nous avons simulé quatre types de situations

hypothétiques post-recalage, à savoir : i) des déplacements et ii) des déformations de l'organe dans des régions de dose homogènes et hétérogènes. Il s'agissait ici d'évaluer l'impact de ces situations en termes de similarité de Dice et de DRA.

Une dose de planification provenant d'un patient réel, obtenue sur TPS, a servi comme base à la simulation. L'organe patient simulé propagé dans l'espace commun était constitué d'une sphère, *Fantôme Numérique Original* (FNO), placée à chaque fois dans des régions de dose homogènes ou de gradient de dose. La première expérience consistait à placer le fantôme sphérique sur l'isobarycentre de la prostate du patient considéré, d'un volume similaire à celui de celle-ci. Ceci fait, un *Fantôme Numérique Mouvant* (FNM) de même volume, simulant l'organe cible, a parcouru, avec un pas de 2 mm, une bande de 100 mm de longueur centrée sur le centre de gravité du FNO. A chaque étape, nous avons calculé le Dice Score entre le FNO et le FNM puis la DRA entre leurs HDV normalisés respectifs, évalués à partir de la dose reçue par le patient. La deuxième expérience était similaire, à ceci près que le FNO était cette fois sur le barycentre de la vessie, c'est-à-dire dans une zone de gradient de dose. L'expérience a été répétée dans les trois directions de l'espace. Quant à la troisième et quatrième études, nous avons repris le fantôme centré sur la prostate et sur la vessie respectivement (dose homogène et hétérogène). Cette fois-ci l'organe cible simulé, centré au même endroit, changeait de volume et prenait une forme d'ellipsoïde 3D, en faisant varier la longueur d'un demi-axe entre la moitié et le double du rayon original avec un pas de 1 mm, et ce, successivement dans les trois directions de l'espace.

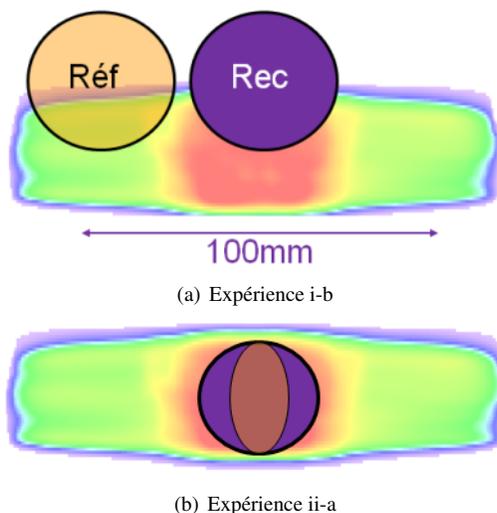


FIGURE 3 – Expériences sur fantômes numériques.

2.3.2 Recalage non-rigide de type FFD sur des données réelles

Nous avons mené une étude sur 24 patients atteints d'un cancer de la prostate et ayant passé un examen TDM de planification. Toutes les images TDM présentent des coupes de 2mm d'une résolution de 512x512 pixels d'1mm dans le plan axial. Pour chacun de ces différents patients, les organes ont été seg-

mentés manuellement par un même observateur expert selon le protocole clinique de radiothérapie. Dans cette étude, seuls seront considérées les images TDM et les segmentations de la prostate, de la vessie et du rectum.

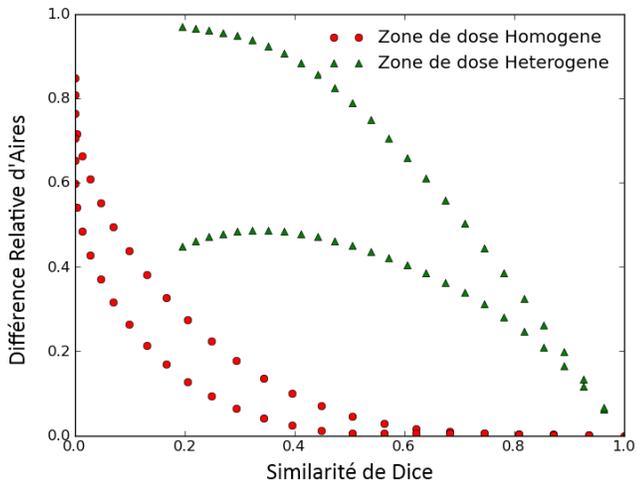
Nous avons suivi une méthode de validation croisée de type « Leave-one-out » : chacune des 24 images TDM ayant été choisie comme Espace Commun de Référence, sur lequel les 23 autres ont été mises en correspondance en utilisant en utilisant une stratégie de recalage dense basé uniquement sur l'intensité des images TDM, stratégie se décomposant en deux étapes principales : (i) un recalage affine « Block Matching »[5] restreint par masque à la segmentation du contour du patient (éliminant ainsi les informations superflues de l'environnement), dont l'algorithme multi-résolution consiste à prendre un bloc sur l'image de référence et de chercher, sur l'image à recalcer, le bloc de même taille présentant la similarité la plus élevée au sens de la corrélation croisée normalisée. (ii) aux résultats du recalage affine, nous avons ensuite appliqué un recalage non-rigide de type « Free-Form Deformation »(FFD) [6, 7], basé sur le calcul multi-résolution de B-Splines cubiques ; la transformation est alors définie à partir d'une grille de points de contrôle, points auxquels seront associées les B-Splines. Finalement, une fois la transformation calculée entre les images patient et de référence, elle a été appliquée aux organes (prostate, vessie et rectum) et à la matrice de dose afin de les propager dans l'espace de référence par interpolation (*plus proche voisin* pour les organes ; *trilinéaire* pour la dose). La DRA a alors été calculée entre les HDV normalisés (Organe patient propagé ; Dose patient propagée) et (Organe de référence ; Dose recalée patient), ainsi que la similarité de Dice a alors été calculée entre un organe du patient de référence et celui recalé d'un autre patient.

3 Résultats

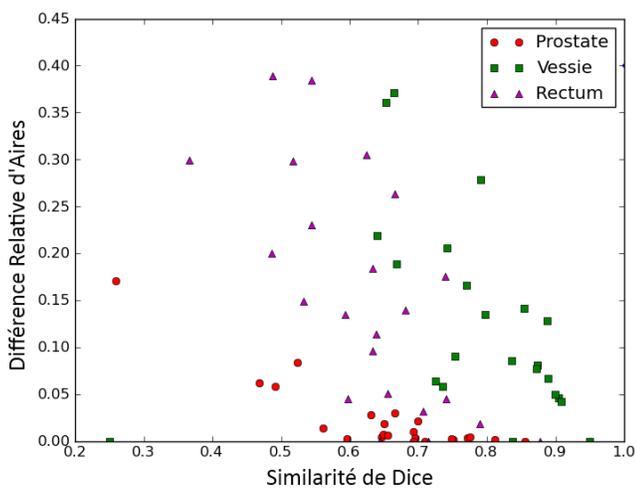
La figure 4 illustre quelques résultats des expériences menées au moyen des fantômes numériques et de l'évaluation de la déformation de la distribution de dose après un recalage de type FFD. Sur la figure 4(a), outre la non-linéarité de la relation similarité de Dice/DRA, nous pouvons constater qu'il peut y avoir plusieurs DRA différentes pour une même similarité de Dice, et ce à la fois pour deux zones différentes comme la prostate/vessie (une même similarité de Dice au niveau de la prostate et de la vessie peut traduire plusieurs DRA différentes), mais également pour une même zone, notamment dans les zones de dose hétérogène comme la vessie. Nous pouvons par ailleurs constater (Fig. 4(b)) que si, dans une zone de dose homogène (prostate), les DRA sont proches de 0 et leurs variances relativement faibles, ce n'est pas le cas dans des zones de gradient de dose. En effet, les DRA et leurs variances sont nettement plus élevées avec une DRA médiane supérieure à 0,2.

4 Conclusion

Au vu des résultats obtenus, il apparaît clairement que dans le cas d'une mise en correspondance de distribution de dose, la seule utilisation de la similarité de Dice ne valide pas complè-



(a) Fantôme numérique : Relation Similarité de Dice/DRA - Expérience I



(b) Données réelles : DRA par Organe pour un atlas

FIGURE 4 – Quelques résultats des expériences sur fantômes numériques et sur données réelles.

tement la déformation de la dose. Ainsi, particulièrement en présence d'un gradient de dose, des différences apparaissent grâce à la DRA, différences indétectables avec la seule similarité de Dice : ces deux métriques sont donc complémentaires. La mesure de validation de la mise en correspondance doit aussi reposer sur un critère dosimétrique. La métrique basée sur les Différences Relatives d'Aires appliquées aux Histogrammes Dose-Volume normalisés décrite dans cet article nous permet de quantifier l'erreur de mise en correspondance de dose pour les études comparatives de populations.

En outre, cette métrique pourrait également être utile à d'autres travaux à venir comme par exemple pour évaluer influence sur les HDV des segmentations manuelles inter-observateur des organes. Une autre problématique sera de s'intéresser aux différents aspects de la conservation, après transformation, de la distribu-

tion de la dose, des Histogrammes Dose-Volume, des Histogrammes Dose-Masse mais aussi d'étudier par exemple l'influence du jacobien de la transformation sur celle-ci.

Parallèlement, nous travaillons également sur des stratégies de recalage élastique dense, iconiques ou hybrides utilisant au moins partiellement l'information apportée par les segmentations des organes, afin d'obtenir des résultats les plus précis possibles tant en termes de similarité de Dice qu'en termes de DRA.

Références

- [1] Ashesh Jani, Christopher Hand, Charles Pelizzari, John Roeske, Lani Krauz, and Srinivasan Vijayakumar. Biological-effective versus conventional dose volume histograms correlated with late genitourinary and gastrointestinal toxicity after external beam radiotherapy for prostate cancer : a matched pair analysis. *BMC Cancer*, 3(1) :16, 2003.
- [2] Wilma D. Heemsbergen, Abraham Al-Mamgani, Mar-nix G. Witte, Marcel van Herk, Floris J. Pos, and Joos V. Lebesque. Urinary obstruction in prostate cancer patients from the dutch trial (68 gy vs. 78 gy) : Relationships with local dose, acute effects, and baseline characteristics. *International Journal of Radiation Oncology*Biolog*y*Physics*, 78(1) :19 – 25, 2010.
- [3] M.A. Ebert, A. Haworth, R. Kearvell, B. Hooton, B. Hug, N.A. Spry, S.A. Bydder, and D.J. Joseph. Comparison of dvh data from multiple radiotherapy treatment planning systems. *Phys Med Biol*, 55(11) :N337–46, 2010.
- [4] Arno Klein, Jesper Andersson, Babak A. Ardekani, John Ashburner, Brian Avants, Ming-Chang Chiang, Gary E. Christensen, D. Louis Collins, James Gee, Pierre Hellier, Joo Hyun Song, Mark Jenkinson, Claude Lepage, Daniel Rueckert, Paul Thompson, Tom Vercauteren, Roger P. Woods, J. John Mann, and Ramin V. Parsey. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *NeuroImage*, 46(3) :786 – 802, 2009.
- [5] S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache. Reconstructing a 3d structure from serial histological sections. *Image and Vision Computing*, 19(1-2) :25 – 31, 2001.
- [6] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations : Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18 :712–721, 1999.
- [7] Marc Modat, Gerard R. Ridgway, Zeike A. Taylor, Manja Lehmann, Josephine Barnes, David J. Hawkes, Nick C. Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine*, 98(3) :278 – 284, 2010. HP-MICCAI 2008.