HDR

intitulée

Intégration d'une plateforme Monte Carlo et Imagerie par Résonance Magnétique pour augmenter la qualité des données médicales dans le cadre de la radiothérapie personnalisée

présentée devant

L'Université de Lille 1 Sciences et Technologies

pour obtenir

Le Grade D'Habilité à Diriger la Recherche

par

Nick Reynaert

Soutenue le

15 janvier 2017, à Villeneuve d'Ascq

Ecole doctorale :
Sciences pour l'ingénieur (SPI)

Laboratoire :
Institut d'Electronique, de Microélectronique, et de Nanotechnologie, (IEMN), USTL, Villeneuve d'Ascq, France

Jury :

Manuel Bardiès	Professeur, CRCT, Toulouse	Rapporteur
David SARRUT	Directeur de recherche, CNRS, Lyon	Rapporteur
Rochdi MERZOUKI	Professeur, Polytech'Lille, Lille1	Rapporteur
Dirk Verellen	Professeur, VUB, université de Bruxelles	Examinateur
Tuami LASRI	Professeur, IEMN, Lille1	Examinateur
Fabrizio CLERI	Professeur, IEMN, Lille1	Garant

Intégration d'une plateforme Monte Carlo et Imagerie par Résonance Magnétique pour augmenter la qualité des données médicales dans le cadre de la radiothérapie personnalisée

Nick Reynaert

Institut d'Electronique, de Microélectronique, et de Nanotechnologie (IEMN) — Avenue Henri Poincaré, 59491, Villeneuve d'Ascq

Centre Oscar Lambret — 3 rue Frédéric Combemale, 59000

Résumé

Les travaux présentés dans ce manuscrit, se situent dans le contexte de la personnalisation des traitements en radiothérapie, qui demande de construire des bases de données avec une qualité élevée (« Rapid learning »), afin d'évoluer vers un système informatique d'aide à la décision clinique (Clinical Decision Support).

Le premier outil qui va sérieusement augmenter la qualité des traitements et des données cliniques est l'imagerie par résonance magnétique (IRM). L'introduction de l'IRM dans chaque étape du traitement et plus spécifiquement dans la planification des traitements va déboucher sur une précision géométrique plus élevée et la possibilité de s'affranchir du scanner (tomodensitomètre) TDM (CT). Plusieurs méthodes de génération des images pseudo-CT à partir de l'IRM sont décrites dans ce manuscrit (3 articles). La première méthode est basée sur la combinaison d'un atlas et des recalages déformables, la deuxième sur une conversion directe des intensités IRM vers CT, utilisant des séquences spécifiques (ZTE, DUTE) utilisant le contenu d'hydrogène des tissus du patient pour définir les sections efficaces des interactions rayonnement-matière.

La deuxième brique, indispensable pour augmenter la qualité des données cliniques, c'est notre plateforme de calcul Monte Carlo utilisée pour la reconstruction de la dose effectivement délivrée au patient. La dose délivrée peut varier significativement de la dose planifiée à cause du positionnement, des mouvements du patient, du mouvement des organes et des déformations en cours de traitement. Il est extrêmement important de corréler les effets cliniques (toxicité et efficacité des traitements) à cette dose délivrée, afin de construire des modèles prédictifs. Deux articles sur ce sujet sont introduits dans ce manuscrit.

Dans une troisième partie, plus fondamentale, deux études sont introduites concernant la prescription et le rapport de la dose prescrite, un paramètre très important dans la standardisation des données. Tout d'abord nous proposons une solution concernant la conversion entre la dose dans le milieu et la dose dans l'eau, et en second, nous démontrons comment il faut calculer et optimiser la dose pour prescrire dans le cadre des traitements des lésions pulmonaires.

Dans le chapitre final les briques sont combinées afin d'expliquer comment continuer la recherche en radiothérapie vers des systèmes d'aide à la décision clinique pour la personnalisation des traitements futurs.

Mots Clés

Imagerie par résonance magnétique, simulations Monte Carlo, rapid learning, Big Data, radiothérapie personnalisée

Integration of a Monte Carlo platform and magnetic resonance imaging to increase the quality of medical data in the framework of personalized radiotherapy

Summary

The work presented in the current manuscript, fits into the context of the personalization of radiotherapy treatments, which requires building databases containing high quality data ("Rapid learning") to evolve towards a clinical decision support system (CDSS).

The first tool which will enormously increase the quality of the treatments and the corresponding data is magnetic resonance imaging (MRI). The introduction of MRI in each stage of the radiotherapy treatment chain and more specifically in the treatment planning will result in a higher geometric precision and the ability to bypass the need of CT scanners in radiotherapy. Several methods of generating pseudo-CT images from MRI data are described in this manuscript (3 papers). The first method is based on the combination of an atlas and deformable image registration methods, while the second uses a direct conversion of MRI intensities to CT Hounsfield units, using specific sequences (ZTE, DUTE) extracting the hydrogen content of the patient's tissue to define cross sections of the interactions between the ionizing radiation and the tissues of the patient.

The second brick, essential to increase the quality of clinical data, is our in-house Monte Carlo platform used for the reconstruction of the actually delivered dose to the patient. This delivered dose may differ significantly from the planned dose because of positioning errors, patient movement, and geometrical changes in between treatment sessions (e.g. patients losing weight). It is extremely important to correlate the clinical effects (toxicity and treatment efficacy) with this delivered dose when building predictive models. Two papers on this subject are introduced in the current manuscript.

In the third part, two fundamental studies are introduced regarding dose reporting and prescription, a very important parameter in the standardization of clinical data. First, we propose a solution for the conversion between dose to medium and dose to water (solving a debate in scientific literature since 2000), and secondly, a paper demonstrating how to calculate, optimize and prescribe dose for lungs treatments, which is again a topic that has led to a long discussion in literature and in clinical practice.

In the final chapter the different bricks are combined to explain how to continue our research in radiotherapy towards clinical decision support systems for the personalization of future treatments.

Key words

Magnetic resonance imaging, Monte Carlo simulations, radiotherapy, rapid learning, Big Data, Clinical decision support, personalized medicine

Table des matières

Liste des abréviations

COL : centre Oscar Lambret

CT (TDM): computerized tomography / tomographie numérique

CTV : clinical target volume / volume cible anatomo-clinique

DQA : Delivery quality assurance (contrôle d'un plan de traitement dans un fantôme)

DICOM : digital image and communications in medicine

DRR : digitally reconstructed radiograph / radiographie reconstruite numériquement

dUTE : double UTE / séquence UTE à double temps d'écho

DVF : deformation vector field / champ de déformation

DVH (ou HDV) : dose volume histogram

DQPRM : Formation nationale des physiciens médicaux en France

FOV : field of view / champ de vue

Gy : unité de dose (Gray), 1 Gy = 1 $J.kg^{-1}$

GTV : gross tumor volume / volume tumoral macroscopique

GUI : graphical user interface / interface utilisateur graphique

HDV (ou DVH) : histogramme dose volume

HU (ou UH) : Hounsfield units

IGRT : image-guided radiation therapy / radiothérapie guidée par l'image

IMRT (ou RCMI) : intensity-modulated radiation therapy

IRM (ou MRI) : imagerie par résonance magnétique

IVDT : image value to density table / courbe d'étalonnage du CT

Linac : linear accelerator / accélérateur linéaire d'électrons

LTSI : laboratoire de traitement du signal et de l'image (Rennes)

MAE : mean absolute error / erreur absolue moyenne

MC : Monte Carlo

MCDE : Monte Carlo dose engine

MCTP : Monte Carlo treatment planning / Planification de traitement Monte Carlo

ME : mean error / erreur moyenne

MLC : multi-leaf collimator / collimateur multi-lames

MR: magnetic resonance

MRI (ou IRM): magnetic resonance imaging

MU (ou UM) : monitor unit

MV: megavolt

MVCT : megavolt CT / tomographie numérique par faisceau megavolt

OAR: organe à risque

ORL: oto-rhino-laryngologique

pCT: pseudo-CT

PET (ou TEP): positron emission tomography

PTV: planning target volume / volume cible prévisionnel

RCMI (ou IMRT) : radiothérapie conformationnelle avec modulation d'intensité

RF : radio-fréquence

RM: résonance magnétique

ROI: region of interest / région d'intérêt

RTP : radiotherapy treatment planning / planification de traitement en radiothérapie

SBRT : stereotactic body radiation therapy / stéréotaxie extra-crânienne

SRS : stereotactic radiosurgery / stéréotaxie intracrânienne

TE : temps d'écho

TDM (CT): Tomodensitomètre

TEP (ou PET) : tomographie par émission de positons

TPS: treatment planning system / système de planification de traitement

TR : temps de répétition

UH (ou HU): unités Hounsfield

UM (ou MU) : unité moniteur

UTE: ultra-short time echo (sequence)

VMAT: volumetric modulated arc therapy / arc-thérapie avec modulation d'intensité

xml : extensible markup language

ZTE: zero time echo (sequence)

Liste des figures

Introduction

Le but de la radiothérapie est d'irradier les cellules cancéreuses tout en épargnant les organes sains autour de la cible. La radiothérapie est souvent faite en combinaison avec la chimiothérapie et/ou la chirurgie, mais peut également être un traitement exclusif. A peu près la moitié des traitements en radiothérapie sont à but curatif. Le reste des traitements sert à pallier les symptômes de la maladie (palliatif). On peut utiliser des faisceaux externes générés par un accélérateur linéaire (Linac) ou des sources radioactives (^{60}Co par exemple). On peut également introduire des sources radioactives dans le corps du patient (intra-cavitaire ou interstitiel). Dans ce travail on ne va considérer que la radiothérapie externe la plus courante, utilisant des faisceaux de photons générés par un accélérateur linéaire.

1. La chaîne de la radiothérapie

Afin de rendre la suite compréhensible, il est indispensable d'introduire les étapes d'un traitement de radiothérapie. Un schéma est montré en figure 1.

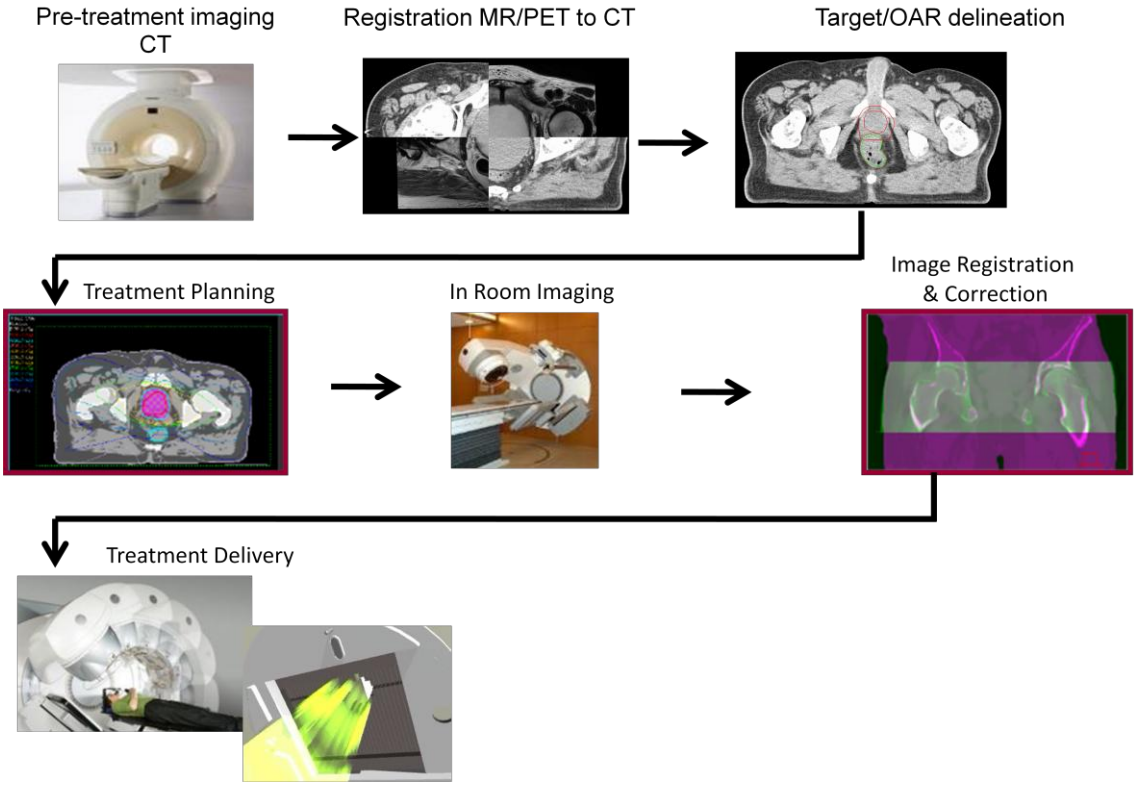


Figure 1 : Chaîne de traitement en radiothérapie. Une fois qu'on a décidé de traiter le patient en radiothérapie (utilisant des images CT, IRM ou TEP diagnostiques) le patient doit d'abord passer par une simulation. Les images du simulateur (CT) sont ensuite combinées avec les images diagnostiques recalées. Puis les volumes cibles et les organes à risque sont délinéés (contourés) par le radiothérapeute et la prescription est faite. Le dossier arrive en dosimétrie où l'équipe de physiciens médicaux et de dosimétristes optimisent le plan de traitement. Le traitement consiste en plusieurs séances. A chaque séance, le patient est positionné par un

recalage entre les images issues du système d'imagerie embarqué et les images prévisionnelles déterminées pendant la dosimétrie. Le patient est finalement irradié.

Les étapes montrées dans la figure 1 sont résumées ici :

- Simulation virtuelle du traitement : Le patient est mis en position de traitement avec ses systèmes de contention (matelas individuel, masque thermoformé, …) afin d'optimiser la reproductibilité du positionnement pendant les séances de traitement. Un scan 3D du patient est pris dans cette position. Des tatouages sont marqués sur la peau du patient pour faciliter son repositionnement (utilisant des lasers). On peut prendre plusieurs scans avec plusieurs résolutions, avec ou sans produit de contraste, ou à plusieurs phases du cycle respiratoire (scan 4D).
- Parce-que le CT n'offre pas toujours un contraste assez élevé, on va combiner les images du simulateur CT avec des images IRM ou TEP avant de délinéer les organes. L'IRM offre un meilleur contraste dans les tissus mous et le TEP et l'IRM offrent des images fonctionnelles de la tumeur (d'abord pour mieux voir la tumeur, mais aussi possiblement pour définir des zones qui sont plus actives dans la tumeur). Les images sont recalées utilisant un recalage rigide ou déformable (élastique) afin de les définir dans le même système de coordonnées.
- Le radiothérapeute va délinéer ou contourer les volumes cibles et les organes à risques qu'il faut protéger pendant le traitement. Souvent les images IRM et/ou TEP sont utilisées pour définir la cible et les contours sont propagés sur le CT. Le médecin fait la prescription : il définit la dose qu'il souhaite au niveau de la cible (objectifs du traitement) et les contraintes à respecter au niveau des organes à risque.
- Le dossier entre en dosimétrie. Les physiciens médicaux et/ou dosimétristes optimisent la dosimétrie en utilisant un logiciel de planification de traitement (le TPS, treatment planning system). Pendant ce processus, dépendant du type de traitement (voir plus loin), le dosimétriste définit des faisceaux de façon semi-automatique et le logiciel calcule la distribution de dose dans le patient. Le patient est modélisé en utilisant les images CT. Les unités Hounsfield (niveaux de gris dans les voxels du CT) sont converties en densité (relation bi-linéaire). En utilisant ce modèle du patient, le TPS peut calculer la distribution de dose au moyen de plusieurs algorithmes (voir plus loin). Puis le dosimétriste optimise cette distribution de dose en modifiant (manuellement ou automatiquement) les poids et les formes des faisceaux. Des histogrammes dose volume (représentant la dose reçue à des pourcentages de volume pour chaque région d'intérêt) sont utilisés pendant l'optimisation et pour l'évaluation de la qualité du traitement.
- Une fois optimisée, la dosimétrie (le plan de traitement) est validée par un physicien et un radiothérapeute et le dossier est introduit dans un logiciel « Record et Verify » qui fait la connexion entre la dosimétrie et les machines de traitement.
- Le traitement est conventionnellement fait en plusieurs fractions pour des raisons radio-biologiques. Les tissus sains récupèrent plus facilement que la tumeur après une irradiation grâce à des mécanismes de réparation de l'ADN des cellules. Alors, au lieu de donner la dose de prescription (par exemple 70 Gy) en une fois, on va irradier 35 fois le patient avec une dose de 2 Gy (35 jours). Pendant chaque séance, il faut correctement positionner le patient, utilisant les lasers/tatouages, les systèmes de contention utilisés pendant la simulation du traitement, et l'imagerie intégrée dans la machine de traitement. Les images obtenues pendant le traitement sont comparées avec les images de simulation pour positionner le

patient. Une fois bien positionné le patient est irradié. Comme on va le voir plus loin, il y a des machines qui vont suivre les mouvements du patient pendant le traitement.

- Ce qui ne figure pas dans le schéma et est expliqué plus loin, sont les contrôles de qualité qui sont appliqués. Dépendant du type de traitement, un deuxième calcul de la distribution de dose, utilisant un algorithme de calcul de dose indépendant, est fait (le « double calcul »), une mesure sur fantôme du traitement est faite, et/ou une mesure de la dose in-vivo (utilisant par exemple des dosimètres au niveau de la peau du patient).

2. La radiothérapie : les différentes techniques

La radiothérapie externe a évolué énormément ces dernières décennies. L'introduction de l'imagerie 3D par TDM (CT, inventé en 1972 par Cormach et Hounsfield) a largement boosté la précision des traitements. Pour la première fois on était capable de visualiser la géométrie de la tumeur et les organes autour en 3D. Mais, il y a encore dix ans, une grande partie des patients étaient encore traités en 2D (utilisant des images de projection en 2D). Ce n'est que depuis 5-7 ans que les simulations des traitements sont systématiquement faites avec un CT-simulateur.

La technique d'irradiation a aussi beaucoup évolué. Au début le patient était irradié avec un nombre de faisceaux limité utilisant des formes de faisceau simples (des champs carrés par exemple). De cette façon un grand volume autour de la cible était irradié avec une dose homogène et beaucoup d'organes sains étaient irradiés avec. Pour éviter des effets négatifs dans ces organes la dose à la cible était limitée et les traitements étaient rarement curatifs. L'introduction du collimateur mutilames (MLC, voir figure 2) a introduit la radiothérapie conformationnelle. Un exemple d'un tel traitement est donné en figure 3.

Figure 2: MLC : Chaque lame peut être positionnée individuellement, offrant la possibilité de sculpter le faisceau conformément avec la géométrie de la cible.

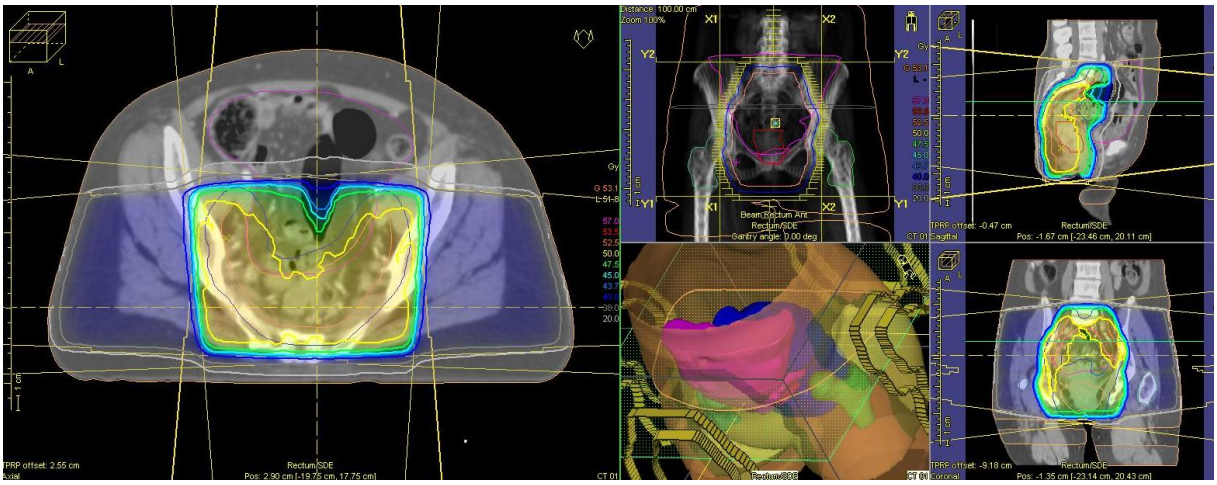


Figure 3: Traitement de radiothérapie conformationnelle utilisant le MLC pour mieux conformer la dose à la cible.

Pour augmenter encore plus la précision, on a introduit la radiothérapie conformationnelle avec modulation de l'intensité (RCMI), montrée à la figure 4.

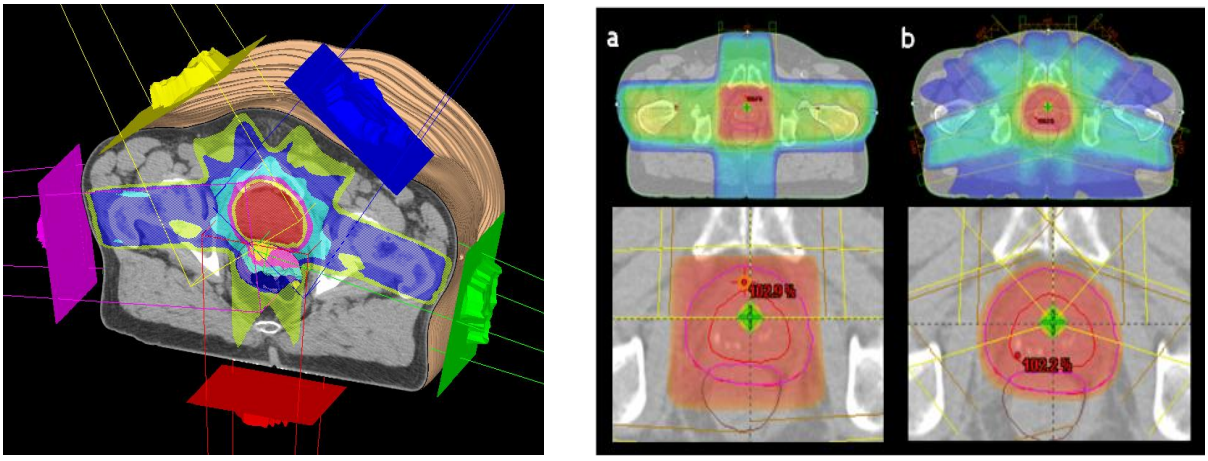


Figure 4: La RCMI : l'intensité de chaque faisceau est modulée en superposant plusieurs faisceaux par direction, afin de sculpter la distribution de dose autour de la cible (traitement prostate) protégeant ainsi la vessie et le rectum de façon optimale. La figure à droite montre clairement l'avantage de la RCMI (à droite) par rapport au techniques conformationnelles (à gauche).

La figure 5 permet une bonne visualisation des différentes techniques de traitements.

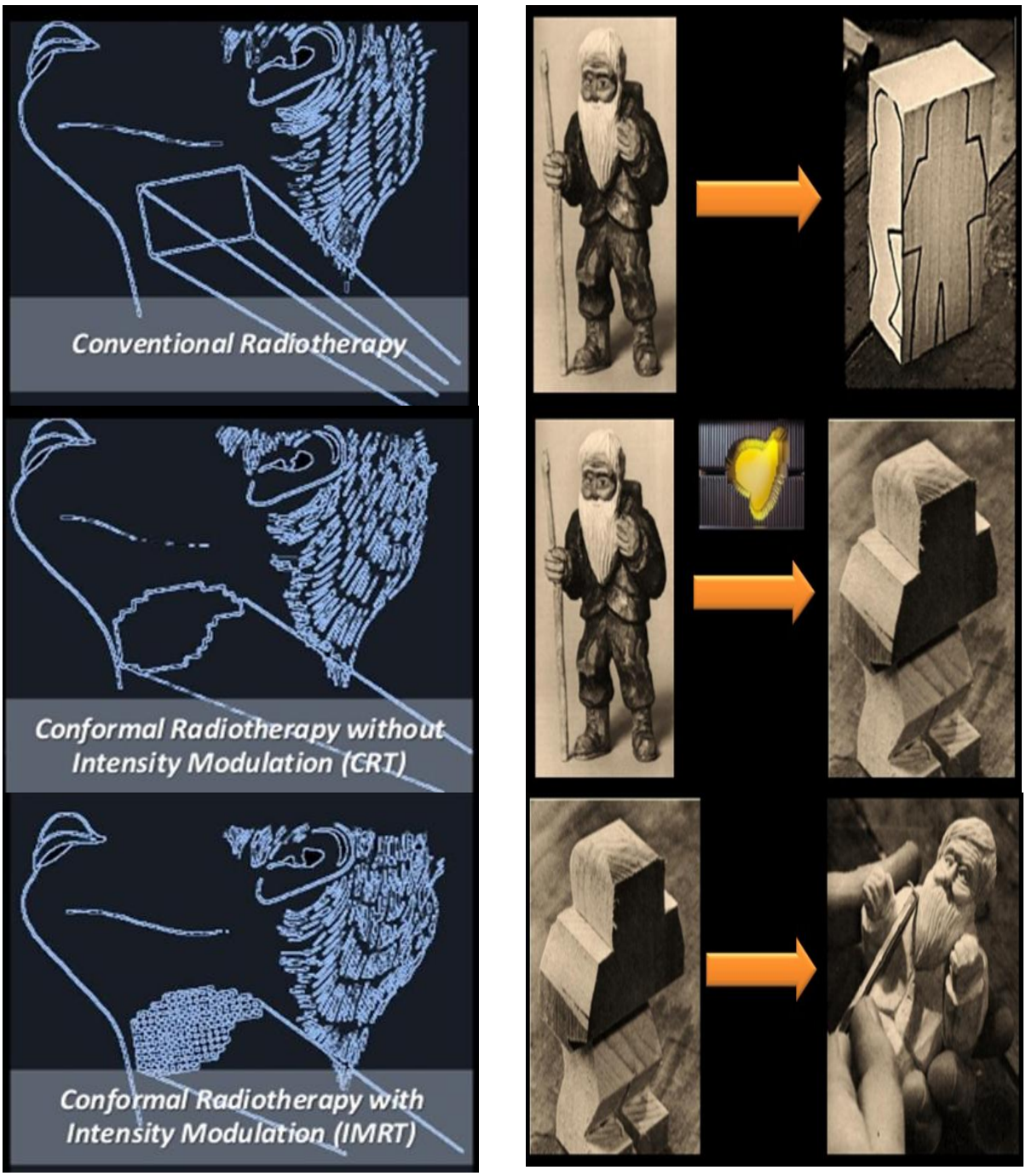


Figure 5 : Une vue alternative des différentes techniques d'irradiation. Du haut en bas on a la radiothérapie conventionnelle, après la radiothérapie conformationnelle et puis la radiothérapie conformationnelle avec modulation de l'intensité.

L'expérience en RCMI a montré que pour certaines cibles la qualité du traitement augmente en fonction du nombre de faisceaux. En limite on peut utiliser un nombre de faisceaux infini utilisant une rotation continue. De cette façon on a introduit la RCMI en Arc (VMAT).

La **Tomothérapie**™ (Accuray) est une machine dédiée à la RCMI. Le tube à rayons X tourne continument autour du patient qui avance petit à petit à travers le faisceau (irradiation hélicoïdale, figure 6). Cette machine est optimale pour les tumeurs allongées dans le sens longitudinal du patient. On peut facilement irradier plusieurs tumeurs au sein d'un seul traitement. Au centre Oscar Lambret à Lille, il y a trois machines de ce type. Cette machine sera donc le sujet de plusieurs études décrites dans ce manuscrit.

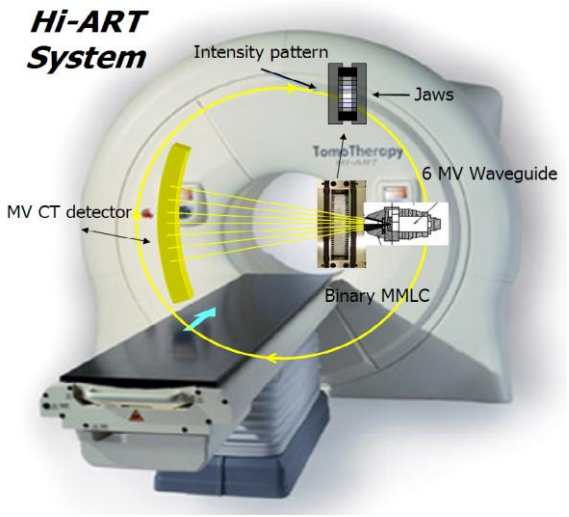


Figure 6: La Tomothérapie. Le Linac tourne en continu autour du patient pendant que la table avance à travers le champ d'irradiation. Ce champ est délimité par des mâchoires et un MLC binaire. L'intensité de chaque faisceau est modulée en utilisant le temps d'ouverture individuel de chaque lame.

Déjà depuis plusieurs décennies, il y avait la volonté d'irradier de petites cibles dans le crâne. Ceci demande une précision géométrique encore plus élevée qu'en RCMI, ce qui a mené à l'introduction de la stéréotaxie intracrânienne où on utilise de petits faisceaux (5-20 mm). Une option est d'installer un micro MLC (des lames plus fines qu'en RCMI) dans un Linac conventionnel. Une autre alternative est d'utiliser des machines dédiées comme le Gammaknife™ (Elekta), le Cyberknife™ (Accuray), le Novalis™ (Brainlab), … L'irradiation de petites cibles, où on irradie peu de tissu sain, offre aussi la possibilité de limiter le nombre de séances (hypofractionner). Au lieu de fractionner le traitement en 30 séances (ce qu'on fait conventionnellement, pour que les tissus sains puissent récupérer entre les séances par réparation des dommages de l'ADN) on donne la dose totale en 3 séances par exemple (ou même en une seule pour le Gammaknife). Cette idée est aussi intéressante pour des cibles radio-résistantes à l'extérieur du crâne. C'est la stéréotaxie extra crânienne (SBRT ou Stereotactic body radiotherapy).

Depuis 2007 un **Cyberknife** est installé au Centre Oscar Lambret à Lille. Cette machine est montrée à la figure 7. Le Linac (6 MV) est monté sur un robot Kuka et peut bouger autour du patient avec 6 degrés de liberté. Plusieurs systèmes de collimation sont disponibles. Il y a les collimateurs fixes (12 cônes avec un diamètre entre 5 et 60 mm). Le collimateur Iris (diaphragme) a les mêmes ouvertures que les collimateurs fixes mais il a l'avantage que l'on peut modifier le diamètre de l'ouverture sans entrer dans la salle et sans devoir changer de collimateur. En principe l'Iris offre la possibilité d'utiliser une gamme d'ouvertures continue mais cette option n'est pas prévue dans le système de

planification. Dans la nouvelle génération de Cyberknife (déjà commandé mais pas encore installé au Centre Oscar Lambret à Lille) un MLC est introduit comme collimateur. Pour le positionnement initial du patient, un système d'imagerie qui consiste en deux tubes à rayons X montés au plafond et deux détecteurs intégrés au sol, offre deux projections orthogonales du patient. Cette configuration (combinaison système d'imagerie, le robot et les collimateurs) offre une précision millimétrique. Pour définir la position sur les images rayons X, il faut une zone contrastée dans l'environnement de la tumeur. Pour le crâne, la solution est triviale car la géométrie est considérée comme rigide. Pour les autres localisations, la situation est plus compliquée. On utilise souvent les vertèbres (XSight Spine^{TM}) ou des fiduciels implantés dans le foie par exemple. Pour les traitements poumon on peut souvent utiliser le contraste entre la cible et le tissu pulmonaire (XSight Lung^{TM}).

Pour les indications extracraniennes, il faut aussi tenir compte des mouvements du patient pendant l'irradiation. Le système d'imagerie est utilisé pour corriger la position du robot en fonction des mouvements du patient en temps réel. Pour les mouvements corrélés à la respiration (traitements poumon, foie, reins, …) il y a le système Synchrony^{TM} qui consiste à positionner des diodes (des markers externes) au niveau de la poitrine du patient en combinaison avec un détecteur infrarouge monté au plafond au niveau des pieds du patient (voir figure 7). Parce qu'on ne peut pas utiliser le système d'imagerie rayons X en temps réel (les tubes chaufferaient trop et on donnerait trop de dose au patient), on va, avant le traitement, générer un modèle de corrélation entre les mouvements internes visibles sur les images rayon X et le mouvement des diodes. Ce modèle est vérifié régulièrement et corrigé (si nécessaire) pendant le traitement.

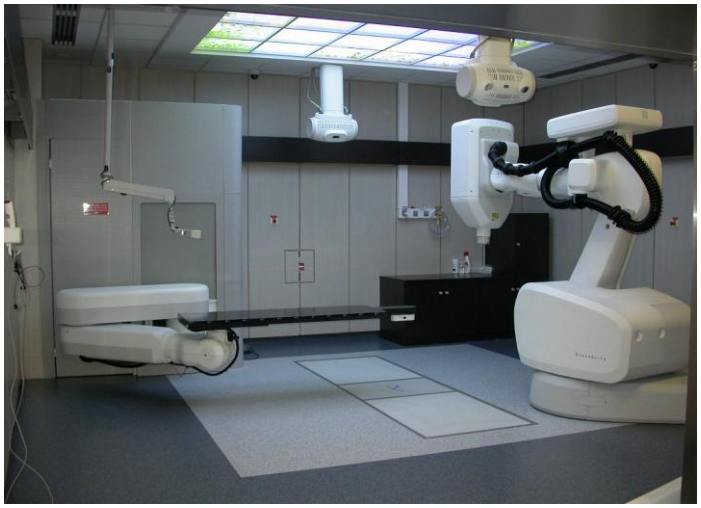


Figure 7: Salle de traitement du Cyberknife : Le Linac est monté sur un robot Kuka et peut bouger avec 6 degrés de liberté autour du patient sur des nœuds prédéfinis. La table est robotisée avec 5 degrés de liberté. Des images de projection sont prises utilisant les deux tubes rayons X montés au plafond et les deux détecteurs intégrés au sol.

3. Les différentes techniques de contrôles qualité en radiothérapie

Comme déjà évoqué plus haut, il est important de garantir la qualité des traitements. C'est la responsabilité des physiciens médicaux. Le but est de diminuer les incertitudes qui sont multiples et qui sont énumérées sans être exhaustive, comme :

- Incertitudes liées à la machine de traitement: le débit de dose peut varier, les tailles de champs, la position des lasers, la position du système d'imagerie intégré. Les mouvements de la machine (le bras, la table, …). Pour ces incertitudes on fait des contrôles journaliers, hebdomadaires, mensuels, trimestriels, annuels … qui sont écrit dans des procédures règlementaires. Même une machine qui est parfaitement réglée et contrôlée introduit une incertitude géométrique inhérente qui peut aller de 1 mm jusqu'à 1 cm dépendant de la machine [1].
- Etalonnage de la machine en dose absolue : Les physiciens utilisent des chambres d'ionisation (des détecteurs spécifiques) étalonnées dans des laboratoires primaires ou secondaires. La mesure faite avec une telle chambre doit être convertie en dose utilisant des protocoles comme celui de l'IAEA, l'AAPM, … [2].
- Incertitudes liées au patient : l'anatomie du patient peut changer pendant un traitement qui prend jusqu'à 7 semaines. Le patient peut bouger pendant le traitement, et il va en tout cas respirer. Ces effets peuvent induire une grande différence entre la dose planifiée et la dose effectivement administrée (délivrée).
- Incertitudes liés au TPS : il y a plusieurs paramètres qu'il faut contrôler avant d'introduire un TPS en routine, comme décrit dans plusieurs protocoles [3]. Un ingrédient important ce sont les mesures des profils de dose dans une cuve à eau qui sont faits pendant le « commissioning » (mise en service) du TPS. Ces mesures sont intégrées dans le TPS et directement ou indirectement utilisées par le TPS pour calculer la distribution de dose pour les patients. La qualité de ces mesures est extrêmement importante. En plus, comme déjà évoqué, il y a plusieurs algorithmes de calcul de dose disponibles. Certains TPS utilisent des algorithmes simples qui ne font qu'utiliser les tables des mesures dans la cuve à eau avec des facteurs de correction pour corriger le fait que la géométrie du patient est difficilement comparable avec une cuve à eau. D'autres utilisent des algorithmes beaucoup plus avancés comme le Monte Carlo comme on va le voir plus loin [4]. Et quelle que soit la précision de l'algorithme utilisé il est légalement obligatoire pour chaque plan de traitement (où c'est « techniquement faisable ») de faire un double calcul utilisant un logiciel indépendant.
- Incertitudes liées à la définition initiale de la cible (le GTV ou Gross tumor volume [5]). Souvent le contraste dans les images CT ne suffit pas pour clairement voir la cible. Dans ce cas il y a déjà une grosse incertitude introduite au début de la chaine de traitement qu'on ne peut plus corriger. Dans ce cas, des machines qui sont très précises (comme en SBRT) vont encore induire des erreurs plus importantes en irradiant très précisément à côté de la cible. Une solution est la combinaison du CT avec l'IRM ou le TEP. Mais dans ce cas il y a une incertitude introduite pendant le recalage entre les différentes images qui peut aller jusqu'à 2 mm pour les cas simples (le crâne par exemple) et beaucoup plus dans le thorax ou le pelvis. Par ailleurs une marge est définie autour du GTV pour tenir compte des cellules cancéreuses non-visibles par l'imagerie, qui donne le CTV. Pour la plupart des indications il n'y pas de consensus concernant le CTV. Cet aspect sera discuté plus loin.

La suite de ce paragraphe va focaliser sur les techniques de contrôle qualité pertinentes pour ce manuscrit.

3.1 Le double calcul

Chaque plan de traitement doit être validé par un calcul de dose indépendant, lorsque c'est « techniquement faisable ». En pratique, on voit que tous les traitements conformationnels sont effectivement validés en utilisant des logiciels très simplifiés [6]. Par contre, les techniques plus avancées comme la RCMI ou la SBRT sont beaucoup moins validées parce-que c'est plus compliqué ou qu'il faut des outils avancés qui ne sont pas toujours disponibles. En plus, les logiciels de double calcul utilisés sont souvent tellement approximatifs, que les différences que l'on voit entre le TPS et le double calcul sont toujours causées par le double calcul lui-même. Au centre Oscar Lambret, on utilise le logiciel IMSURE^{TM} (Standard Imaging) pour valider les traitements conformationnels depuis 2012. Pour les traitements Cyberknife (SBRT) et Tomotherapy (RCMI), il n'y avait pas de logiciel disponible. C'est pour cette raison que nous avons développé une plateforme Monte Carlo capable de valider les plans de ces deux techniques en routine. Un logiciel que j'avais déjà développé à l'université de Gand [7] a été complètement reprogrammé pour son utilisation en clinique à Lille (voir plus loin). Il faut néanmoins réaliser qu'un double calcul ne valide que le calcul de dose planifiée et que ça ne teste pas du tout la machine.

3.2 Les DQAs

Pour les traitements RCMI et SBRT il y a la possibilité de copier le plan de traitement sur un fantôme, composé de plastique tissu équivalent dans lequel on peut introduire un détecteur, comme une matrice de chambres (voir figure 8).

Figure 8 : Exemple d'un fantôme (PTW) utilisé pour les validations des plans (les DQAs) en Tomothérapie.

Le plan est copié sur le fantôme homogène et la distribution de dose est recalculée avec le TPS. On va ensuite mesurer la dose dans le plan de la matrice de chambres d'ionisation et comparer les deux distributions (dans le TPS ou dans un logiciel indépendant), comme montré à la figure 9. Cet exemple est fait en utilisant le TPS. On a la possibilité de comparer des profils de dose, ou de faire un test gamma. Ce test valide en même temps le fait que la machine peut correctement délivrer le plan et que le TPS peut correctement calculer la dose pour ce plan. Par contre, parce-que le fantôme est homogène, on ne démontre pas que le TPS calcule la dose correctement dans le patient, où la géométrie est plus compliquée et plus hétérogène. Au centre Oscar Lambret on valide encore

systématiquement tous les patients traités en Tomothérapie et ce, déjà depuis 7 ans. Pour le Cyberknife, on ne suit pas cette procédure par manque de temps (la machine est très occupée à Lille), mais là on va plutôt régulièrement faire des tests end-to-end qui testent les différents systèmes de suivi des mouvements (XSight Lung, XSight Spine, …).

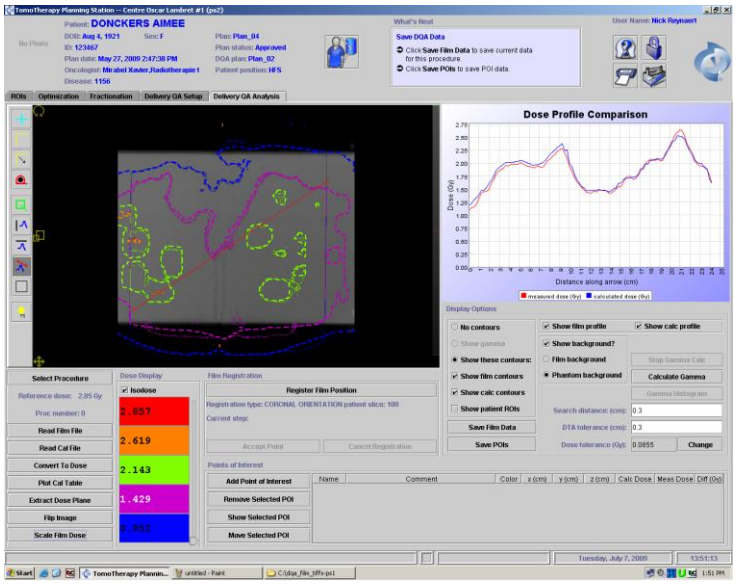


Figure 9 : Comparaison de la mesure dans le fantôme avec le calcul du TPS pour un traitement en Tomothérapie.

3.3 Reconstruction de la dose délivrée

Afin de mieux comprendre les effets des traitements il est important de précisément connaître la dose qu'on a réellement délivrée au patient. Plusieurs effets vont introduire des différences entre la dose planifiée et la dose délivrée :

- La machine ne fait pas toujours exactement ce qui était prévu (petites erreurs de MLC par exemple). Et dans le cas du Cyberknife, le robot va explicitement suivre les mouvements et par définition utiliser d'autres positions que celles planifiées.
- Le patient n'est pas toujours positionné exactement au même endroit. La géométrie interne du patient peut varier entre les séances et pendant les séances.
- Il y des mouvements du patient pendant les séances. Il s'agit de mouvements aléatoires (intestins par exemple) ou de mouvements corrélés à la respiration.
- La géométrie du patient change d'une semaine à l'autre. Les patients traités pour un cancer ORL vont souvent beaucoup maigrir, ce qui peut modifier la dose délivrée. Pour les patients pelvis, on voit souvent une réduction importante du volume cible. Dans ce cas, des organes comme le rectum peuvent entrer dans le champ d'irradiation.

La reconstruction de la dose d'une séance individuelle revient à une reconstruction de la géométrie pendant le traitement (si possible en 4D) et d'une reconstruction des faisceaux effectivement utilisés (et la synchronie entre les deux). Pour déterminer la dose totale effectivement délivrée, il faut accumuler les distributions de dose des différentes séances, ce qui demande l'utilisation de recalages déformables. Plus de détails sont donnés plus loin.

4. Le Monte Carlo en radiothérapie

4.1 Introduction

Parmi les algorithmes de calcul de dose utilisés en radiothérapie, ceux basés sur l'algorithme Monte Carlo sont considérés comme plus précis [4]. Comme tous les algorithmes Monte Carlo, on utilise des nombres pseudo-aléatoires en combinaison avec des distributions de probabilité d'interaction (sections efficaces des interactions entre photons et électrons avec les atomes de la machine et du patient). En dosimétrie le code Monte Carlo va résoudre l'équation de transport des photons/électrons de Boltzmann en utilisant des méthodes stochastiques. Les ingrédients d'une simulation Monte Carlo en dosimétrie (en général et appliqué à la radiothérapie) sont :

- Les interactions photons, électrons – matière (physique atomique)
- Description de la source : modélisation du Linac en radiothérapie
- Description de la géométrie : le patient modélisé en utilisant le CT
- Les scores : détermination de la distribution de dose, incertitudes, spectres

Une simulation Monte Carlo pour la radiothérapie consiste en deux calculs montrés à la figure 10. La source revient à simuler des électrons (ayant une distribution d'énergie et une distribution spatiale) qui frappent une cible de tungstène, où le rayonnement de freinage (photons) est généré.

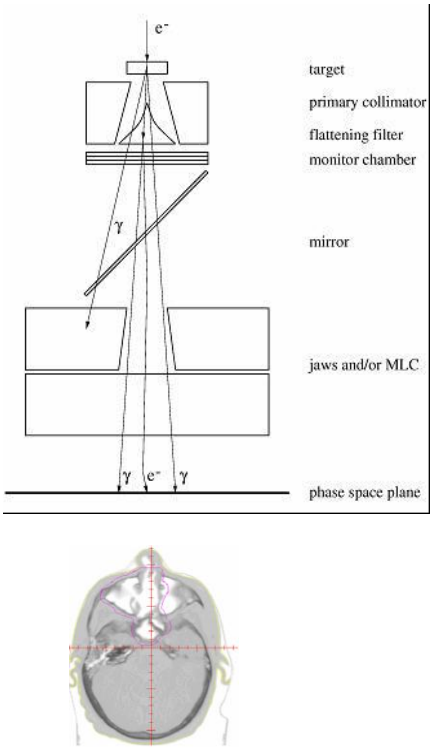


Figure 10 : Les composants principaux d'un calcul Monte Carlo en radiothérapie : simulation de la tête du Linac et du patient.

Puis, ces photons sont suivis dans la géométrie du Linac. Suivant leurs interactions, des particules secondaires sont générées et sont suivies aussi. Le résultat de ce premier calcul est un fichier « phase-space » qui contient les paramètres (énergie, direction, position, poids statistique) de chaque particule qui a traversé ce plan. Ce fichier est utilisé comme input pour calculer la distribution

de dose dans le patient. Le modèle du patient est déterminé en utilisant les images du scanner. Pour chaque voxel, les unités Hounsfield sont converties en densité (relation bi-linéaire) et en composition. Le transport des particules du fichier phase-space dans la géométrie du patient est suivi et l'énergie déposée dans les voxels est accumulée. Le résultat de ce deuxième calcul est la dose en Gy par électron primaire. Un calcul dans un fantôme d'eau pour le faisceau de référence (par exemple un faisceau 10x10 cm^2 à 100 cm de DSP) est simulé pour étalonner le système. Ce facteur d'étalonnage est ensuite utilisé pour convertir la dose par électron primaire obtenue pendant le calcul d'un plan en dose absolue. A cause de leur caractère stochastique, les calculs Monte Carlo demandent beaucoup plus de temps que les algorithmes plus approximatifs utilisés dans les TPS [8][9]. Pour cette raison, beaucoup d'approximations et de techniques de réduction de variance sont introduites dans les algorithmes utilisés en radiothérapie. Des approximations sont appliquées au niveau de la simulation du Linac (des sources virtuelles et l'utilisation du ray-tracing dans les collimateurs) et au niveau du patient. En plus, des filtres (gaussiens) sont appliqués sur la distribution de dose (« denoising ») afin de limiter l'impact du bruit statistique sur les résultats (le bruit peut déformer l'histogramme dose-volume de la cible et donner l'impression que la dose n'est pas parfaitement homogène dans la cible, ce qui est quand-même un des objectifs du traitement). A cause de ces techniques, un système Monte Carlo utilisé dans un TPS reste approximatif pour pouvoir calculer un traitement complet en un temps raisonnable. Pour les systèmes de contrôle qualité, on peut utiliser le Monte Carlo sans trop d'approximations. Le facteur qui va déterminer le temps est le nombre de CPUs disponibles.

4.2 Modélisation des Linacs

La géométrie du Linac (matériaux, dimensions, densités) est donnée par le fournisseur des machines (Accuray dans notre cas pour le Cyberknife et la Tomothérapie). La distribution en énergie et la taille du spot des électrons avant de frapper la cible sont moins connues. Ce sont des paramètres qu'il faut optimiser en utilisant les mesures réalisées dans une cuve à eau. Ceci fait partie de la modélisation de la machine. On utilise des profils latéraux, des rendements en profondeurs et des « output factors » pour un certain nombre de champs (champs ouverts, et petits champs). On va calculer ces fonctions en utilisant le modèle Monte Carlo en variant l'énergie et la taille du spot (deux distribution gaussiennes) jusqu'au moment où les calculs correspondent aux mesures [10].

Cyberknife

La modélisation du Cyberknife est décrite dans l'article de Wagner et al [11] dans lequel nous avons déterminé des facteurs de correction pour les mesures des petits champs du Cyberknife, utilisant la chambre MicroLion™ de PTW. Nous avons d'abord simulé les collimateurs fixes (diamètres de 5 à 60 mm) et après le collimateur IRIS (diaphragme avec diamètre variable). Les simulations sont faites en utilisant BEAMnrc/DOSXYZnrc (EGSnrc) [12][13] et l'IRIS est modélisé avec egspp d'EGSnrc [14]. La géométrie et la composition de tous les composants de la machine ont été fournies par Accuray. La même géométrie est modélisée avec Gate (Geant4) [15] afin d'avoir le choix entre EGSnrc et Gate dans notre plateforme Monte Carlo.

Le Cyberknife M6 est modélisé en collaboration avec le Centre François Baclesse d'Esch sur Alzette (Luxembourg), où ils ont récemment installé cette machine de nouvelle génération. Accuray a de

nouveau fourni la géométrie des collimateurs primaires à forme pyramidale. La machine utilise un MLC pour les traitements RCMI (des champs allant jusque 11.5x10 cm^2). La modélisation du MLC et de la machine en général est faite en utilisant des mesures dans la cuve à eau d'Esch. Pour les simulations des traitements, la difficulté était dans les angulations du collimateur : à chaque nœud, l'angle du collimateur est optimisé en prenant en compte les collisions table-robot, les câbles, le fait de ne pas bloquer le système d'imagerie. Ces données sont aussi fournies par Accuray.

Tomothérapie

Pendant la modélisation de la Tomothérapie, il y avait tout de suite un problème avec la modélisation de l'énergie et de la taille du spot de la machine. Pour la distribution spatiale des électrons primaires une double-gaussienne [16] est utilisée afin d'obtenir des résultats cohérents avec les mesures (les épaules et pénombres des profils pour les champs de 5, 2.5 et 1.0 cm). Cette modélisation est décrite plus en détail plus loin. Pour la simulation des traitements, la définition des temps d'ouvertures des lames en fonction de l'angle du bras doit être bien optimisée. Certains paramètres sont définis au centre de la projection (il y a une projection définie dans le fichier RTPlan, tous les 7 degrés), et d'autres sont définis au début de la projection. Après des résultats cohérents pour les mesures dans la cuve à eau, des tests de faisceaux individuels définis dans le TPS ont été faits pour des densités homogènes. Ce n'est que dans un dernier temps que des plans de traitements ont été recalculés pour une vingtaine de patients, en mixant les indications.

Autres machines

Plusieurs autres machines sont modélisées ou en cours de modélisation afin de généraliser notre plateforme Monte Carlo :

- Le Clinac™ de Varian est modélisé en Gate par un stagiaire du M2 de l'université Lille1. Au Centre Oscar Lambret, on a encore pour le moment deux Clinacs.
- En collaboration avec l'ICO de Nantes (dans le cadre d'un projet PhysiCancer de l'Inca) nous sommes en train de modéliser le Novalis, pour recalculer les plans de traitements des patients SBRT de Nantes.
- En collaboration avec le HUS (le CHU de Helsinki, Finlande) nous venons de commencer la modélisation des machines Elekta.

4.3 Modélisation du patient

Les patients sont modélisés en utilisant leur CT. Les niveaux de gris du CT (les unités Hounsfield) sont corrélés (bi-linéairement) avec la densité des tissus. Pour un calcul de dose des systèmes de planification de traitement utilisés en routine, il n'y a que la densité qui est modélisée. Pour un calcul Monte Carlo on peut aussi modéliser la composition des matériaux de chaque voxel. Cet étalonnage est souvent fait en utilisant des matériaux qui sont tissu équivalents. Une procédure encore plus précise (utilisée dans notre plateforme MC) est l'étalonnage stœchiométrique (Vanderstraeten et al [17], Demol et al [18]). Dans ce cas on va scanner des matériaux qui ne sont pas nécessairement tissu équivalents, mais pour lesquels on connaît la composition en détail. Après on étalonne les HU par

rapport à chaque élément individuel et on calcule les HU des vrais tissus en utilisant la composition et la densité énumérée dans des tables (par exemple l'ICRU). Cette méthode est détaillée plus loin.

4.4 Le plateforme Monte Carlo Moderato

La plateforme Moderato est basée sur le logiciel MCDE [7] que j'ai développé en 2004 à l'université de Gand afin de valider les plans de traitement RCMI de l'hôpital universitaire de Gand. Sur ce projet il y avait 7 thésards (5 à l'université de Gand, 1 en collaboration avec l'UCL de Bruxelles, et 1 à Erasme Rotterdam). MCDE était entièrement basé sur BEAMnrc/DOSXYZnrc. La fonctionnalité nécessaire pour utiliser ces deux logiciels pour un calcul complet d'un plan de traitement était complètement intégrée dans ces logiciels:

- Possibilité de calculer plusieurs faisceaux dans un seul calcul
- Interface Dicom pour importer les données du TPS (CT patient, RTStruct : les contours, RTPlan : la description des faisceaux, RTDose : la distribution de dose calculée par le TPS)
- Introduction d'une grille de dose indépendante des voxels géométriques (question d'efficacité) [19]
- Introduction du moteur de calcul dans un système d'optimisation de traitement en RCMI
- Introduction des méthodes de « denoising » (des filtres qui enlèvent le bruit sans modifier les vrais gradients de dose) [20]
- Modélisation des machines RCMI et VMAT d'Elekta
- Etalonnage stœchiométrique d'un nombre de scanners CT dans plusieurs centres en Europe [17]

Utilisant ce système nous avons recalculé des plans de traitements de plusieurs centres de radiothérapie, ce qui a mené à une vingtaine d'articles scientifiques.

Quand je suis arrivé au Centre Oscar Lambret à Lille (2008) il y avait déjà le Cyberknife et l'installation des deux premières machines Tomothérapie était prévue. Alors la question se posait tout de suite s'il y avait moyen d'utiliser MCDE pour modéliser ces deux machines et valider les plans de traitements. Par contre, il était clair que le logiciel n'était pas utilisable en clinique car il était mal programmé, pas facile à utiliser (pas d'interface graphique par exemple) et pas prévu pour l'ajout rapide d'autres machines. Dans le cadre d'un projet Siric (OncoLille) nous avons obtenu le budget pour employer un informaticien. Il a complètement reprogrammé MCDE (plus professionnellement) en c++/Javascript, et en utilisant les mêmes idées. Entretemps nous avons commencé la modélisation du Cyberknife et de la Tomothérapie comme décrit plus haut. Les demandes qu'on a posées pendant l'analyse initiale du logiciel étaient :

- Un système modulaire pour augmenter la flexibilité et faciliter l'ajout d'autre machines et même d'autres algorithmes de calcul Monte Carlo
- Un système de qualité : bien documenté et avec des tests pour chaque module
- Un système facile à utiliser avec une interface graphique (GUI)
- Un système avec une version routine utilisée pour le double calcul, et une version « recherche » capable de déterminer la dose effectivement délivrée, la possibilité de calculer sur des images IRM, et l'informatisation de la prescription médicale et l'automatisation de la validation des dossiers.

- Un système qui peut être utilisé en routine sans introduire trop de charge de travail pour les dosimétristes et les physiciens

Quelques schémas d'analyse sont montrés en figure 11.

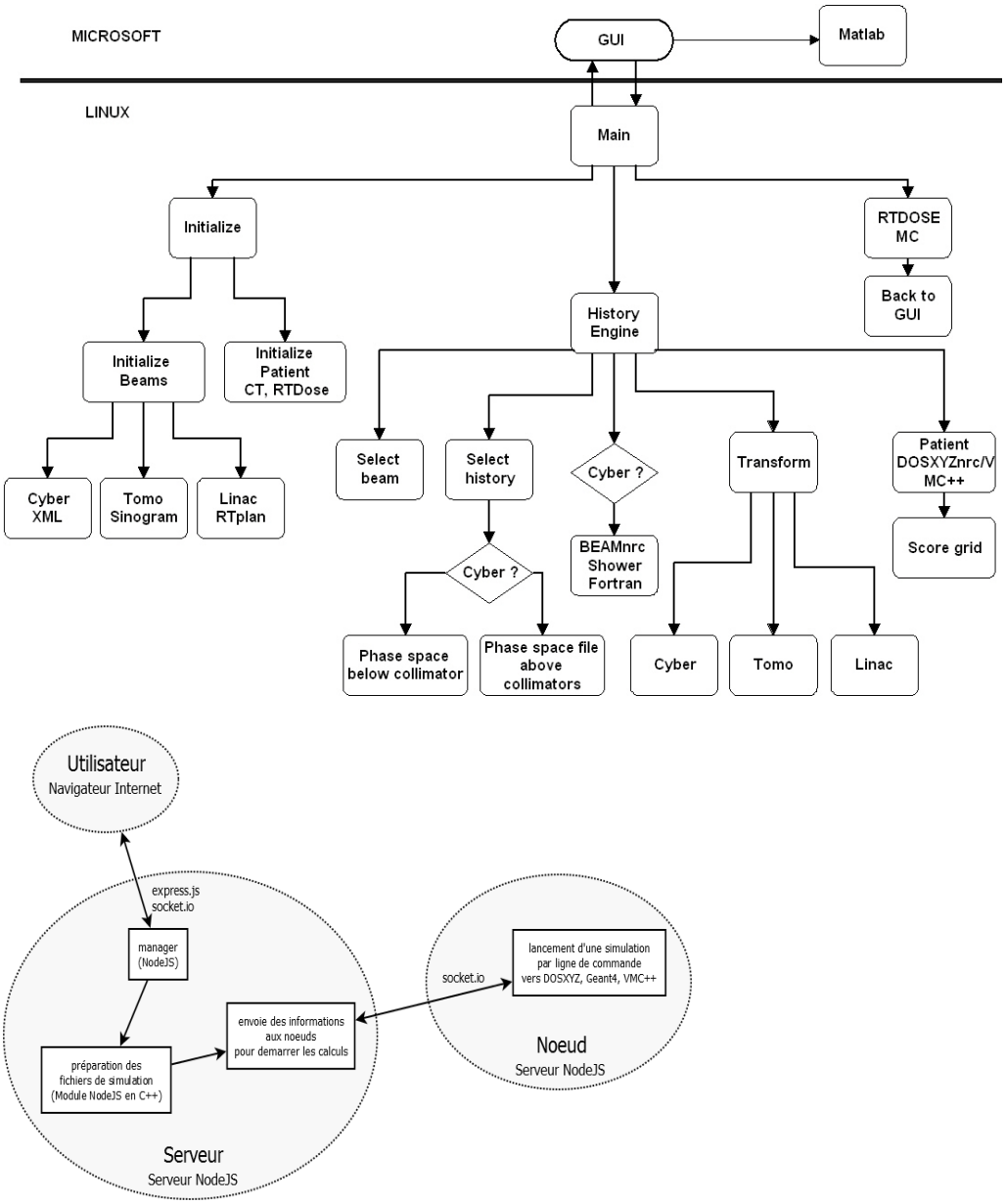


Figure 11 : L'analyse de la plateforme Monte Carlo Moderato. Le premier schéma montre la fonctionnalité des calculs, programmé avec des classes en c++. Le point clef est la centralisation du moteur (History Engine). Ceci offre la possibilité de combiner plusieurs algorithmes Monte Carlo (par exemple BEAMnrc et Geant4). Dans l'application actuelle le GUI n'est pas en Matlab mais en Javascript (page web). En bas, on voit le modèle serveur-nœuds programmé en NodeJS (Javascript), utilisé pour paralléliser les tâches sur les serveurs.

Actuellement, nous utilisons trois serveurs de 64 cœurs et 0.3 Tb de mémoire. Le logiciel a eu un Audit (Black Duck) qui était presque 100 % positif en collaboration avec EuraSanté. Nous avons eu une collaboration avec Intel UK, afin d'optimiser le système (logiciel plus serveur).

En clinique nous avons commencé par valider un grand nombre de cas Cyberknife. Après la modélisation détaillée de la Tomothérapie nous avons validé chaque plan de traitement Tomothérapie et Cyberknife en routine. Un serveur Dicom intégré dans Moderato va automatiquement détecter les exports de dossier des stations TPS et lancer les doubles calculs (aucune intervention du dosimétriste). Une fois le calcul fini (actuellement un double calcul prend 30 minutes), le physicien d'astreinte peut en quelques clics ouvrir, valider et importer le rapport en pdf dans le système Record&Verify. Un exemple de l'interface graphique est montré en figure 12.

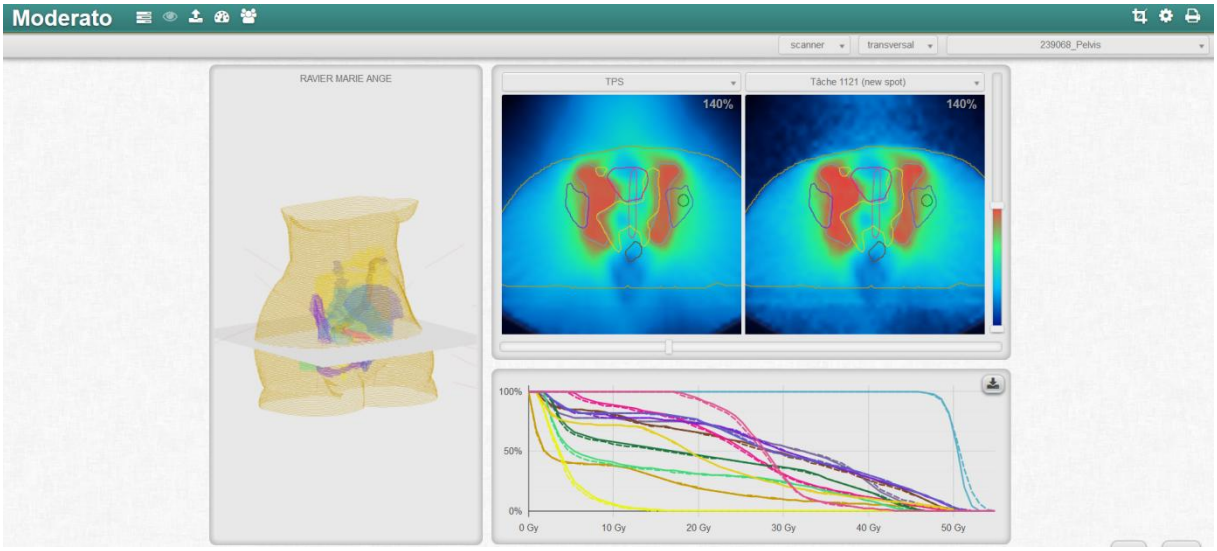


Figure 12 : GUI (page Web) pour visualiser la distribution de dose et les histogrammes dose-volume avec Moderato.

Plus de détails concernant l'utilisation du logiciel sont donnés au chapitre 2.

5. L'IRM en radiothérapie

Depuis son introduction à la fin des années 1970 [21], l'imagerie par résonance magnétique (IRM) est une technique d'imagerie avancée qui est rapidement devenue incontournable en radiothérapie [22]. Le principe de l'IRM repose sur la polarisation des noyaux d'hydrogène par un champ magnétique constant, et sur leur excitation à des niveaux d'énergie supérieurs par des pulses électromagnétiques radiofréquences. Lorsque ces noyaux retournent à leur état stable, ils émettent des ondes radiofréquences qui sont détectées par des antennes composées de bobines. Les différences dans le temps de retour à l'état stable, notamment par les temps de relaxation T_1 et T_2 ayant pour origine respectivement les interactions spins-réseaux et spins-spins, caractérisent le type de tissus. Le processus est défini par les équations de Bloch :

$$\frac{dM_x(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_x - \frac{M_x(t)}{T_2}$$
$$\frac{dM_y(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_y - \frac{M_y(t)}{T_2}$$
$$\frac{dM_z(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_z - \frac{M_z(t) - M_0}{T_1}$$

Dans ces équations le vecteur M(t)(M_x(t), M_y(t), M_z(t)) est la magnétisation nucléaire ; ϒ est le rapport gyromagnétique ; B(t)(B_x(t), B_y(t), B_z(t)) est le champ magnétique. B_z(t) est la sommation de B_0 (le champ magnétique statique : couramment 1.5 ou 3T) et le gradient ΔB_z(t) utilisé pour encoder la position en z ; M_0 est la magnétisation « steady state » : pour t → ∞ M_0 est orienté en z. Alors le temps de relaxation T_1 montre la vitesse de relaxation en direction z déterminé par l'interaction entre le spin du noyau et le réseau. T_2 par contre représente la relaxation dans le plan transverse (XY) et spécifique des interactions spins-spins. Sans ces relaxations (T_1 et T_2 → ∞) ses équations deviennent les équations de Larmor. Il faut bien insister que les équations de Bloch sont des équations macroscopiques qui décrivent l'état de l'ensemble des spins dans un volume.

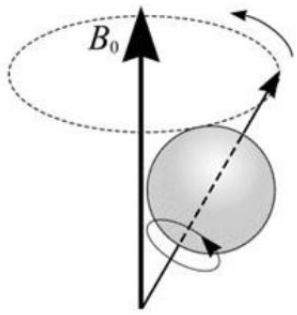


Figure 13 : Précession d'un noyau autour d'un champ magnétique B_0.

Comme montré en figure 14 la fréquence de précession d'un noyau autour d'un champ magnétique statique B_0 est donné par f = ϒ x B_0. Dans cette situation les spins on deux états : up ou down (parallèle ou antiparallèle avec le champ B_0). La magnétisation nette est la différence entre les ups et les downs (figure 14). Le signal IRM résultant dépend du nombre de ups, et par conséquent de la densité des protons dans les tissus. Alors, en théorie, en l'absence d'effets de relaxation (T_1 et T_2 → ∞), un IRM pourrait servir à déterminer la densité de protons dans les voxels du patient. Cependant, on va le voir au chapitre 3, il est extrêmement compliqué de « désactiver » ces relaxations.

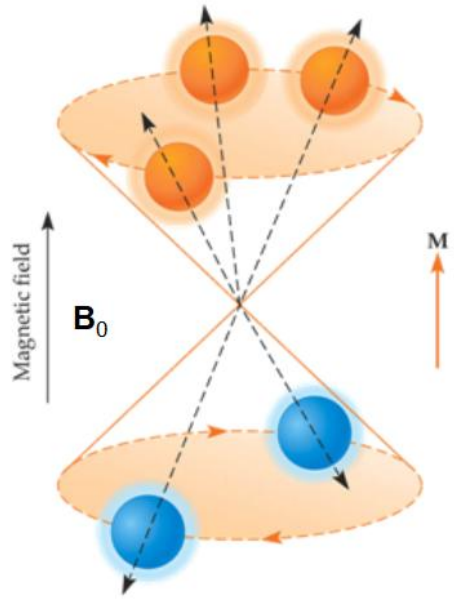


Figure 14 : Magnétisation nette dans la direction de B_0.

Si on lance un pulse radio fréquence avec la même fréquence que la fréquence de Larmor d'un noyau, le spin va tourner (produit vecteur) dans le plan perpendiculaire (figure 15). La durée du pulse détermine l'angle du tilt (souvent 90 ou 180 degrés).

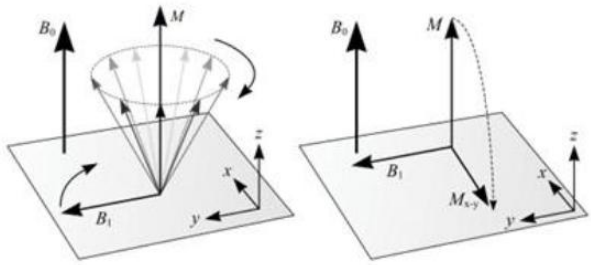


Figure 15 : Rotation de la Magnétisation avec un pulse radio fréquente.

Après un pulse de 90 degrés (l'exemple montré en figure 16) le vecteur M est orienté et fait des rotations autour de l'axe longitudinal dans le plan transverse. Un signal avec la même fréquence est émis par le noyau et détecté dans la bobine de réception de signal. L'amplitude de ce signal est maximale lorsque le vecteur M est dans le plan transverse. L'interaction avec B_0 par contre, va de nouveau l'orienter en z et le signal dans la bobine (l'enveloppe du signal) va diminuer et devenir zéro une fois que la magnétisation sera de nouveau orientée en z.

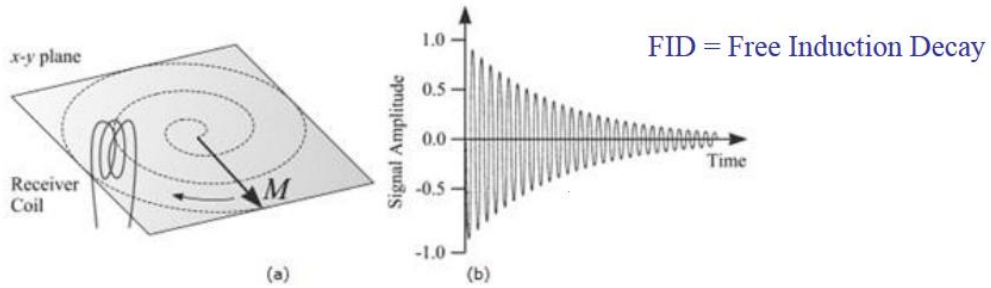


Figure 16 : La magnétisation mesurée dans la bobine de réception.

Les temps de relaxation T_1 et T_2 sont définis à la figure 17. Le T_1 est défini comme le temps qu'il faut avant que M récupère 63 % de sa valeur initiale. T_2 est le temps qu'il faut au signal dans le plan transverse pour diminuer de 37 % par rapport à sa valeur juste après le pulse à cause du déphasage des spins. Par définition les spins vont se déphaser à cause des hétérogénéités dans le champ magnétique local, alors pour maximiser le signal on applique un deuxième pulse de 180 degrés qui va compenser ce déphasage et générer un signal « écho » après un temps d'écho (TE). Le signal S mesuré dans la bobine de réception après l'écho est donné par l'équation :

$$S(TR,TE) \propto \rho \{1 - e^{-TR/T_1}\}\{e^{-TE/T_2}\}$$

T1W T2W

Le T_1 et le T_2 peuvent être utilisés pour obtenir différents contrastes dans les images IRM, surtout avec l'application des produits de contraste comme le gadolinium (un produit qui va raccourcir le T_1 des tissus).

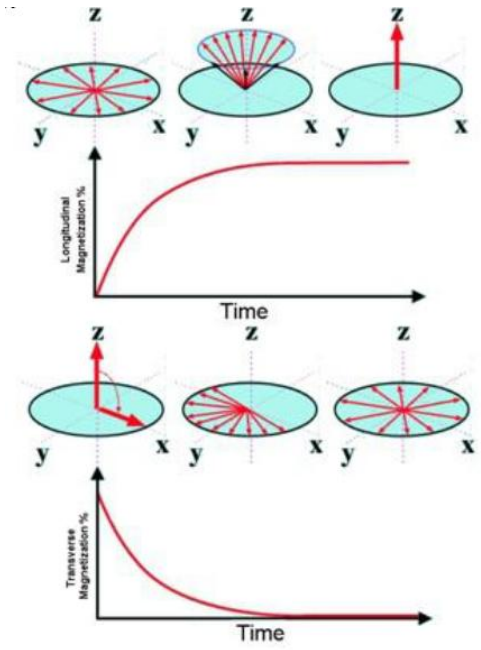


Figure 17 : T_1 est défini par le temps qu'il faut avant que la magnétisation reprenne le sens de B_0 près un pulse. T_2 est déterminé par le déphasage des spins dans le plan transverse.

L'IRM est construite sur le principe de séquence, un enchaînement déterminé de pulses radiofréquences, de gradients de champ magnétique, et de fenêtres d'acquisition du signal, répété périodiquement, comme le montre la Figure 18 dans le cas d'une séquence écho de spin. Chaque séquence représente un arrangement particulier de ces composants qui permet de coder le signal selon une fonction spatiale dans le plan de Fourier, et est définie entre autres par le temps d'écho et le temps de répétition (le temps entre deux enchaînements périodiques). Un post-traitement approprié de l'espace de Fourier fournit une image anatomique dont l'intensité des voxels dépend de la densité de protons, des temps de relaxation T_1 et T_2, et des paramètres de la séquence. Cette intensité est exprimée en niveaux de gris et n'a pas d'unité. Selon la séquence et les paramètres utilisés, l'intensité du signal peut dépendre en proportion majeure du T_1, du T_2, ou de la densité de protons des tissus, et le signal est alors pondéré par une de ces grandeurs. On peut pondérer les images par les valeurs T_1 des tissus, utilisant un temps d'écho court et un temps de répétition court (et l'inverse pour pondérer l'image avec T_2). Par contre, si on veut déterminer la densité des protons (important dans le cadre de la radiothérapie en IRM seule, voir chapitre 3) il faut utiliser un temps d'écho « ultra » court (pour enlever la pondération T_2) et un temps de répétition très long (pour enlever la pondération T_1). Les images pondérées en T_1 sont utilisées en particulier pour délinéer le contour des différents tissus et peuvent être améliorées en utilisant un produit de contraste comme le gadolinium. Les images pondérées en T_2 sont des indicateurs sensibles à la détection des cancers et permettent de caractériser la plupart de ses phases. Elles permettent également de distinguer les tissus bénins des tissus malins [23].

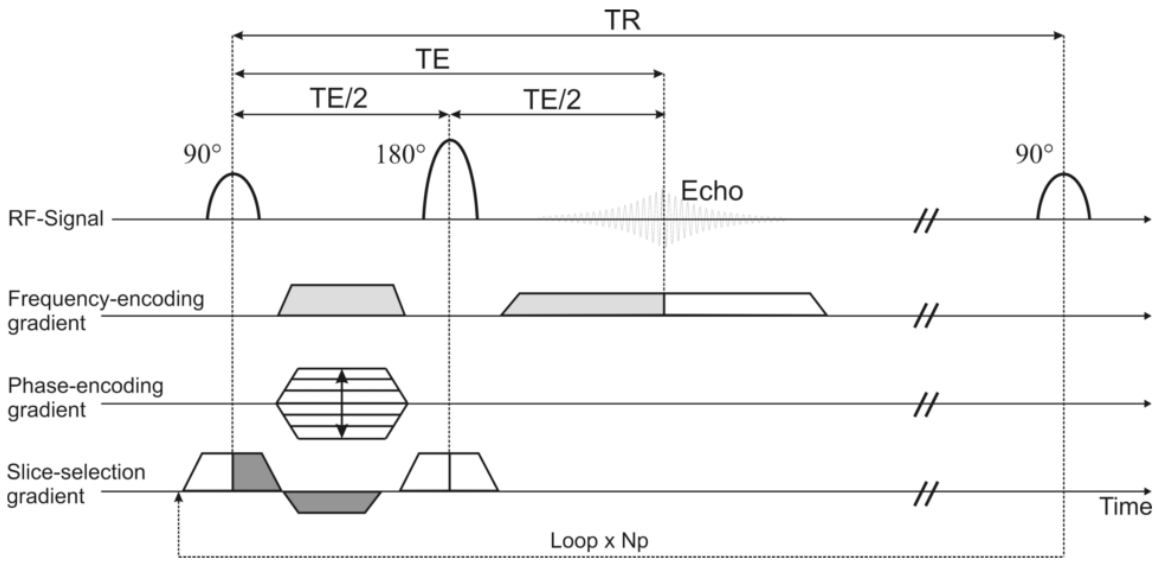


Figure 18 : Diagramme de séquence d'une séquence écho de spin classique. Un premier pulse RF bascule l'aimantation à 90° dans le plan sélectionné par le gradient de sélection de coupe. Les gradients d'encodages de la fréquence et de la phase permettent de sélectionner l'espace du plan de Fourier à remplir lors de la lecture du signal (de l'écho). La séquence écho de spin introduit un pulse RF de 180° afin de récupérer la dispersion de la phase ayant pour origine les inhomogénéités de champ. Cet enchaînement est répété N_p fois pour remplir l'espace de Fourier d'une coupe. Image reproduite de Jung et Weigel [24].

De ce fait, l'IRM permet de mieux discerner les tumeurs que le CT, et a donc été intégrée dans les pratiques de la radiothérapie en tant qu'image diagnostique. Le volume tumoral est délimité plus précisément sur une image IRM que sur une image CT, que ce soit dans le cas des tumeurs de la prostate [25][26], des lésions cérébrales [27], ou des tumeurs tête et cou [28][29]. Ces études

montrent que les variations intra et inter-observateurs sont diminuées dans le cas de l'utilisation de l'IRM [30].

Chapitre 2 : Validation de la dose planifiée et délivrée

1. Introduction

1.1 Le Double calcul

Le logiciel IMSURE est utilisé au Centre Oscar Lambret à Lille pour le double calcul des dosimétries conformationnelles (Clinac). Par contre, ce logiciel n'est pas utilisable pour la Tomothérapie et est trop approximatif pour le Cyberknife. Et c'est effectivement pour ces traitements plus complexes, avec des gradients de dose plus importants, souvent hypofractionnés et délivrant des doses plus importantes dans les organes à risques (traitements curatifs) qu'il faudrait faire un double calcul, et ça en utilisant un logiciel qui est par définition plus précis que le TPS même. Historiquement on était déjà contents de trouver les grosses erreurs. Maintenant, on aimerait obtenir une précision plus élevée (2-3 %). Dans le TPS du Cyberknife (Multiplan^{TM}), il y a deux algorithmes de calcul de dose : un algorithme approximatif appelé le Ray-tracing et un algorithme Monte Carlo. Le Ray-tracing va directement appliquer les tables de mesures utilisant des corrections (1D « scaling » avec la densité). Cet algorithme donne des erreurs majeures par rapport au Monte Carlo dans les poumons principalement (voir figure 19). En routine clinique, l'algorithme ray-tracing est toujours utilisé dans les phases d'optimisation et de prescription, sauf pour les poumons où l'algorithme Monte Carlo est utilisé pour la prescription. L'algorithme Monte Carlo du TPS est principalement utilisé comme double calcul. Evidemment, on peut se poser des questions concernant l'indépendance entre les deux calculs. Mais pendant longtemps, il n'y avait que ça.

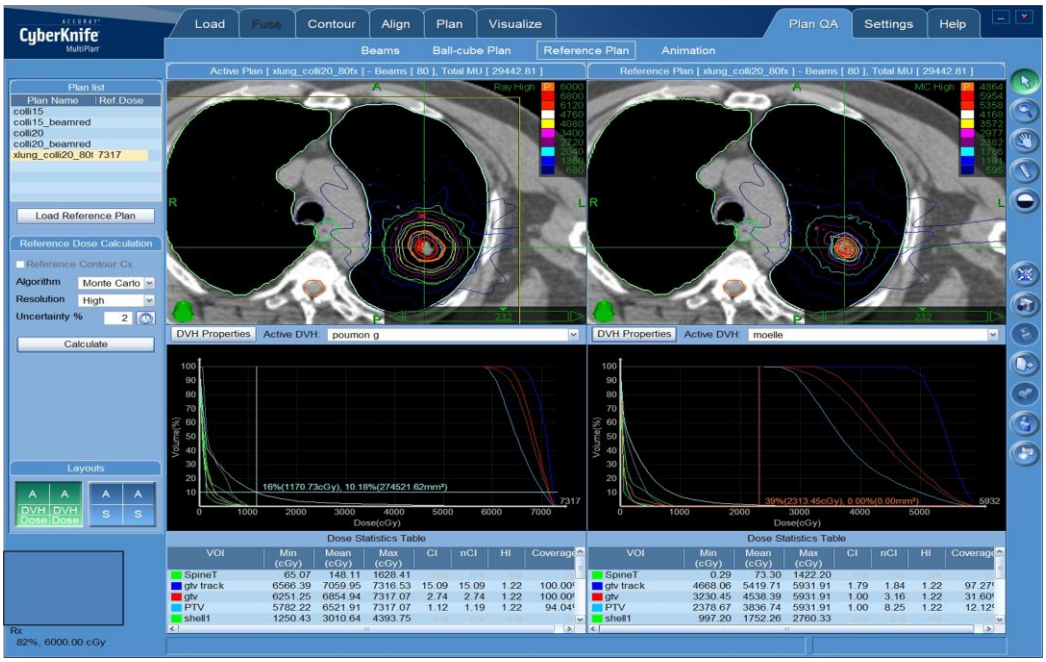


Figure 19 : L'interface du TPS Multiplan. Comparaison d'un calcul entre les algorithmes ray-tracing (gauche) et Monte Carlo (droite). L'optimisation est faite en ray-tracing et la dose est normalisée à 60 Gy dans 95% du PTV.

Les grosses erreurs qu'on observe dans les poumons ne sont pas spécifiques à l'algorithme ray-tracing de Multiplan, on le retrouve pour chaque algorithme qui ne prend pas en compte le transport des électrons secondaires. Pour cette raison, il y a maintenant un consensus stipulant qu'il faut utiliser un algorithme de type B prenant en compte les électrons secondaires (comme le Monte Carlo, superposition/convolution, collapsed cone, Acuros^{TM} de Varian, …) [4]. Dans notre cas, on utilise l'algorithme MC de Multiplan pour faire les prescriptions dans les poumons. Puis la question se pose de savoir comment faire le double calcul dans ce cas. L'inversion des rôles des deux algorithmes (utilisant le ray-tracing comme double calcul) ne semble pas raisonnable comme on peut le voir à la figure 19. Il y a clairement le besoin d'un nouveau système de double calcul indépendant.

Pour la Tomothérapie, il n'y avait pas de double calcul possible dès l'installation des machines. Pour chaque patient, un contrôle prétraitement est fait (DQA). Un DQA valide surtout la délivrance du plan de traitement par la machine, tel que le TPS l'avait calculé. Les mesures sont faites dans un fantôme homogène ; une géométrie qui de correspond pas du tout avec celle du patient. Cette mesure est un complément du double calcul mais ne peut pas le remplacer. C'est pourquoi la demande d'introduire un logiciel de double calcul pour la Tomothérapie était aussi importante.

C'est pour toutes ces raisons que nous avons décidé de modéliser les Linacs du Cyberknife et de la Tomothérapie en Monte Carlo et de reprogrammer MCDE [7] afin d'introduire un système de haute qualité, facile à utiliser, flexible vers le futur au Centre Oscar Lambret : Moderato (comme décrit plus en détail dans le premier chapitre). Dans Moderato, nous avons le modèle du Cyberknife installé au COL [31], mais aussi celui du nouveau Cyberknife (le M6^{TM}) avec MLC. Pour la Tomothérapie, il y a toutes les options récentes comme les mâchoires dynamiques. Les Clinacs (Varian) sont modélisés aussi, mais Moderato n'est actuellement pas utilisé pour le double calcul de ces machines. En collaboration avec l'ICO de Nantes (dans le cadre d'un projet PhysiCancer de l'INCA) on est en train de modéliser le Novalis^{TM} (Brainlab) et nous venons de commencer la modélisation des machines Elekta pour le HUS de Helsinki (Finlande). Le but est de généraliser le logiciel et de lancer une start-up au COL afin de commercialiser Moderato (en collaboration avec la société Aquilab). Cette start-up devrait activer un service de Double calcul à distance accessible aux autres centres de radiothérapie en France et ailleurs. Le fonctionnement de ce service est décrit plus en détail à la fin de ce chapitre. Les méthodes utilisées dans Moderato sont expliquées dans le premier article introduit dans ce chapitre.

1.2 Vers la dose délivrée

Un double calcul a pour but de valider la dose provisoire calculée par le TPS. Par contre, à cause des incertitudes de positionnement journalier du patient, les mouvements internes et externes du patient, les changements systématiques de l'anatomie du patient en cours de traitement (amaigrissement par exemple), les erreurs au niveau de la machine, … la dose effectivement délivrée au patient peut largement différer de la dose planifiée. La méthode d'estimation de la dose délivrée dépend largement de l'information disponible pendant le traitement. Pour les traitements sur des machines conventionnelles, l'imageur portal est de plus en plus utilisé pour reconstruire la dose « in-vivo ». Pour la Tomothérapie et Cyberknife il faut par contre d'autres méthodes.

1.2.1 La dose délivrée par séance au Cyberknife

Le Cyberknife utilise un système d'imagerie (deux tubes à rayons X montés au plafond, et deux détecteurs intégrés dans le sol) pour suivre les mouvements du patient avec plusieurs systèmes d'asservissement (« tracking »). Pour les géométries rigides (le crâne par exemple) on peut considérer que la dose planifiée sera une bonne estimation de la dose délivrée. Par contre, dans le thorax, les mouvements sont corrélés avec la respiration du patient (poumons, foie, reins, …). Pour ces traitements, il y a le système Synchrony^{TM} (décrit au Chapitre 1). Le robot va « respirer » avec le patient et, par définition, les positions utilisées pendant le traitement ne seront pas celles planifiées dans le plan de traitement. Ces positions sont sauvegardées dans des fichiers logs avec une résolution de 20 ms. Les positions des fiduciels internes et des marqueurs externes sont sauvegardées aussi. Le seul inconvénient est qu'il n'y a pour l'instant pas d'imagerie 4D (pas même 3D) disponible pendant le traitement. On est alors obligé de partir des images 4D CT de simulation (les images utilisées pendant la planification du traitement).

Afin de reconstruire la dose délivrée [32][33][34][35], on peut partir du fichier log qui contient les positions du robot. Si on sélectionne une position dans ce fichier, le temps est utilisé pour chercher la position des marqueurs correspondante. Cette position est utilisée pour estimer la phase respiratoire, et ainsi le CT correspondant pour lancer le calcul Monte Carlo du faisceau. La distribution de dose obtenue est alors projetée sur le scan 3D (le scan de planification) en utilisant des vecteurs de déformation entre les deux scans (calculés avec ITK/Elastix [45][46]). On peut ainsi cumuler la dose de tous les faisceaux présents dans le fichier log. La résolution temporelle peut être augmentée par rapport à la résolution du scan 4D (5 ou 10 phases par cycle respiratoire) en interpolant la géométrie entre les phases (en utilisant de nouveau des vecteurs de déformation). En pratique par contre, cette méthode ne serait pas du tout efficace en temps de calcul. Alors on part plutôt de la géométrie. La résolution dans le temps va déterminer combien de géométries on va considérer (par interpolation entre les phases si besoin). Et puis, pour chaque géométrie on va utiliser le temps dans les fichiers logs pour déterminer les faisceaux qui étaient « actifs » (on prépare un RTPlan qui est spécifique pour chaque phase respiratoire considérée). Puis, le nombre de calculs Monte Carlo à effectuer est déterminé par le nombre de géométries considérées. Cette méthode est appliquée « manuellement » dans le cadre de la thèse de Marie Charoy, et décrit dans le deuxième article introduit dans ce chapitre (thèse finalisée en juin 2014). La méthode est maintenant automatisée dans Moderato. Dès le moment qu'il y a des fichiers 4D CT et des fichiers logs détectés pendant l'import du dossier, Moderato va automatiquement déterminer la dose délivrée et la comparer avec la dose planifiée.

1.2.2 La dose délivrée par séance en Tomothérapie

Il n'y pas de fichiers logs disponibles pendant les traitements de Tomothérapie. Le MVCT est utilisé pour reconstruire le sinogramme effectivement délivré. Ceci ne concerne que la position des lames. Les autres paramètres comme l'angle du bras ou la position de la table ne sont pas sauvegardés. Par contre, il y a dans le sinogramme l'information concernant la synchronisation du MLC avec la table et le bras.

Pour reconstruire la géométrie du patient nous pouvons utiliser les images MVCT du jour, prises juste avant le traitement. D'un coté, cela donne plus d'information que sur le Cyberknife, où il n'y a pas

d'information 3D disponible. De l'autre, il n'y pas d'information du tout pendant l'irradiation, et les mouvements du patient pendant le traitement ne sont pas connus. Le calcul de la dose délivrée est plus simple, car il faut juste appliquer le sinogramme reconstruit sur le MVCT du jour. La seule difficulté (plutôt informatique) est qu'il faut compléter le scan MVCT avec des coupes du kVCT, parce qu'on ne scanne jamais le volume en entier.

1.2.3 L'accumulation de la dose de toutes les séances

Pour toutes les modalités de traitement, une fois qu'on a déterminé la dose délivrée pour chaque séance, il reste encore à accumuler la dose totale du traitement. Pour le Cyberknife, par manque d'information 3D, ça revient à sommer les doses reconstruites pour chaque séance sur le CT de simulation. Pour la Tomothérapie, où il y a un scan 3D de chaque séance, il faut déformer le MVCT sur le kVCT et propager la dose sur le kVCT avant de faire la sommation. En théorie, il n'y a pas vraiment de difficultés. En réalité, par contre, il y a le problème que pour certaines déformations du patient, la masse des tissus n'est pas toujours conservée. Lors de périodes d'amaigrissement par exemple, des voxels vont disparaître. Pendant les traitements ORL, les parotides changent en volume, en masse et en densité, et il devient quasi-impossible d'accumuler correctement la dose dans les voxels. C'est aussi vrai dans le cas où la cible diminue de masse (et de volume) d'une séance à l'autre ou par rapport au scan de planification (le kVCT). Par définition, une déformation est basée sur une conservation de la masse.

Dans ce cas la quantité « dose » n'a plus de sens et la question devient très fondamentale : est-ce qu'il ne faut pas définir une autre quantité, plus liée à la biologie des tissus ? Pour les parotides pendant les traitements ORL, on pourrait envisager de suivre le pourcentage de volume qui reçoit une dose maximale, ce qui est considéré comme biologiquement important. On pourrait aussi suivre la dose moyenne de ces structures, car elles sont considérées comme des organes parallèles. Mais on sent tout de suite qu'on ne peut pas généraliser et qu'il faut presque chercher une solution spécifique pour chaque organe.

1.2.4 Les articles

Le premier article donne la description de Moderato comme système de double calcul de la dose planifiée et délivrée. L'accent est mis dans cet article sur la façon dont le système est introduit et utilisé en routine, sans ajouter plus de travail à ceux qui préparent les plans de traitements.

Le deuxième article offre le résumé de la thèse de Marie Charoy : l'évaluation de la précision du système d'asservissement « Synchrony tracking » pour les traitements du foie. Le but initial de sa thèse était de déterminer l'impact des rotations et des déformations pendant ces traitements. Le système Synchrony ne considère que les translations du barycentre des fiduciels implantés dans le foie. Les rotations et déformations ne sont pas prises en compte.

2. Article 1

Clinical implementation of a Monte Carlo based treatment plan QA platform for validation of Cyberknife and Tomotherapy treatments

N. Reynaert[1], B. Demol[1], M. Charoy[1], S. Bouchoucha[1], F. Crop[1], A. Wagner[1], T. Lacornerie[1], F. Dubus[1], E. Rault[1], P. Comte[1], R. Cayez[1], C. Boydev[1], D. Pasquier[2], X. Mirabel[2], E. Lartigau[2], T. Sarrazin[1]

1: Department of Medical physics, Centre Oscar Lambret and University Lille 1

2: Academic Department of Radiation Oncology, Centre Oscar Lambret, and University Lille 2

Abstract

Purpose: The main focus of current paper is the clinical implementation of a Monte Carlo based platform for treatment plan validation for Tomotherapy and Cyberknife, without adding additional tasks to the dosimetry department.

Methods: The Monte Carlo platform consists of C++ classes for the actual functionality and a web based GUI that allows accessing the system using a web browser. Calculations are based on BEAMnrc/DOSXYZnrc and/or GATE and are performed automatically after exporting the dicom data from the treatment planning system. For Cyberknife treatments of moving targets, the log files saved during the treatment (position of robot, internal fiducials and external markers) can be used in combination with the 4D planning CT to reconstruct the actually delivered dose. The Monte Carlo platform is also used for calculation on MRI images, using pseudo-CT conversion.

Results: For Tomotherapy treatments we obtain an excellent agreement (within 2 %) for almost all cases. However, we have been able to detect a problem regarding the CT Hounsfield units definition of the Toshiba Large Bore CT when using a large reconstruction diameter. For Cyberknife treatments we obtain an excellent agreement with the Monte Carlo algorithm of the treatment planning system. For some extreme cases, when treating small lung lesions in low density lung tissue, small differences are obtained due to the different cut-off energy of the secondary electrons.

Conclusions: A Monte Carlo based treatment plan validation tool has successfully been implemented in clinical routine and is used to systematically validate all Cyberknife and Tomotherapy plans.

Key words: Monte Carlo, QA, treatment planning, delivered dose

Introduction

Currently, radiotherapy treatment plans are often validated using independent monitor unit calculation algorithms [1-4]. In France e.g., there is a legal obligation to apply this QA technique for all treatments when such a calculation is “technically feasible”. This leads to a systematic validation of conformal treatments, using a limited number of beams, while the more complicated techniques, such as IMRT, VMAT and SBRT, are not

often validated systematically. This is because few commercial tools are available for these techniques, or because the precision of these systems is limited. Furthermore, most monitor unit validation tools only recalculate dose at the isocenter, while for the advanced techniques a more detailed 3D dose distribution is needed, not only focusing on the PTV, but also paying attention to the organs at risk (OARs). Although there are a couple of commercial tools available, providing 3D plan validations, these do not provide a solution for Tomotherapy and Cyberknife.

During the last decades, Monte Carlo codes have become easily available and have been developed for quality control of radiotherapy treatment plans [5-8]. Furthermore, in the last decade, an increasing number of treatment planning systems are using Monte Carlo based algorithms [9,10]. As these algorithms are by nature relatively slow, a number of approximations are applied, sometimes leading to hybrid systems. Especially the beam model is often largely approximated [10]. This is unfortunate as the ability of modeling the beam in detail is the main advantage of a Monte Carlo algorithm, when comparing e.g. to collapsed cone algorithms that already provide an adequate precision for dose calculation within the patient geometry. Because of these approximations it remains important to validate treatment planning systems, with e.g. independent Monte Carlo algorithms using an accurate beam model [11]. The main challenge is not the implementation of such an algorithm but rather the introduction of such a tool in clinical routine. The Monte Carlo platform used in our department is based on our in-house Monte Carlo based dose calculation software MCDE [12]. This system was used at the University of Ghent to validate IMRT plans in different radiotherapy departments [13-15]. In MCDE the program DOSXYZnrc [16] was reprogrammed as a component module for BEAMnrc [17] and a dicom interface was added to translate the RTPlan file to a BEAMnrc input file (all beams were handled in one single input file). This demanded a reprogramming of certain parts of BEAMnrc/DOSXYZnrc. The RTStruct and RTDose files exported from the TPS were also imported in MCDE to allow the application of a statistics-based stop taking into account the statistical noise in all organs of interest [18]. A denoising algorithm was developed [19] and an independent scoring grid was introduced [20] to increase efficiency. The system was used to validate a large number of H&N treatments and was compared to Pinnacle, Helax, AAA and Peregrine [11,13-15]. The main problem of MCDE was that all QA specific modifications were applied inside BEAMnrc/DOSXYZnrc. This lack of flexibility led to the decision to reprogram the program completely, focusing on user-friendliness, flexibility and quality. This recently lead to a new system (Moderato) that allows an easy handling of new treatment units and even switching between different Monte Carlo engines. But the main focus of current paper is to describe how such a system was introduced in clinical routine in a smooth way that didn't add additional tasks to the dosimetrists.

Methods and Materials

Description of the Monte Carlo platform

As current Monte Carlo platform is based on MCDE it was originally constructed around BEAMnrc/DOSXYZnrc while focusing on flexibility. The main flow of a Monte Carlo plan validation is illustrated in figure 1.

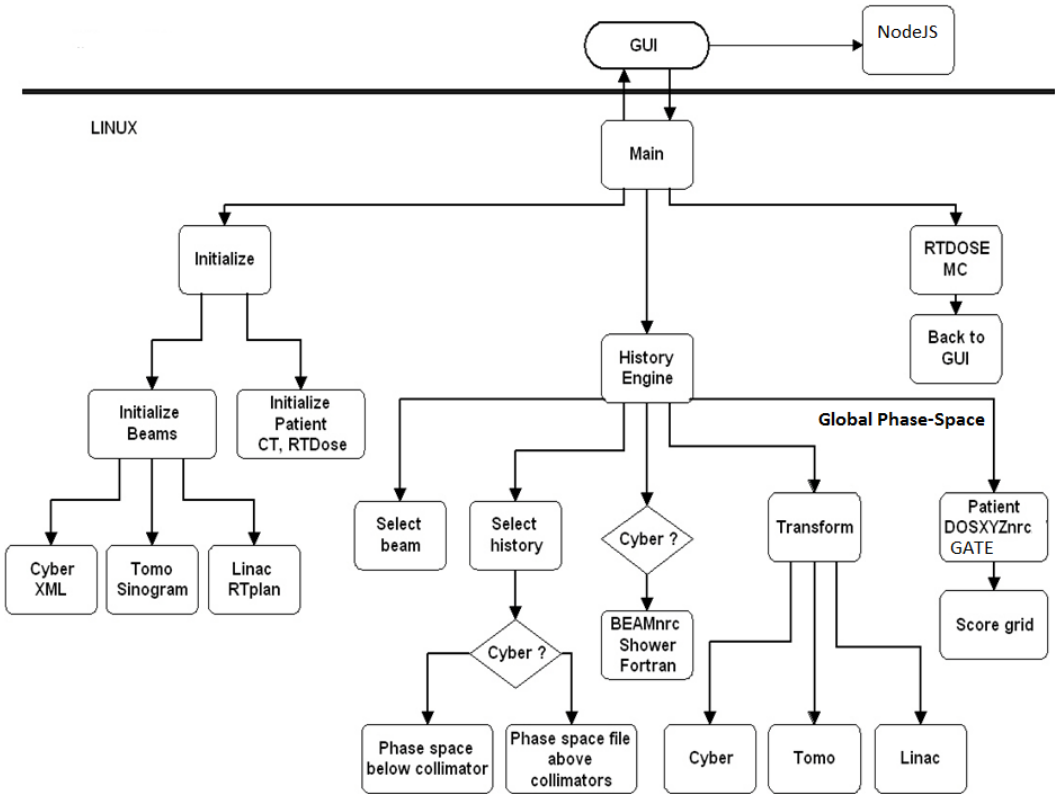


Figure 1. Flow chart of Moderato. The GUI is accessible via a web browser and allows uploading patient data, launching calculations, visualizing the isodoses and the DVHs, and printing a report. The main functionality is programmed in C++. The history engine is centralized in Moderato allowing easy switching between different Monte Carlo codes. For Cyberknife phase-space files can be pre-calculated for the two sets (fixed and Iris) of 12 collimators, consequently the BEAMnrc calculation step is replaced by simply reading these files. The initialization of Tomotherapy treatments is possible using the binary sinogram but also via the RTPlan file. The transformation depends on the degrees of freedom (rotation angles, table movement) of each modality. The patient calculation can be performed using DOSXYZnrc or GATE.

The process can be summarized as follows:

- Initializing geometry and BEAMnrc input files using the dicom/xml files exported from the TPS
- Generating a global phase-space file at the exit plane of the treatment head that combines particles from all beams (can be considered as a cylindrical phase-space surrounding the patient)
- Performing the dose calculation in the patient geometry for each sub-task (MC calculation is ran in parallel on several CPU cores)
- Combining and converting the obtained dose files for all sub-tasks and make the results available in the GUI (the web browser)

The first and the last tasks are handled by a master process (that is also providing the Javascript GUI). The middle part is handled by a number of slave processes and parallelized over multiple CPU cores. Currently 3 servers, each having 64 CPU cores are combined to parallelize the Monte Carlo calculations. Each slave sub-process runs on a single core and generates phase-space by launching for example a BEAMnrc calculation; transforms this phase-space in dicom coordinates using the beam parameter information, and performs the patient calculation. In such a way each process simulates the complete treatment and the obtained dose files are merged by the master process. As Moderato is simply launching the Monte Carlo calculations by using default commands, one can easily switch between different Monte Carlo engines, by generating the specific input files. No modifications were applied to the Monte Carlo engines. Currently BEAMnrc/DOSXYZnrc and GATE [21] are implemented in Moderato. This allows for example to generate the phase-space using BEAMnrc, while calculating the patient dose distribution, using Gate. The IAEA phase-space format is used to ensure

compatibility between the different codes. Ctcreate (BEAMnrc) is used to convert the CT Hounsfield units to a voxelized geometry (density and material composition). The image value to density table (IVDT) is obtained using the stoichiometric calibration method as described in [22, 23]. But other, e.g. TPS and CT specific, IVDT curves are available as well and are selected in the GUI when launching the calculation.

Modeling of the different treatment modalities

1. Cyberknife

The Cyberknife$^{(R)}$ (Accuray, Sunnyvale, CA, USA) was modeled by Wagner et al. [24]. Both the Iris and fixed collimators were modeled and phase-space files for the two sets of 12 collimators were pre-calculated and stored for utilization in Moderato, i.e. for these calculations the phase-space generation step is replaced by simply reading these phase-space files while saving the transformed particles to the global patient specific phase-space file as described above. Calibration was performed by linking the Monte Carlo calculation for the 6.0 cm field at SAD= 80 cm (1.5 cm depth) with the actual calibration in the water phantom. During the treatment plan calculation, absolute dose is obtained by multiplying the Monte Carlo dose (Gy/primary history) with the obtained calibration factor and the total number of monitor units in the plan. The Cyberknife M6^{TM} (including the Incise^{TM} MLC) was modeled as well. The Cyberknife M6 has a modified (pyramid shape) primary collimator allowing the maximal fieldsize (11.5 cm x 10 cm) when using the MLC. To model the MLC the diagrams provided by Accuray were used. The MLC data are also available in the plan xml file.

2. Tomotherapy

The Tomotherapy unit (Accuray, Sunnyvale, CA, USA) was modeled using the drawings provided by Accuracy. A double Gaussian spot as described in [25] was needed to obtain agreement between measurements and MC calculations for all jaw openings (1.0, 2.5 and 5.0 cm). Recent treatment options such as the dynamic jaws and TomoDirect were included. The transport through the MLC can be performed using full Monte Carlo transport or by a more efficient ray-tracing technique. The ray tracing considers the detailed geometry (i.e. the tongue and groove, and the leaf bank tilt) of the leaves. The total distance travelled through tungsten is used to calculate an exponential attenuation factor used in a russian roulette process. This method, which is in principle less accurate as scatter in the MLC is completely neglected, is much faster than the full Monte Carlo option and is the default option when recalculating a patient plan in clinical routine. The modeling of the MLC (geometry and position of the leaves) was based on measurements in the water phantom and treatment planning system calculations of individual beams traversing heterogeneous phantoms. The position of the leaf sides was read from the machine archive file for the individual machines (in our center, 3 Tomotherapy units are used in clinical routine). Afterwards, the commissioning of the global model of the linac head in Moderato was performed on 25 patients of different clinical indications to include treatments with different field sizes, modulation factors and pitch. For this commissioning the agreement with the TPS was considered as the metric. As the Tomotherapy TPS is using an accurate dose calculation algorithm (collapsed cone) and as the accuracy of the TPS has been validated by a large number of DQA measurements, we consider that on average, the TPS will provide accurate results. So we determined a distribution of differences between TPS and Moderato and verified if, on average both systems are in agreement. At the same time outliers could be investigated in detail.

For absolute calibration, the same method as described for the Cyberknife was applied, while replacing the number of monitor units by the treatment time.

Calculation Statistics and voxel sizes

Calculation time can be determined by simply defining the number of primary histories to be simulated, or by applying a statistic stop as described in [18]. When using this last option, the user defines the required uncertainty level and the normalization dose for each region of interest. For example, one can ask for 2 % of the prescription dose in the PTV and for 3 % of 40 Gy in the spinal cord. In a first iteration the Monte Carlo code will determine the uncertainty for 1e5 primary histories and then estimates the number of histories to be added to obtain the required uncertainty level and launches a second iteration. The code runs until all uncertainty levels are met in 95 % of the voxels in each region of interest included in the statistic stop. As this procedure is not efficient, it was only run for a couple of cases for each treatment modality, or for very specific cases. In clinical routine, the number of histories is defined and is only depending on the treatment modality. For Cyberknife 10^8 primary histories are simulated, while for Tomotherapy 10^7 are adequate. This leads to typical uncertainties of 2 % in 95 % of the voxels in the PTV. The number of particles in the phase-space file, just before entering the patient geometry, varies between treatment modalities and even large variations are possible when comparing different Cyberknife treatment plans, as this is largely depending on the collimator sizes used. Typically, the number of particles in these phase-space files is half of the number of primary histories. Moderato allows the user to select the number of times these particles need to be recycled in the patient calculation (on the order of 10). Calculation times on 40 CPU cores are about 15 minutes for the Cyberknife (using pre-calculated phase-space) and 45 minutes for Tomotherapy. For the Cyberknife M6 with MLC, phase-space cannot be pre-calculated and calculation time becomes comparable to Tomotherapy calculations. The dose is scored in the CT voxels. The resolution of the CT dataset can be lowered in the GUI. For Tomotherapy calculations the voxel size of the CT data exported from the Tomotherapy TPS (the CT data containing the treatment couch) is maintained (1.5x1.5x2 mm^3). For Cyberknife, depending on the number of CT slices the original resolution is maintained or divided by 2 in all directions. The slices are always 1 mm and the in-plane resolution is always 1 mm or even below 1 mm (for the head e.g.). For the Cyberknife simulation, the skin contour is used to delimit the number of voxels to be kept in memory. In the visualization GUI the dose can be recalculated in larger spherical voxels, which comes down to smoothing. This process is only used when the results are really influenced by noise (large PTV tail e.g.).

Delivered dose reconstruction

During a Cyberknife treatment using the Synchrony™ tracking mode for liver or lung tumors, several log files are written. These files contain the actual robot position and the position of the external markers and internal fiducials. For liver treatments the center of mass of the fiducials will serve as a surrogate for the tumor motion (most liver lesions are not visible on the two orthogonal x-ray images). The position of external markers and robot are measured at a frequency of 25 Hz (data were extracted at 10 Hz). The link between external markers and internal fiducials is updated at a frequency chosen by the therapists. The actual contents of the log files have not been validated, as it seems complicated to determine the actual robot position and angles independently. One can imagine that the errors will be very limited as these Kuka robots are known to be extremely precise, as they have been used for several decades in the car industry. In any case, we believe that the usage of these log files will give us a closer approximation of the actually delivered dose, even if these files would contain small errors.

Moderato is able to reconstruct the dose for these moving geometries in an automated way. From the moment the 4D CT (consisting of 5 or 10 phases) of the patient is uploaded into the system, Moderato automatically performs a 4D calculation. When the log files are available as well they are automatically used. If not, the complete treatment plan is applied on every phase of the 4D CT and the dose is warped and summed on the planning CT using ITK [26] and Elastix [27] deformable registration software. This ignores a possible interplay effect between robot and patient movements though [28]. When the log files are present, the synchrony

between marker data (external marker or internal fiducial) and robot position is used to link the actual robot position to the different breathing phases. Then, for each CT phase a specific dose calculation is performed, using only the "active" beams (taking into account the actual position of the robot). Again a dose summation using deformable registration is performed to reconstruct the actually delivered dose. It is also possible to increase the time resolution by interpolating the deformation maps to generate intermediate CT datasets, although this has not proven to be clinically relevant for all studied cases. As shown in figure 2, any baseline shift is taken into account. The selection of the respiratory phase is based on the minimum and maximum level, defined in figure 2. These values follow the trend of the baseline.

For Tomotherapy treatments there are no specific log files, but one can calculate dose on the daily MVCT using the reconstructed sinogram (available from the patient archive). This does not take into account any patient movements during the treatment.

Figure 2. Information available in the log files. The x-axis provides the time in ms (with an arbitrary origin), while the y-axis shows the breathing amplitude in arbitrary units. The graph illustrates how the base line drift is taken into account when selecting the individual breathing phases. The respiratory phase is selected by scaling between the min and the max line (taking into account the difference between inspiration and expiration). As these lines are following the drift in the base line, this will be correctly taken into account when selecting the phase. An example of this selection is shown in the zoom below. The "E" denotes the start of the exhale and the "I" that of the inhale.

Introduction of Moderato in clinical routine

The main purpose of the software was to have an independent dose recalculation for Cyberknife and Tomotherapy plans. In our department a systematic dose recalculation is performed for the conformal plans (performed on our Varian™ Clinacs), using the commercial software package IMSURE QA™ (Standard Imaging, Middleton, USA) that recalculates the monitor units in the prescription point. For Tomotherapy and Cyberknife

treatments a more accurate system is needed providing a full 3D dose validation. Therefore, the MCDE algorithm, programmed at the University of Ghent, was reprogrammed, using an object oriented strategy, as explained above. But next to that we decided that this system shouldn't add additional work in the dosimetry department. Once a plan is finalized it is exported to a dicom server running on the Linux host of Moderato. The directory containing the incoming dicom files is continuously scanned and from the moment all data for a specific patient are available the system launches a Monte Carlo calculation automatically. The status of all cases can be visualized on the so-called Dashboard of Moderato projected on a big screen in the planning room. This window also allows a quick visual check if the export was correctly performed and if all dicom files are available. Once the calculation is finished and the results are available a flag is activated in our patient flow system (RT-Flow^{TM} Surgiqual Institute, Grenoble, France), indicating that the patient plan is ready for validation by the physicist and the physician. One simple click activates the visualization screen containing the isodoses, the DVH data and a table containing dose-volume information for relevant dose-volume points of all delineated organs. This sheet can then directly be included in a Record and Verify system. At the same time the system is removed from the Dashboard. This procedure allows a systematic validation of all Cyberknife and Tomotherapy patients, using the 192 CPU cores on the 3 Linux servers.

MRI-only dose calculation in Moderato

It is well known that MRI provides an optimal soft-tissue differentiation and allows more accurate target volume delineation in modern radiotherapy [29]. Current dose calculation algorithms require electron density, thus MRI and CT data need to be registered, increasing uncertainty, when propagating contours. A direct dose calculation on MRI images can reduce the uncertainty and also simplifies the treatment planning process. Recently introduced radiotherapy specific MRI systems allow positioning the patient in the actual treatment position [29]. A number of research groups have focused on converting MRI datasets to pseudo-CT images that can be used for attenuation corrections in PET/MRI or for treatment planning, assigning bulk densities [31], atlas-based methods [31] or direct conversion of MRI grey levels, using dedicated MRI sequences (e.g. ZTE or DUTE) [32].

As described in a recently submitted paper [33], we are currently working on several scientific projects on MRI-only treatment planning. Moderato plays an important role in the dosimetrical validation of the pseudo-CTs, generated by the atlas-based deformation method, which is an optimization of the method originally introduced by Dowling et al. [31]. We are specifically focusing on Head and Neck patients and thus on the impact of air cavities when using small beams (Cyberknife and Tomotherapy).

Results

Modeling and commissioning of our 3 Tomotherapy units

As the profiles and PDDs in the watertank were almost superimposing perfectly (within 1 %) and as this paper is not really focusing on these modeling issues a detailed comparison is not shown. Instead, we prefer to describe the setup and results we obtained for a direct comparison between TPS and Moderato for individual beams on a phantom. One example is shown in figure 3.a but the same test was performed for all field sizes (1.0, 2.5 and 5.0 cm). A density override was used both in the TPS and in Moderato excluding the impact of the IVDT for these test cases. A worst case scenario was considered by defining a lung density of 0.1 g/cm^3.

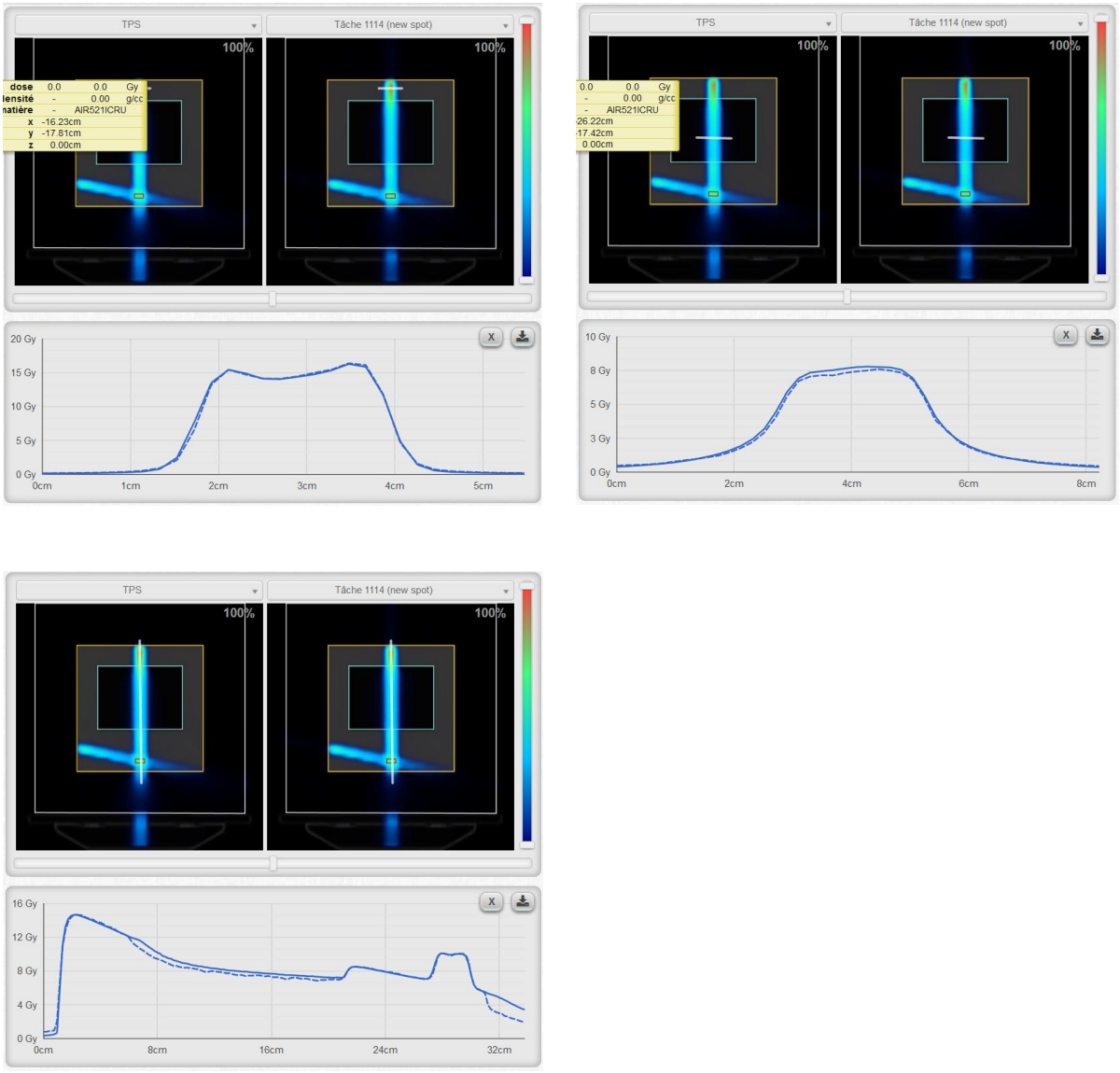


Figure 3. The dose visualisation GUI of Moderato, illustrating the direct comparison between Moderato and the Tomotherapy TPS for an individual beam on a "lung" phantom (the second beam is added because the TPS does not allow defining a single beam). Two profiles are shown. The first profile illustrates the agreement between TPS and Moderato in water. In the lung cavity a difference is observed (confirmed by the PDD below). This difference is because of the limitations of collapsed cone algorithms in this low density (0.1 g/cm^3) region. The differences in the air surrounding the phantom are caused by a different IVDT, but these differences are not relevant.

The profile illustrated in figure 3 is in the direction perpendicular to the leaf and jaw motion, testing the modeled leaf sides, which has proven to be the most critical parameter (actual position read from the xml machine archive, but also the Tongue and Groove). The agreement obtained was (in absolute dose) as good as perfect for all studied cases (within 1 %). The leaf sides extracted from the xml files of our 3 Tomotherapy units were almost identical. So a common model was used for the 3 units.

A comparison of the PDD through the lung cavity is an interesting test for the collapsed cone algorithm used in the Tomotherapy treatment planning system. Even for this very low lung density the precision of the TPS is more than acceptable (see figure 3.b).

To determine the impact of the leaf sides as a function of offset a "picket fence" test was planned on the same phantom as described above. Dose was systematically prescribed to several PTVs while minimizing the dose in the regions between the GTVs in order to force the system to close intermediate leaves.

The result of this comparison is shown in figure 4, again comparing absolute dose, illustrating a perfect agreement (within 1 %).

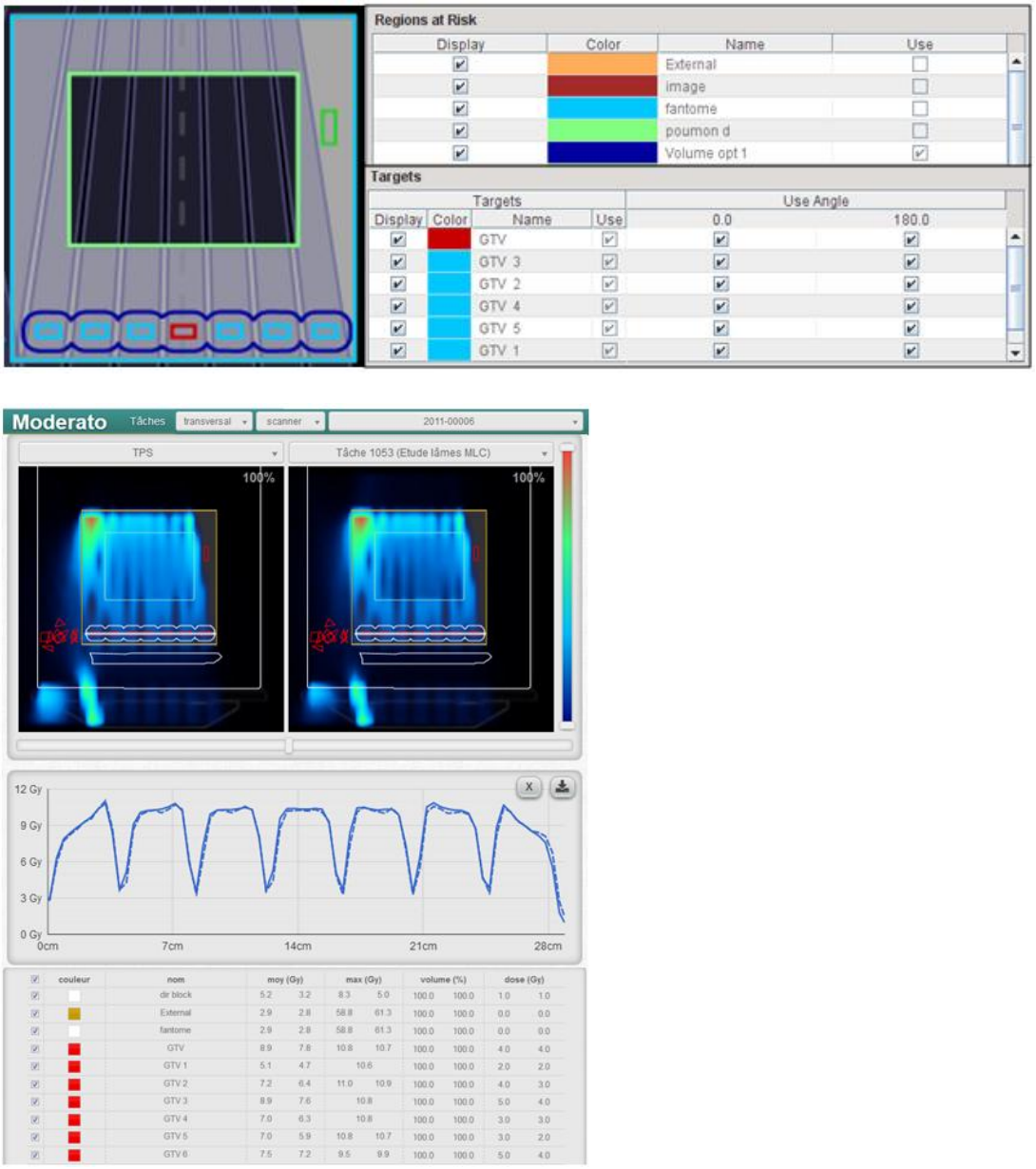
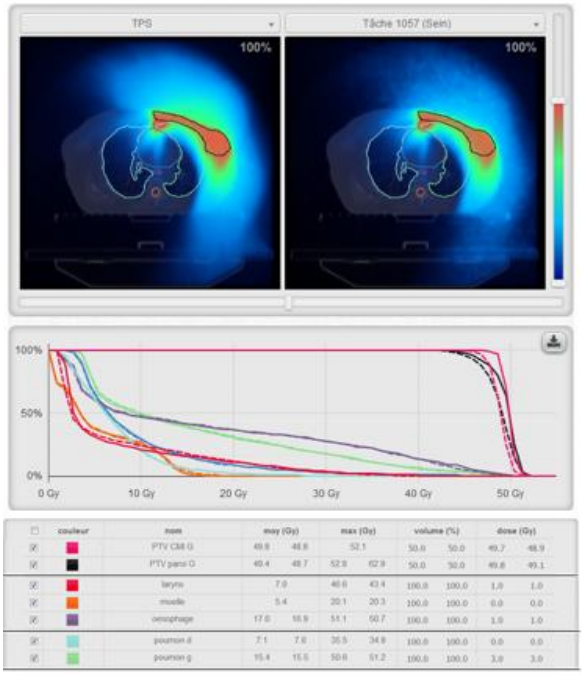


Figure 4: Picket fence test. The image above shows the TPS interface and the definition of the 7 GTVs. Below the results, illustrating an almost perfect agreement between Moderato and TPS. The travel direction of the MLC is perpendicular to the plane. This test is evaluating the leaf sides and the position (perpendicular to the leaf motion). A test parallel to the leaf motion is not necessary as a binary MLC is used.

These results encouraged us to start recalculating patient plans. Again the comparison is absolute and we used the stoichiometric IVDT curve obtained by Demol et al [22]. This curve differs from the IVDT used in the TPS to ensure an independent dose recalculation. One example of such a comparison is shown in figure 5.a for a breast treatment.



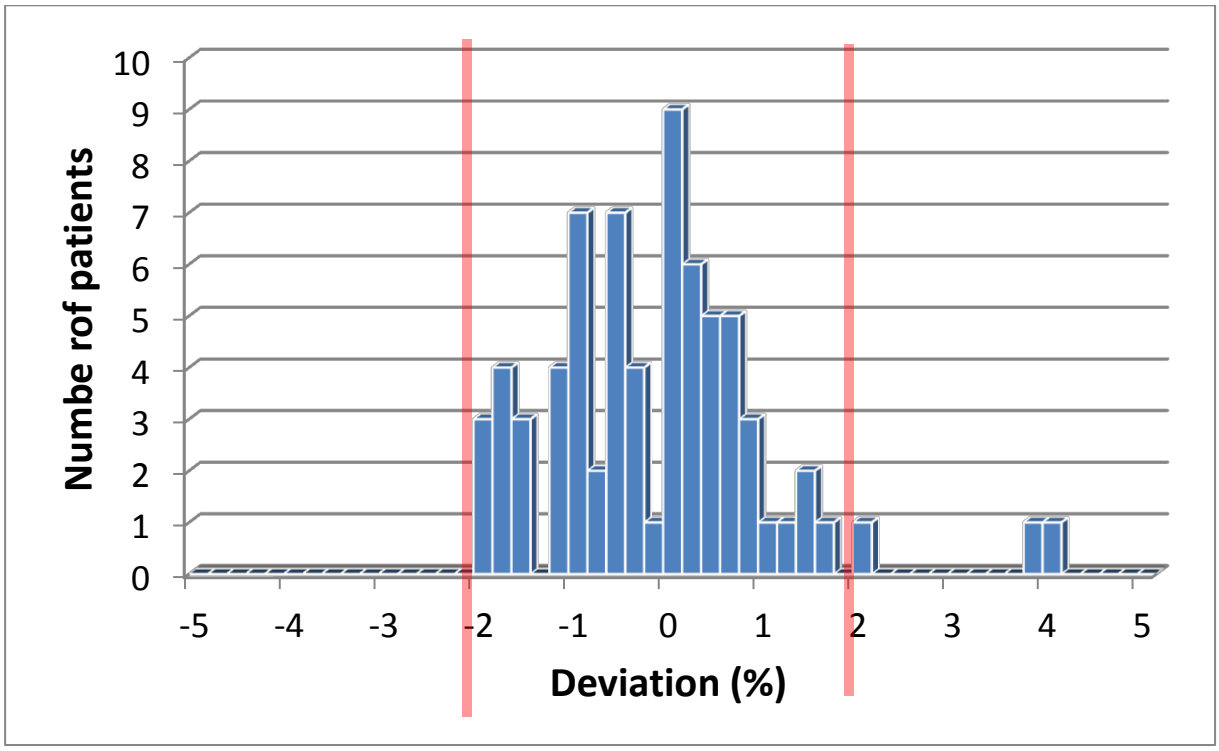


Figure 5. Comparison for a breast patient (above). Below a histogram for 25 targets in 23 patients (mixing different clinical indications) illustrating the agreement of the Tomotherapy model in Moderato. For one patient (having two PTVs) a difference of 4 % is observed because of a problem with the CT data. For all other patients, an agreement between TPS and Moderato within 2 % is obtained.

A study on 23 patients (25 PTVs) is shown in figure 5c. For all patients an agreement between TPS and Moderato within 2 % was obtained for all relevant dose-volume parameters. A gamma test for these cases is considered irrelevant as there is no positioning uncertainty as all doses are calculated on the same scoring grid. For all studied cases, the ray-tracing option provided almost identical dose distributions to the full Monte Carlo

simulation through the MLC (within 1 %). As the Tomotherapy system uses a binary MLC (individual leaves are always fully open or fully closed), the leaf scatter is less important, which explains the fact that the ray-tracing provides correct results.

For one specific breast case an important difference (4 %) between TPS and Moderato was observed. As all treatment parameters were identical to a number of other breast cases, this difference was really patient specific. The only difference with the other patients was the large CT reconstruction diameter (70 cm). Inspecting the scan revealed high HUs in a ring surrounding the patient (leading to voxels in air having a density > 0.1 g/cm^3). Removing this CT artifact again led to a perfect agreement.

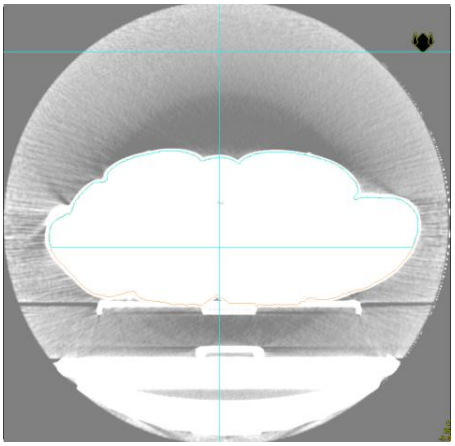


Figure 6. CT data for a large reconstruction diameter (70 cm), giving rise to a ring artifact modeled as high density air in the TPS. The air surrounding the patient has an increasing density when increasing the diameter. At the extreme diameter of 70 cm air has a density of 0.1 g/cm^3. This leads to an overestimation of the attenuation of the beam before entering into the patient. As this effect depends on the IVDT, a large difference (4 %) was obtained between Moderato and the TPS for this specific patient.

Modeling and commissioning of the Cyberknife unit

The modeling of the Cyberknife unit was described by Wagner et al. [24]. In current paper only clinical applications are shown. For most cases the agreement was within 2 %, when comparing to the Monte Carlo results obtained using the Cyberknife TPS (Multiplan). Multiplan allows a direct comparison between the two calculation algorithms available in the TPS (ray-tracing and Multi Plan Monte Carlo). Remark that for the Cyberknife M6, Multiplan uses a third algorithm namely a Pencil Beam algorithm for the MLC plans. Two examples, for different indications, are shown in figure 7.

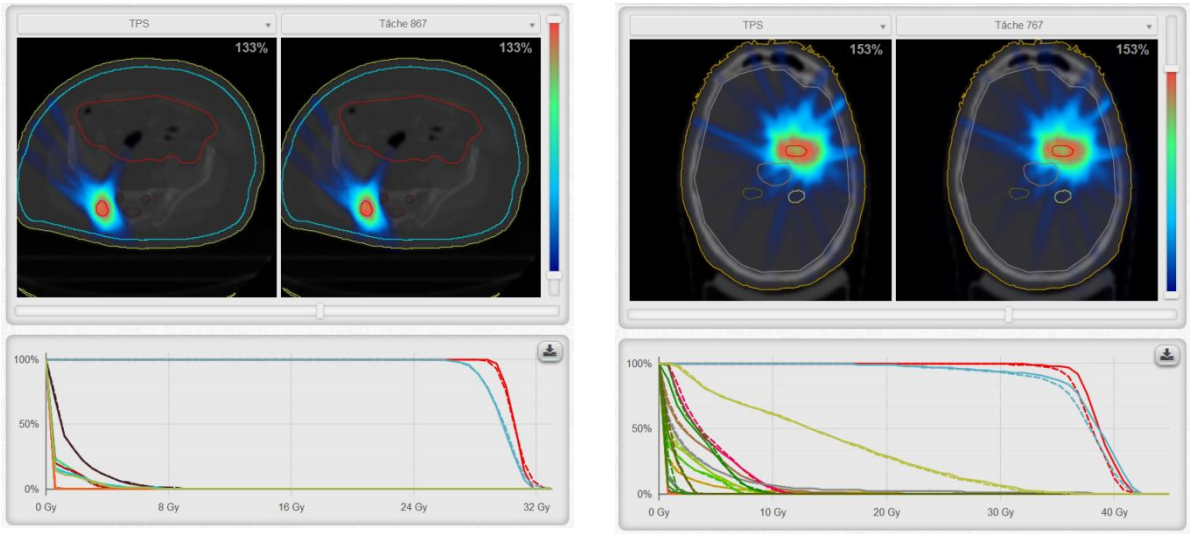
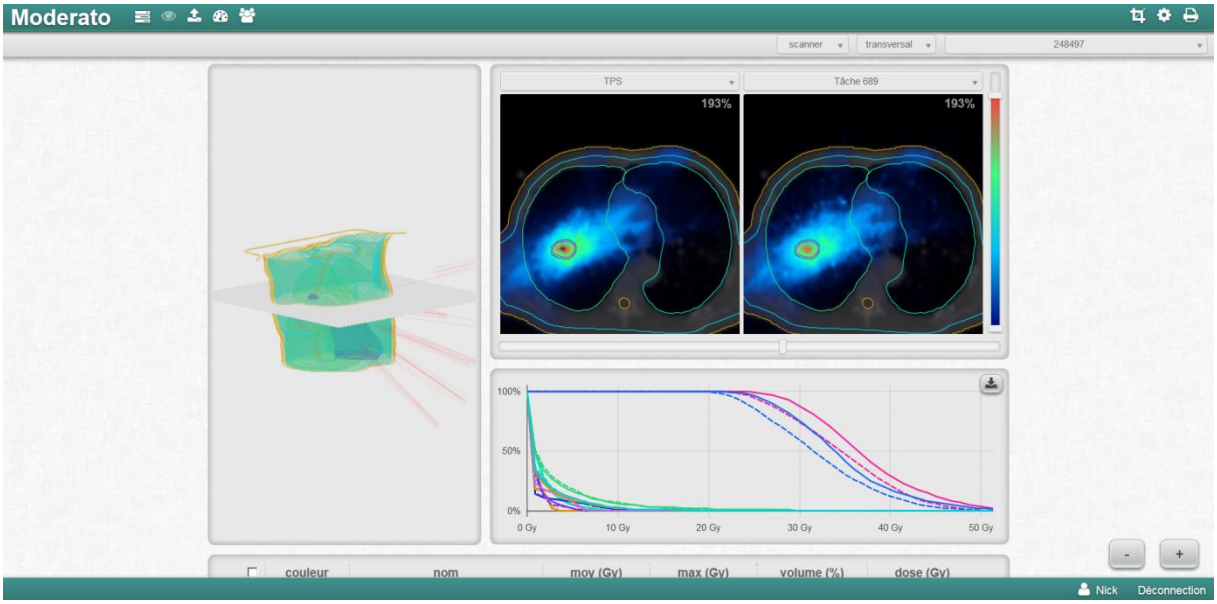


Figure 7. Comparison between Moderato and Cyberknife TPS for two Cyberknife treatments. The first case (left) is a bone metastasis in the pelvic region (iliac), while the second is a brain metastasis. Especially for the first case, small collimators are used (down to 7.5 mm). For the first case the agreement is within 0.5 % at all dose-volume levels. For the second case a difference of 1.8 % is observed for the PTV50.

Even for lung patients, when the difference between the ray-tracing and Monte Carlo options in Multiplan can be very large, the agreement between Multiplan Monte Carlo and Moderato is within 2%. Only for very extreme cases, namely small lung tumors surrounded by low density lung tissue a more important difference is obtained (PTV50 differs by 6 %), because of the different energy cut-off value used for the secondary electron transport (see figure 8).

Repeating the Moderato calculation with a higher ECUT value (50 keV instead of 10 keV) resolves most of the problem, although there still remains a small difference. This is probably because of other approximation applied in the TPS Monte Carlo algorithm. For these extreme cases, these differences can be considered as clinically acceptable.



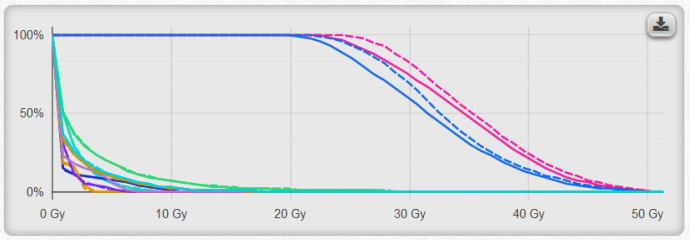


Figure 8. Comparison between Multiplan Monte Carlo (left on isodose plot, and solid line in DVH) and Moderato for a Cyberknife treatment of a small lung lesion surrounded by low density lung tissue, illustrating the impact of using different energy cut-off values for secondary electrons. A PTV50 difference of 6 % is observed between TPS and Moderato. In the figure below, the original Moderato results, calculated with an ECUT of 10 keV (solid line), are compared to a calculation using ECUT = 50 keV. Modifying ECUT does explain most of the deviation.

4D dose calculation: delivered dose

An example for a specific liver patient treated on Cyberknife is provided, showing a comparison between the planned and the actually delivered dose deformed on the planning CT (see figure 9). The differences are very small (within 2 % of the prescribed dose), because of the limited motion, which is probably because of the treatment belt, systematically used for liver patients in our department. This example illustrates the feasibility of the above described method, all results were obtained automatically.

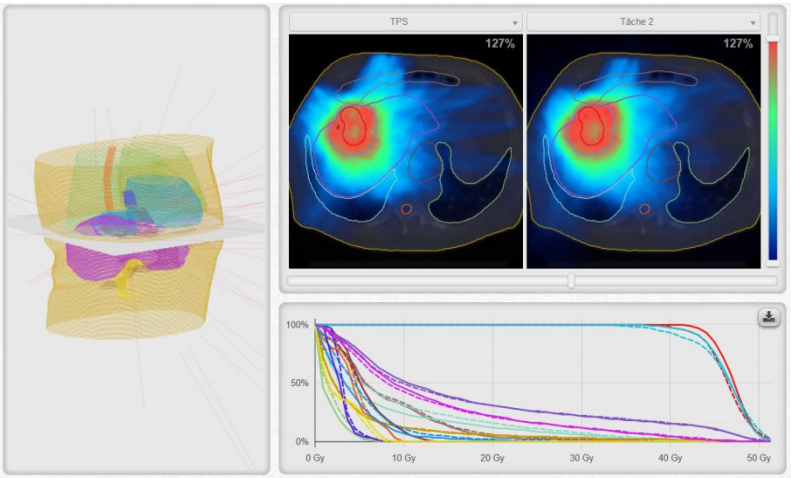


Figure 9. Comparison between planned and delivered dose (both obtained using Moderato) for a liver lesion treated on Cyberknife using the Synchrony tracking system. For the GTV50 a 1.5 % difference is observed (delivered dose < planned dose), while in the colon delivered dose is 4 % higher when normalizing to the local dose (1.5 % when normalizing to the prescription dose). This is the impact of tracking the tumor without taking into account the deformations (beams getting closer to the OAR than during planning).

Dose calculations on MRI

An example of a Monte Carlo calculation on a pseudo-CT dataset is shown in figure 10. The atlas-based method provides in general an agreement within 2 %, when comparing DVHs on actual CT and pseudo-CT. As described in a dedicated paper, for very specific cases [34], when the deformation method cannot accurately reproduce the patient geometry, more important deviations (> 5 %) can be obtained though. One specific example is a patient having part of the skull removed. As illustrated in figure 11, for this case, a simple atlas-based method

can never provide an accurate pseudo-CT. For these atypical anatomies the conversion method should fall back to a direct conversion of MRI grey levels, using tissue segmentation e.g.. The same can be said about air cavities that can vary largely from one patient to another.

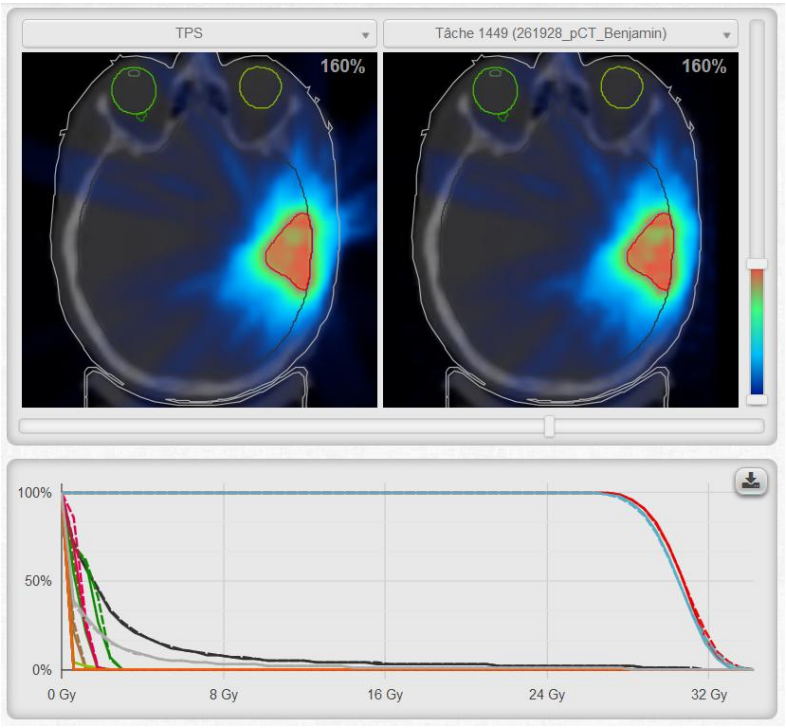
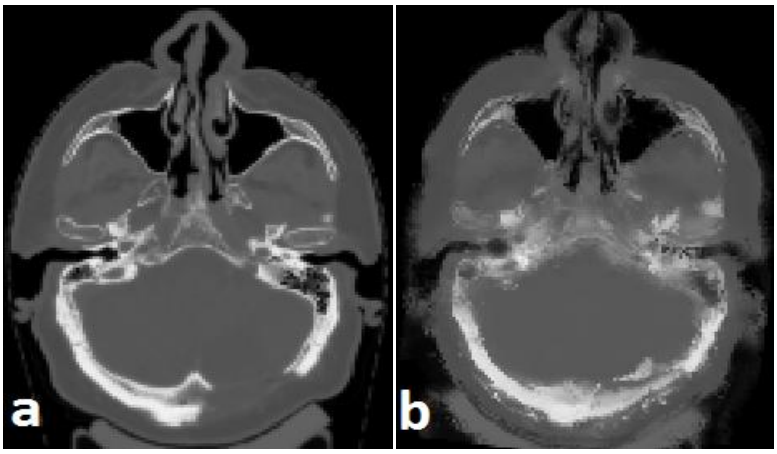


Figure 10: Comparison of dose distributions obtained on a pseudo-CT compared to the real CT, using Moderato. The DVHs are all in agreement within 0.5 %, proving the point that the pseudo-CT can be considered equivalent to the actual CT and would lead to an identical treatment plan. Comparable results are obtained for 90 % of our head and neck patients.



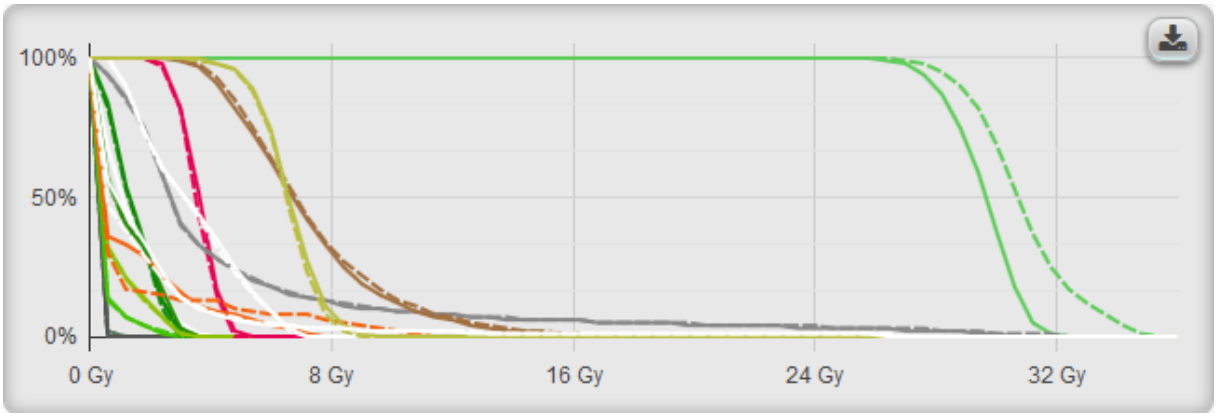


Figure 11: Comparison between actual CT (a) and pseudo-CT (b) for a patient that underwent surgery resulting in the removal of part of the skull. The atlas-based deformation method is not able to reproduce the hole in the skull in the pseudo-CT, which clearly leads to important dose deviations (> 5 %) in the GTV (located next to the hole), as shown in the DVHs below (full line = actual CT).

Discussion

An independent dose validation is a legal requirement in France, for all treatments when this is technically feasible. Up to now, all conventional conformal treatments are validated using a commercial system such as IMSURE QA. Although the system is easy to implement in clinical routine it lacks precision when more advanced treatments need to be validated, as it is using an approximate algorithm to validate a more advanced dose calculation engine. Many radiotherapy departments have developed in-house software systems for these validations, mostly based on Monte Carlo algorithms. To our knowledge most of these systems are not used on a daily basis though, or in any case not for every individual patient. This demands more than a precise algorithm, namely a robust user-friendly system that does not add additional burden to the workflow in the dosimetry department. This is the focus of current paper. Moderato is a Monte Carlo platform that can be accessed by using a simple web page, for uploading patient data, launching Monte Carlo calculations and evaluating the plan comparison in a very user-friendly way. The automation has even further increased the ease of use. Simply exporting the finalized plan data (which is a task that is already performed when exporting to the PACS) to a specific dicom server; launches the calculation using default parameters, and the status can be followed in real-time. The only additional task is handled during the plan validation by the physicist, who

needs to open the visualization page of the GUI and print the report for inclusion in our R&V system (four clicks).

It is often a point of discussion if treatment plans provided by modern TPS algorithms should still be validated by an independent algorithm, as these systems often use advanced dose calculation algorithms such as collapsed cone or even Monte Carlo. Current paper describes the importance of this validation. Even for the small number of patients and even knowing that the dose calculation algorithm used in the Tomotherapy TPS is very precise, an independent dose verification tool has already proven useful for this specific case, illustrating an IVDT problem for large reconstruction diameters.

One can even consider adding additional functionality to the software. Currently, we are working on the automation of the dose prescription step (PTV dose and OAR constraints) which will allow Moderato to automatically validate the plan quality and to highlight specific constraint violations in the table below the DVHs in the Moderato GUI.

A next step is to focus on delivered dose, using all available information saved during the treatment sessions (such as daily images, log files, reconstructed sinogram …). The main limitation of current technique is that everything is based on the 4D planning CT, which is just a snapshot of the breathing motion, not necessarily representative for motion during the actual treatment. Once 4D cone beam CT or even 4D MRI will be available during the treatment, the precision of our dose reconstruction method can even be further increased. As we are interested in a database system containing all patient treatments (“rapid learning”) [34] to construct clinical decision support systems, we need to ensure that the quality of the data is guaranteed. Replacing planned by delivered dose is an important step towards high quality data and can be linked to clinical outcome and toxicity data. The usage of MRI is a second important brick in the construction of a predictive system. That’s why we need to be able to calculate on MRI images and why this is also an important ingredient of Moderato. A dose calculation on a pseudo-CT is not specific though for our Monte Carlo tool, as any TPS can be used for this purpose. We are currently working on a more direct introduction of the MRI images into Moderato. One option could be to convert MRI grey values of a dedicated MRI sequence (ZTE e.g.) into cross sections or by using the hydrogen content as an intermediate parameter as explained in Demol et al. [22].

Conclusion

In current paper the introduction in clinical routine of a Monte Carlo based platform for quality control in radiotherapy was described. The main focus is the possibility to verify the planned dose distribution for every individual patient without adding additional burden to the dosimetrists and medical physicists of the department. This is obtained by a high degree of automation. This allows in routine validation of all patients treated by Tomotherapy and/or Cyberknife, two treatment modalities that demand very precise dose calculation algorithms. Compared to a conventional monitor unit validation tool that only provides the dose in the isocentre, the MC platform provides 3D dose information in the form of isodose information and DVH data. The possibility of switching between different Monte Carlo engines allows simulating different treatment modalities, always using the most adapted algorithm. Next to that, log files, daily images and 4D CT data can be used for calculation of actually delivered dose and dose can also be calculated on MRI. For the moment we are still using pseudo-CT data, but in the near future a more direct link between MRI grey levels and Monte Carlo cross sections will be introduced.

Acknowledgements

Part of the work presented in current paper has been financed by a Siric OncoLille project and a “PhysiCancer” project of the INCA in France

Bibliography

[1] Chen L, Chen LX, Huang SM, Sun WZ, Sun HQ, Deng XW. Independent verification of monitor unit calculation for radiation treatment planning system. Chin J Cancer. 2010; 29(2):217-22

[2] Georg D, Nyholm T, Olofsson J, Kjaer-Kristoffersen F, Schnekenburger B, Winkler P, Nyström H, Ahnesjö A, Karlsson M. Clinical evaluation of monitor unit software and the application of action levels. Radiother Oncol. 2007; 85(2):306-15

[3] Stern RL, Heaton R, Fraser MW, Goddu SM, Kirby TH, Lam KL, Molineu A, Zhu TC; AAPM Task Group 114. Verification of monitor unit calculations for non-IMRT clinical radiotherapy: report of AAPM Task Group 114. Med Phys. 2011; 38(1):504-30

[4] Chan J, Russell D, Peters VG, Farrell TJ. Comparison of monitor unit calculations performed with a 3D computerized planning system and independent "hand" calculations: results of three years clinical experience. J Appl Clin Med Phys. 2002; 3(4):293-301

[5] Fan J, Li J, Chen L, Stathakis S, Luo W, Du Plessis F, Xiong W, Yang J, Ma CM. A practical Monte Carlo MU verification tool for IMRT quality assurance. Phys Med Biol. 2006; 51(10):2503-15

[6] Ahmad SB, Sarfehnia A, Paudel MR, Kim A, Hissoiny S, Sahgal A, Keller B. Evaluation of a commercial MRI Linac based Monte Carlo dose calculation algorithm with geant 4. Med Phys. 2016; 43(2):894.

[7] Fonseca TC, Campos TP. SOFT-RT: Software for IMRT simulations based on MCNPx code. Appl Radiat Isot. 2016; pii: S0969-8043(15)30400-0

[8] Jabbari I, Monadi S. Development and validation of MCNPX-based Monte Carlo treatment plan verification system. J Med Phys. 2015; 40(2):80-9

[9] Chetty IJ, Curran B, Cygler JE, DeMarco JJ, Ezzell G, Faddegon BA, Kawrakow I, Keall PJ, Liu H, Ma CM, Rogers DW, Seuntjens J, Sheikh-Bagheri D, Siebers JV. Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning. Med Phys. 2007; 34(12):4818-53

[10] Reynaert N, Van der Marck S C, Schaart D R, Van der Zee W, Van Vliet-Vroegindeweij C, Tomsej M, Jansen J, Heijmen B, Coghe M and De Wagter C. Monte Carlo Treatment Planning for photon and electron beams (Topical Review), Rad. Phys. Chem. 2007; 76(4):643-686

[11] Reynaert N, Coghe M, De Smedt B, Paelinck L, Vanderstraeten B, De Gersem W, Van Duyse B, De Wagter C, De Neve W, Thierens H. The importance of accurate linear accelerator head modelling for IMRT Monte Carlo calculations. Phys Med Biol. 2005; 50(5):831-46

[12] Reynaert N, De Smedt B, Coghe M, Paelinck L, Van Duyse B, De Gersem W, De Wagter C, De Neve W, Thierens H. MCDE: a new Monte Carlo dose engine for IMRT. Phys Med Biol. 2004; 49(14):N235-41

[13] Sterpin E, Tomsej M, De Smedt B, Reynaert N, Vynckier S. Monte carlo evaluation of the AAA treatment planning algorithm in a heterogeneous multilayer phantom and IMRT clinical treatments for an Elekta SL25 linear accelerator. Med Phys. 2007; 34(5):1665-77

[14] Paelinck L, Smedt BD, Reynaert N, Coghe M, Gersem WD, Wagter CD, Vanderstraeten B, Thierens H, Neve WD. Comparison of dose-volume histograms of IMRT treatment plans for ethmoid sinus cancer computed by advanced treatment planning systems including Monte Carlo. Radiother Oncol. 2006; 81(3):250-6

[15] Vanderstraeten B, Reynaert N, Paelinck L, Madani I, De Wagter C, De Gersem W, De Neve W, Thierens H. Accuracy of patient dose calculation for lung IMRT: A comparison of Monte Carlo, convolution/superposition, and pencil beam computations. Med Phys. 2006; 33(9):3149-58

[16] Walters B, Kawarakow I, Rogers D W O., DOSXYZnrc users manual, NRCC Report No. PIRS-794revB, 2004; unpublished

[17] Rogers D W O, Walters B, Kawrakow .I BEAMnrc user manual, NRCC Report No. PIRS-0509ArevK, 2006; unpublished

[18] Vanderstraeten B, Olteanu AM, Reynaert N, Leal A, De Neve W, Thierens H. Evaluation of uncertainty-based stopping criteria for monte carlo calculations of intensity-modulated radiotherapy and arc therapy patient dose distributions. Int J Radiat Oncol Biol Phys. 2007; 69(2):628-37

[19] De Smedt B, Fippel M, Reynaert N, Thierens H. Denoising of Monte Carlo dose calculations: smoothing capabilities versus introduction of systematic bias. De Smedt B, Fippel M, Reynaert N, Thierens H. Med Phys. 2006; 33(6):1678-87

[20] De Smedt B, Vanderstraeten B, Reynaert N, De Neve W, Thierens H. Investigation of geometrical and scoring grid resolution for Monte Carlo dose calculations for IMRT. Phys Med Biol. 2005; 50(17):4005-19

[21] Sarrut D, Bardiès M, Boussion N, Freud N, Jan S, Létang JM, Loudos G, Maigne L, Marcatili S, Mauxion T, Papadimitroulas P, Perrot Y, Pietrzyk U, Robert C, Schaart DR, Visvikis D, Buvat I. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. Med Phys. 2014; 41(6):064301. doi: 10.1118/1.4871617.

[22] Demol B, Viard R, Reynaert N. Monte Carlo calculation based on hydrogen composition of the tissue for MV photon radiotherapy. J Appl Clin Med Phys. 2015; 16(5):5586

[23] Vanderstraeten B, Chin PW, Fix M, Leal A, Mora G, Reynaert N, Seco J, Soukup M, Spezi E, De Neve W, Thierens H. Conversion of CT numbers into tissue parameters for Monte Carlo dose calculations: a multi-centre study. Phys Med Biol. 2007; 52(3):539-62

[24] Wagner A, Crop F, Lacornerie T, Vandevelde F, Reynaert N. Use of a liquid ionization chamber for stereotactic radiotherapy dosimetry. Phys Med Biol. 2013; 58(8):2445-59

[25] Chen Q, Chen Y, Chen M, Chao E, Sterpin E, Lu W. A slit method to determine the focal spot size and shape of TomoTherapy system. Med Phys. 2011; 38(6):2841-9

[26] Johnson H J, McCormick M, Ibánez L The ITK software guide. Kitware Inc.,. 2013 www.itk.org

[27] Klein S, Staring M, Murphy K, Viergever M A, Pluim J P W, elastix: a toolbox for intensity based medical image registration, IEEE Transactions on Medical Imaging. 2010; 29(1), 196 - 205

[28] Litzenberg DW, Hadley SW, Tyagi N, Balter JM, Ten Haken RK, Chetty IJ. Synchronized dynamic dose reconstruction. Med Phys. 2007; 34(1):91-102

[29] Devic S. MRI simulation for radiotherapy treatment planning. Review paper. Int J Radiat Oncol Biol Phys. 2010; 78(5):1555-62.

[30] Pasquier D, Betrouni N, Vermandel M, Lacornerie T, Lartigau E, Rousseau J. MRI alone simulation for conformal radiation therapy of prostate cancer: technical aspects. Conf Proc IEEE Eng Med Biol Soc. 2006; 1:160-3

[31] Dowling JA, Lambert J, Parker J, Salvado O, Fripp J, Capp A, Wratten C, Denham JW, Greer PB. An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy. Int J Radiat Oncol Biol Phys. 2012; 83(1):e5-11

[32] Johansson A, Karlsson M, Nyholm T. CT substitute derived from MRI sequences with ultrashort echo time. Med Phys. 2011; 38(5):2708-14.

[33] Demol B, Boydev C, Korhonen J, Reynaert N. Dosimetric characterization of MRI-only treatment planning for brain tumors in atlas-based pseudo-CT images generated from standard T1-weighted MR images. Med. Phys. 2016; Under revision

[34] Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, Zegers CM, Carvalho S, Leijenaar RT, Nalbantov G, Oberije C, Scott Marshall M, Hoebers F, Troost EG, van Stiphout RG, van Elmpt W, van der Weijden T, Boersma L, Valentini V, Dekker A. 'Rapid Learning health care in oncology' - an approach towards decision support systems enabling customised radiotherapy'. Radiother Oncol. 2013; 109(1):159-64.

3. Article 2

The impact of different sources of uncertainty on stereotactic radiotherapy of liver tumors using real-time tracking

M. Charoy
Department of Medical Physics, Oscar Lambret Cancer Center, Lille, France
T. Lacornerie
Department of Medical Physics, Oscar Lambret Cancer Center, Lille, France
P. Jordan
Accuray
A. Wagner
Department of Medical Physics, Oscar Lambret Cancer Center, Lille, France
X. Mirabel
Department of Radiation Oncology, Oscar Lambret Cancer Center, Lille, France
N. Reynaert
Department of Medical Physics, Oscar Lambret Cancer Center, Lille, France

Abstract

Purpose: The liver is a mobile organ that undergoes significant movements and deformations during the respiratory cycle. In order to reduce errors introduced by respiratory motion, the majority of hepatic lesions treated on the CyberKnife® system are tracked using the Synchrony® respiratory tracking system, using implanted fiducial markers. The movements of the target are considered identical to those of the fiducials. The correlation between target and fiducial motion must be verified to ensure the accuracy of this type of treatment.
Materials and Methods: Deformable image registration was used to define the GTV on the CT images of the different breathing phases, which allows comparing the motion of fiducials and GTV. The dosimetric impact of real-time tracking was determined, for different GTV-to-CTV margins and compared to the uncertainties of GTV delineation on CT and MRI. As real-time tracking without 4D plan optimization is not robust, the impact on the dose received by the organs at risk was evaluated. The impact of rotations and deformations (not taken into account by the Synchrony tracking system) was evaluated as well. All dose calculations were performed using our in-house Monte Carlo platform that allowed simulation of different tracking methods.
Results: For all patients, the movements of the center of mass (COM) of fiducials and target are the highest in the z-direction (superior-inferior) with maximal amplitude of 14.3 mm. The movements are globally low, thanks to the usage of a belt during treatment delivery. The impact of tracking on the dose received by the organs at risk is limited. Taking into account rotations and deformations into the tracking system does not further improve target coverage. The added value of real time tracking depends on the GTV to CTV margin. When no margin is applied (direct GTV to PTV margin) tracking really provides a better target coverage. Uncertainties related to target delineation are of the same order as those related to not tracking the tumor and are especially important when no CTV margin is defined. It was also shown that movements on 4D PET and 4D CT are not well correlated, illustrating that 4D CT is not an optimal tool for evaluating the breathing motion.
Conclusion: Real-time tracking is very accurate even though rotations and deformations are not considered. When the fiducials are correctly positioned around the lesion, fiducial tracking is as accurate as direct target tracking. But the added value of tracking depends on the applied CTV margin. Accuracy of target delineation of liver tumors should be further increased, otherwise tracking does not make much sense.

Key words: breathing movements - CyberKnife - correlation model – liver- target delineation

Introduction

Uncertainties of various types affect the quality of stereotactic radiotherapy. Some uncertainties are related to physical processes, such as the conversion of CT data to electron densities and the weaknesses of the dose calculation algorithm. Other uncertainties are inherent to the treatment machine and combined imaging systems, to the patient positioning, and to the target and OARs delineation. All uncertainties can potentially have a major impact on the dose to the target volumes and to critical normal tissues which could result in significant alteration of probabilities of tumor control and normal tissue complications [1].

Accurate structure delineation is vital to the success of radiation techniques, because it is the basis of treatment planning. Increasing target contouring accuracy will reduce the overall uncertainty on the delivered dose. It is particularly important for stereotactic radiotherapy by CyberKnife™ (Accuray Incorporated, Sunnyvale, CA), which allows delivering precisely a high dose per fraction to the target. Delineation of liver lesion and OARs is commonly performed on 3D CT images [2]. The liver lesion is often difficult to distinguish because of the low density difference between a hepatic lesion and healthy liver tissue. When possible, iodine is used to enhance image contrast for improving the delineation reproducibility. The inter-observer variability in target delineation, which can be attributed to many factors including the influence of the observer (specialty, training, and personal bias) and the impact of the imaging modality should also be decreased using contrast enhancement. Access to increasingly innovative imaging technologies enables continuous improvement of lesion visualization. Magnetic resonance imaging (MRI) allows the acquisition of several imaging sequences with specific characteristics. It is well known that MRI offers a better soft tissue contrast than CT. Especially for liver tumors MRI plays a very important role and is often combined with CT for target delineation.

Breathing is another important source of uncertainty in stereotactic radiotherapy, especially for tumors in the abdominal regions, such as the liver. With classical radiotherapy, tumor movements obligate applying a larger margin, e.g. an ITV, using 4D CT information [1]. The Synchrony Respiratory Tracking System® (SRTS), available on the CyberKnife, enables real-time adjustments with changes in the tumor position during the treatment delivery. Several internal fiducials are implanted in the liver and infrared diodes are placed on the patient abdomen. A correlation model between the movements of internal fiducials and external markers is built using infrared camera signal and x-rays images. The x-rays images are taken each 80 seconds during the treatment in order to adjust the model when the patient breathing pattern changes. The Synchrony system allows a sub-millimeter precision of the delivery location of the beam [3][2]. As discussed in some studies, liver motion occurs primarily in the superior-inferior (SI) direction in the order of 10 mm [4][5][6][7]. The movements in the left-right (LR) and antero-posterior (AP) directions are lower. These breathing movements are limited with an elastic abdominal belt during the treatment by CyberKnife. Previous studies show a good correlation between internal fiducial motion and external marker motion for liver [4][8][9]. They report that when the fiducials are widely spaced, the correlation is lower because of the presence of liver deformations. Implanting the fiducials inside the tumor volume is not possible because of the risk of tumourous cells spreading when pulling out the needle. The distance of the internal markers to the target and the liver movements and/or deformations can introduce differences in movement of fiducials and target. The correlation between target and fiducials movements is considered exact when the internal markers are correctly arranged, i.e. the three or

four markers encircle the tumor at a distance of less than 60 mm. The fiducials rotations are calculated during patient free breathing and the patient is positioned taking into account the median rotations. Thus, the CyberKnife robot refers to the translations of the fiducials center of mass to adapt the beam position to track the movements.

The purpose of current study is to determine the quality of the correlation between the fiducials and target movements during the patient breathing cycle and to quantify the correlation uncertainties. A bad correlation could introduce errors in beam positioning and thus a difference between the prescribed dose and the dose actually delivered. A better knowledge of the movements between target and fiducials will allow validating the current Synchrony treatment mode, that is currently ignoring rotations and deformations and to improve it by making adjustments if required. Next to that, the accuracy (in terms of reproducibility) of target delineation for liver tumors is evaluated both for CT and MRI and the uncertainty will be compared to that of the real-time tracking. The dosimetric impact of all uncertainties will be determined using an in-house Monte Carlo platform [13]. Next to that we will focus on the impact of the CTV margin defined.

II. Materials and Methods

II.1 Reproducibility of target delineation using CT

II.1.1 CT Acquisition

The 3D CT images were acquired with the Aquilion^{TM} LB (Toshiba Medical Co Ltd, Tokyo, Japan). The first part of the examination was performed without contrast product and consisted of a large acquisition in low resolution, to determine the field of view. The second image set was acquired with a smaller field of view and a higher resolution. In a second time, OptiJECT 350 (Guerbet, France) product containing 350 mg/mL of iodine, was injected in the patient to enhance the image contrast. Several image series were acquired during all phases of the liver transit, including arterial, portal and delayed phases. The CT device locates the time of the first acquisition after injection through the density value of the abdominal aorta, determined from very regular shots. Commonly, the arterial, portal and delayed phases were respectively acquired around 25 seconds, 1 minute and 3 minutes after the beginning of injection. All 3D CT datasets were acquired in blocked respiration to avoid breathing motion artifacts. The raw data obtained after acquisitions were reconstructed with a slice thickness of 1 mm for injected images and 3 mm for non-enhanced images.

II.1.2 Target delineation on CT images

The target delineation was performed twice on enhanced CT images (portal phase), by a radiation oncologist. The second delineation was performed after a sufficiently long time to ensure that the oncologist is not influenced by a previous visualization of the lesion. A second radiation oncologist has delineated the GTV on the same enhanced CT images, which allows comparing the target contouring reproducibility between two physicians. Routinely, CT images are acquired with injection of contrast product, but the contrast product can give problems of renal toxicity and allergic reaction for some patients. The CT acquisition is performed without contrast enhancement for these cases.

The impact of the contrast product on Gross Tumor Volume (GTV) delineation is studied. The GTV contouring performed on 3D CT images without iodine product injection and on the images of the portal phase, by the same radiation oncologist, was compared. CT acquisitions with and without contrast product have been performed in the same coordinate system during the same examination. No registration is needed between the two image datasets before comparison.

II.2 Reproducibility of target delineation using MRI

II.2.1 MRI acquisition

The MRI device used was a Discovery^{TM} MR750 3.0T (GE Healthcare, Waukesha, WI). The patients were in supine position without immobilization tools. Different sequences were acquired in blocked respiration for liver lesion visualization. Images without contrast product are firstly acquired. Afterwards, a product is injected to the patient to enhance the image contrast. The product can be DOTAREM (Guerbet, France) or MULTI-HANCE (Bracco Imaging, France) and contains 20mL with 0.5 mmol/mL of Gadoteric Acid. The injection was followed by acquisition of several sequences in blocked respiration, such as the T2 enhanced and the Multi ART 20S, which corresponds to 3 acquisition times with 20 seconds between two acquisitions from the arterial time.

II.2.2 Target delineation on MR images

A radiation oncologist performed the target delineation twice. The physician has firstly visualized all available sequences and performed the contours alone, on the sequence that provides the optimal target visualization potentially using other sequences. The sequence chosen for the delineation differs from one patient to another. The best MR sequence to visualize the lesion depends on the patient and the image quality (the sequences mostly used were the Axial T2, Portal Phase, multi ART 20S and the Lavaflex Water).

II.2.3 Radiologist support

The interpretation of MR images is a complicated exercise due to the specificities of different sequences. The intervention of a skilled radiologist to assist the radiation oncologist is needed to have an optimal target delineation. The radiologist reviewed the contours previously performed by the radiation oncologist.

II.3 Evaluation of the delineation reproducibility

Different metrics could be used for the similarity evaluation between two 3D objects. The first metric chosen for this study is the Dice Similarity Coefficient (DSC) that quantifies the spatial agreement of a contender- and reference-Volume Of Interest (VOI) according to the whole volume. The DSC can range from zero to one, where zero represents no agreement and one represents perfect agreement

between the two volumes. The DSC values are obtained using Equation 1, where C_r and C_n are the two structures :

$$DSC = 2\,\frac{\mathrm{Cn} \cap \mathrm{Cr}}{\mathrm{Cn}+\mathrm{Cr}} \qquad (1)$$

The DSC value represents the overlap ratio between two structures but it does not inform on how they overlap. A same DSC value could be obtained between two 3D structures perfectly aligned but with a different absolute volume, as between two 3D structures of the same size but slightly shifted. Two other metrics, the Volume Ratio (Rv) and the distance between the centers of mass (dCOM), have been chosen to distinguish the different possible situations. The overlap evaluation between 3D volumes previously delineated, was performed using custom software developed in Matlab (Mathworks Inc, Natick, MA). A registration is required for the comparison of 3D volumes defined on two different sets of images based on different coordinate systems or acquired at two different times. This is the case for the study of the impact of the contrast product on GTV delineation, using two image sets acquired at different phases of the respiratory cycle. Several registration methods are available. The first performs a translation of the fiducials COM, based on the fiducial coordinates in the two image sets. The second method includes the fiducial rotation in the calculation of the registration matrix. This method was performed using the Absor function, an already existing tool encoded in Matlab. This tool uses Horn's quaternion-based method for finding the rotation, translation, and optionally also the scaling, that best maps one collection of point coordinates to another in a least squares sense. The program requests the structure files in DICOM format and the text files containing fiducials coordinates corresponding to the images as input.

II.3 Correlation movements fiducials and target

II.3.A Patient characteristics

The images of 9 patients (four women, aged from 54 to 74 years, and five men, aged from 64 to 73 years) treated for metastatic liver tumors were analyzed. The number of tumors ranged from one to three, and these were located in different segments in the liver. The patients were treated on the Cyberknife system using the Synchrony respiratory tracking system, with three fractions of 15 Gy. On average four internal markers were implanted percutaneously in healthy liver tissue surrounding the tumor using x-ray image guidance. The gold fiducials have a length of 4 mm and a diameter of approximatively 1 mm. The 4D combined positron emission tomography and computed tomography (PET/CT) images were acquired more than a week after the fiducials implantation to minimize errors due to fiducial migration. The distance between the fiducials and the lesion met the recommendation regarding the maximum of 6 cm except for three targets, as shown in Fig.1. These were the smallest targets in current study.

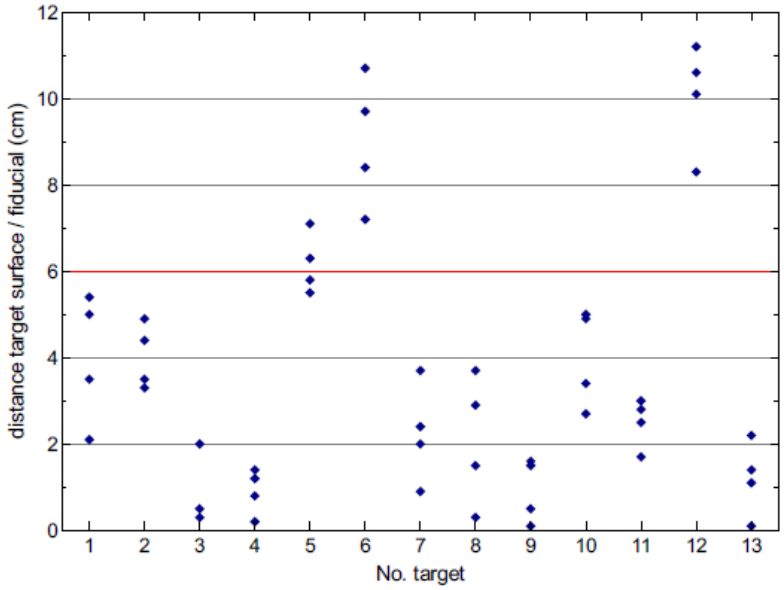


Fig 1. Minimal distance between the segmented volume (SV) surface and the fiducial

No. target	No. patient	Gender	Age	Tumor(s) location	Tumor size (cc)	Radiotracer
1	1	M	64	metastasis SI	29.8	FDG
2	2	F	55	metastasis SVIII	17.3	FDG
3	3	F	54	metastasis SVI	16.3	FDG
4	4	M	73	metastasis SVII	69.5	FDG
5				metastasis SII	3.2	FDG
6	5	M	64	HCC SVII	2.0	FCH
7	6	F	74	HCC junction SV-SVIII	16.4	FCH
8	7	M	65	metastasis SVIII	18.4	FDG
9	8	F	60	metastasis SVIII	27.9	FDG
10				metastasis SIV	7.2	FDG
11				metastasis SVI	20.1	FDG
12	9	M	66	HCC SIII	3.0	FCH
13				junction SIV-SVI	60.7	FCH

Table 1 - List of patient characteristics

No. target	Target COM			Fiducials COM		
	x	y	z	x	y	z
1	0	2	6	1	3	8
2	0	3	2	1	2	6
3	1	8	15	2	7	14
4	0	2	4	1	2	5
5	2	2	4			
6	3	4	9	2	3	6
7	2	1	7	1	4	9
8	2	5	10	1	5	9
9	1	0	6	1	1	7
10	1	0	6			
11	1	0	3			
12	0	2	3	2	4	10
13	1	4	8			

Table 2 - Maximal amplitude of fiducials COM and target COM movements in the directions x, y and z (respectively left-right, anteroposterior and craniocaudal) in millimeters, during the entire respiratory cycle, for the nine patients

II.3.B 4D PET/CT

The first option considered was the usage of 4D PET/CT images. The GTV is visible on the 4D PET images, while the fiducials can be visualized on the 4D CT. The 4D PET/CT images were acquired with the Discovery 690 PET/CT (GE Healthcare, Waukesha, WI) coupled with the Varian® Real-time Position Management™ system (RPM) (Varian Medical Systems, Palo Alto, CA). The RPM system is a video-based system: an infrared camera tracks a block with reflective markers, placed on the patient chest or abdomen, to measure the movements in the antero-posterior (AP) direction.

Two different modes of data acquisition were available with the RPM system; the prospective and retrospective mode. In our study, the series of CT scans are acquired in cine mode, with the retrospective approach. Images are continuously collected during several entire respiratory phases with the patient breathing freely. Each image corresponds to a specific instance of the breathing cycle so that 4D PET and CT images were reconstructed by binning images using phase-based sorting of the RPM system. The retrospective 4D-CT produced 5 or 10 phases according the protocol used. The “0% phase”, the first phase of the cycle, corresponds to the end-inspiration. The images were reconstructed with a slice thickness of 2.5 mm. The patients were scanned in the supine position and the acquisition of 4D images was performed during patient free breathing. The two 4D modalities were acquired sequentially, the 4D CT first. The whole 4D PET-CT scan duration was approximately 20 minutes, including less than a minute for the CT scan and 15 minutes for the PET scan. The time interval between CT and PET acquisition was on the order of one minute. An abdominal compression was used to minimize the breathing amplitude movements during the acquisition. As the target volume is not visible on the CT scans, the GTV was segmented on the PET images and duplicated on the CT images. The Oncentra MasterPlan Planning System (Nucletron, Elekta) was used for segmentation. To have the best visibility of the target volume, an arbitrary threshold was fixed for the image display characteristics “center” and “width”. The PET signal issued from the lesion was the same for all phases because the PET images were sorted in equable temporal phases. The same threshold was used for all phases to compare the characteristics of the target volume. After delineation, the target contours were copied on the CT images and interpolated on the intermediate slices. After the segmentation was performed, the segmented volume (SV) contours of each phase were compared by registration of each CT phase to the primary CT phase. This procedure was carried out in custom software developed in Matlab. Several registration methods were investigated. The first registration performed a translation of the fiducials COM of the two phases, that allowed determining the translation movements similarities between target and fiducials. The second registration method performed a translation of the target COM. This method allowed observing the rotation and/or deformation undergone by the target. The third registration method looked for correction of the fiducials translation and rotation. This method was performed using the Absor function, an already existing tool encoded in MatLab.. The comparison of the results obtained with the first and third registration methods will highlight the role of the rotation correction. A perfect correlation between the SV and the fiducials movement should lead to an exact overlap of the SV contours of all respiratory phases, if the target does not undergo deformations.

II.3.C Using deformable image registration

II.3.C.1 DIR method

This method consists of applying a Displacement Vector Field (DVF) to the target contour of the planning CT to determine the target contour on each phase of the 4D CT images. The DVF was determined by deforming the planning CT to the 4D CT, using ITK/Elastix [10][11], using the BSpline transformation method. As metric a combination of the "AdvancedNormalizationCorrelation" and "TransformBendingEnergyPenalty" was used. The fiducials were removed from the images before deformation to avoid introducing a bias. The contours were propagated on the 4D CT using a MatLab script developed internally.

II.3.C.2 Evaluation of the deformation

The three DVF matrices were applied to the fiducial coordinates. The fiducials of the registered 3D planning CT and those of the target phase (4D CT) should be exactly overlapping. Secondly, the DVFs were applied to the contour points of the kidney, and the obtained volume is compared to the contours manually delineated. The test is performed on the kidney which provides a volume that can easily be contoured reproducibly. The differences between both structures are measured with the absolute volume difference, the COM location difference and the Dice coefficient.

II.3.C.3 Correlation evaluation

To test the accuracy of the Synchrony tracking mode, the correlation between the movements of target and fiducials is evaluated. This comes down to assessing the relative position between target volume and fiducials for each phase. For that, the delineated volume of each respiratory phase was compared to the delineated volume of the primary phase, using a rigid registration based on the location of the fiducials. This procedure was performed using custom software developed in MatLab. The MatLab program allows choosing the fiducials used to perform the registration between two sets of images, mimicking the choice of the fiducials that are "activated" during treatment in the Synchrony mode. Program automation has been performed to enable testing all possible combinations of selected fiducials providing Dice and COM distance for each possible combination. In that way the selection of the best combination of fiducials can be obtained before starting the treatment.

II.4 Monte Carlo dose calculations evaluating the impact of uncertainties

II.4.1 Impact of delineation uncertainties

In order to determine the dosimetric impact, a shift of 6 mm in the 3 directions was applied to the GTV to simulate the target delineation uncertainty. The DVH, calculated using the Monte Carlo software (Moderato) [13], allows determining the dose coverage of the shifted GTV when the PTV is targeted. Only one patient is considered for this test. The applied shift of 6 mm (= global shift of 11 mm) corresponds to the extreme case of the intra-physician contouring uncertainty.

II.4.2 Dosimetric effect of tracking motion on OARs

The dose grids calculated on the individual phases of the 4D CT were warped and accumulated on the planning CT using the DVFs. The dose calculations were performed using the in-house Monte Carlo platform (Moderato). The system allows an independent verification of the dose distribution calculated by the treatment planning systems of Cyberknife and Tomotherapy. The MC dynamic dose calculation based on a treatment plan and on 4D CT images, gives a more realistic result of the dose actually delivered taking into account the robot and patient movements during the treatment. The impact of tracking the target on the dose to the OARs was evaluated and the dosimetric impact of the potential inconsistency between the movements of target and fiducials was determined, by comparing 3D and 4D dose calculations.

II.4.3 Accuracy of fiducial tracking

The accuracy of the Synchrony mode can be evaluated by comparing a MC calculation performed on each phase of the 4D CT with all beams shifted using:

- the COM of the fiducials (simulating Synchrony tracking)

with

- the COM of the target (simulating a direct tumor tracking).

All calculated dose distributions are registered on the planning CT using deformable image registration.

II.4.4 Utility of respiratory tracking

The comparison between the 4D dose calculation with and without tracking of the fiducials COM highlights the utility of respiratory tracking. The same comparison as described above is performed, while summing the dose without deformable image registration.

II.5 Impact of CTV margin

The impact of uncertainties regarding delineation and tracking will largely be determined by the applied CTV margin. In clinical practice a margin of 5 mm is added from GTV to CTV (clinical target). This is (following the ICRU guidelines [12]) partially to take into account uncertainties in GTV delineation and partially to include microscopic spread of tumor cells. In some centers though, no CTV is defined and the PTV is directly determined using a 3mm margin from GTV. A new plan was performed targeting this new PTV (named PTV margin).

II.5.1 Impact of CTV margin on delineation uncertainty

The impact of the uncertainty of the delineation reproducibility of the GTV will be decreased considering the CTV. If one of the GTVs is considered as correct (e.g. GTV2), while the other (GTV1) is the one actually delineated, then one should determine the overlap between GTV1 plus margin (= CTV1) with GTV2, to have an idea regarding the clinical impact. This was systematically investigated for spherical volumes, determining the intersection geometrically, knowing the radius of the two spheres and the distance between their center of mass. This leads to a general model that can be applied afterwards to the clinical cases, studied in current paper. The actual shape of a liver target is not perfectly spherical, but our model provides a good approximation as we are especially interested in a global impact of the applied margins.

II.5.2 Impact of the CTV margin on tracking

For the same reasons the importance of tracking will increase when no CTV margin is used. This was investigated by performing the MC comparisons, described above, for a PTV defined as the GTV with a margin of 5 mm.

III. Results

III.1 Reproducibility of target delineation using CT

As illustrated in figure 2, reproducibility of target delineation is rather poor. The dice coefficient is low (< 0.8) for half of the patients and distances between the centers of mass is > 5 mm for half of the cases. As expected the inter-observer reproducibility is lower than when having the delineation performed twice by the same physician, but even this is far from adequate. The same test was performed on patients that were scanned without contrast product, which lead to even worse results (not shown). Even the ratio of volumes went up to a factor of 2 for certain cases when no contrast product was used. In principle contrast product is always used, unless patients do not tolerate it (allergy). The dosimetric impact of the contrast product was negligible, so the treatment plans were performed directly on the images obtained with contrast product.

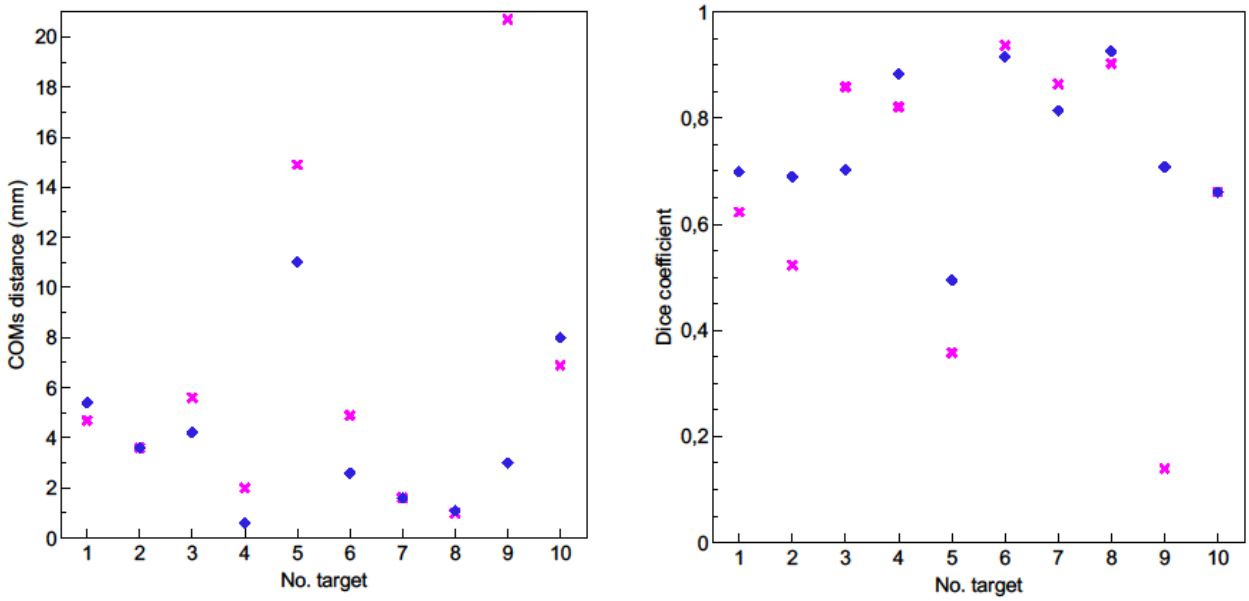


Figure 2 : Distance between COMs (left) and dice coefficient (right) of target volumes delineated on CT with contrast product, twice by the same radiation oncologist (diamonds) and once by two different radiation oncologists (crosses), for all targets.

III.2 Reproducibility of target delineation using MRI

Target delineation was proven to be more reproducible on MRI as shown in figure 3, using one specific sequence.

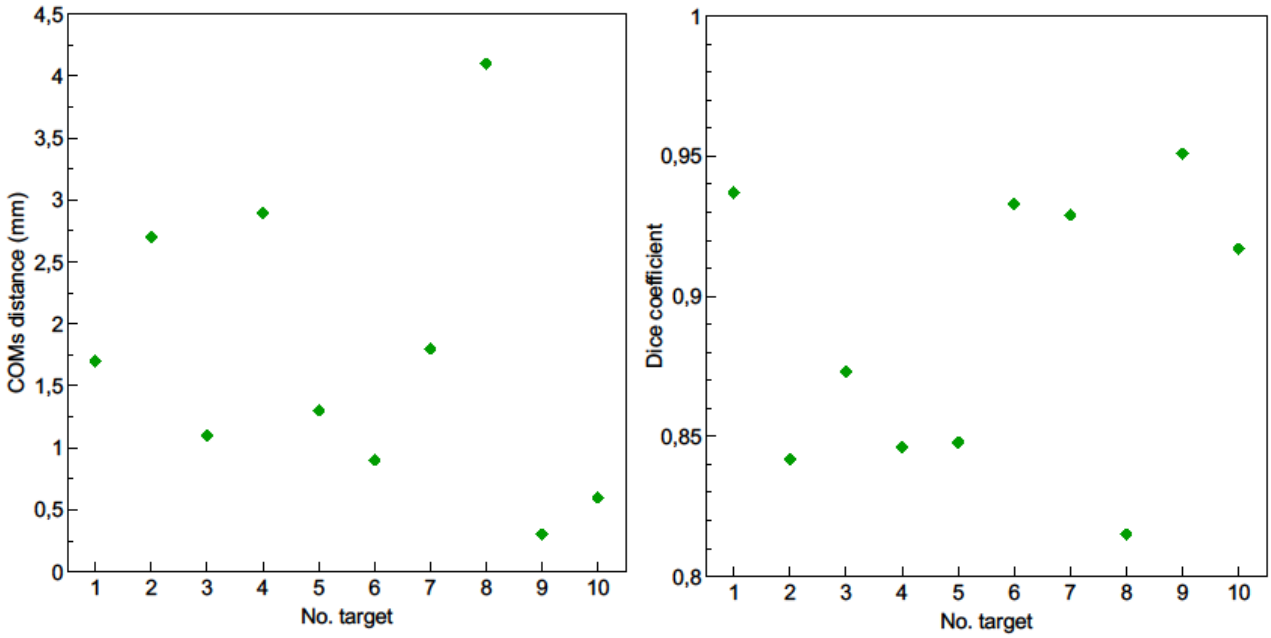


Figure 3 : Distance between COMs (left) and dice coefficients (right) of volumes twice delineated on MRI, by the same radiation oncologist, for all targets.

The dice coefficient is systematically higher than 0.8 and the distance between the centers of mass is much smaller than on CT. There are still a couple of cases where a distance > 3 mm is obtained, illustrating that even on MRI, target delineation for these liver lesions is extremely difficult.

Delineation results are summarized in table 3.

	dCOM (SD) (mm)	dice (SD)	Rv (SD)
comparison of contouring CT without IV - CT with IV	6.1 (4.9)	0.665 (0.271)	1.38 (0.48)
reproducibility of contouring on CT with IV	4.1 (3.3)	0.749 (0.135)	1.09 (0.29)
comparison of contouring by two physicians on CT with IV	6.6 (6.3)	0.669 (0.263)	1.20 (0.64)
reproducibility of contouring on MRI with IV	1.7 (1.2)	0.889 (0.049)	1.13 (0.19)

Table 3: Overview of delineation results using the different imaging modalities. dCOM is the distance between the centers of mass, and Rv is the ratio of the two volumes.

III.3 Correlation movements fiducials and target

III.3.1 Using 4D PET/CT

The GTV was delineated on the 4D PET images while the coordinates of the fiducials were read from the corresponding 4D data. Again target delineation on PET was evaluated. Contours of two targets, of 70 cc and 3 cc, were drawn ten times by the same operator. The results regarding the determination of the delineation reproducibility were summarized in Table 4.

	Relative uncertainty of target volume	Dice coefficient	Absolute uncertainty of distance between the different target volumes contoured (mm)
target #4 mean volume = 70 cc	1.8%	0.99 ± 0.01	0.4
target #5 mean volume = 3 cc	2.3%	0.96 ± 0.01	0.2

Table 4 : Relative uncertainty on absolute volume definition, Dice value and absolute uncertainty on distance between target COMs.

As target delineation on 4D PET is reproducible the movements of fiducials (4D CT) and GTV (4D PET) can be compared on the different breathing phases. Results are shown in figure 6. The movement of fiducials and GTV is clearly not identical. The fiducials seems to be moving with much larger amplitude. At first sight, this might indicate deformations in the liver and the fact that real-time tracking using the center of mass of the fiducials will not lead to a correct tracking of the GTV. A closer inspection of the 4D CT/PET images revealed a non consistency of the two image modalities as images are not obtained simultaneously. The patient is first scanned on the CT in a very short time interval (less than a minute), thus taking a snapshot of the respiration movements. Afterwards the PET images are required, averaging the movements of a large number of breathing cycles (20 minutes). This averaging systematically leads to smaller motion amplitudes.

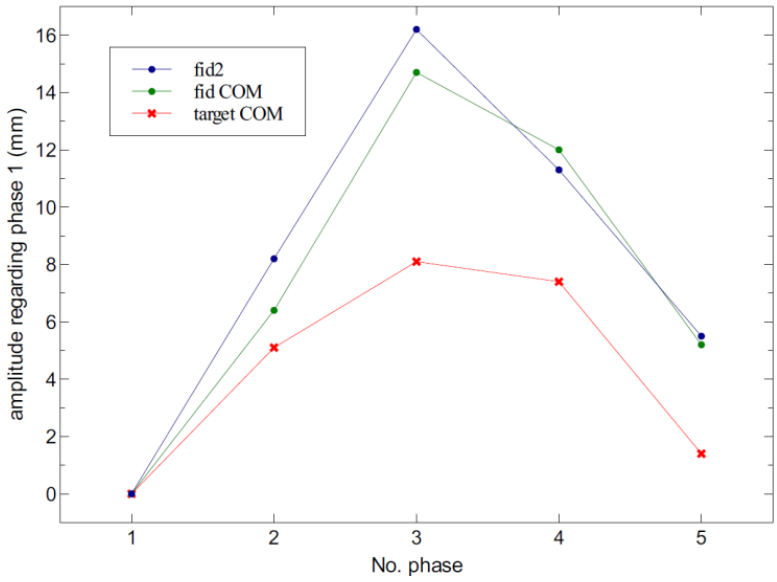


Figure 6 : Example of the motion amplitudes of target and fiducials.

For the current study, 4D CT/PET does not provide enough precision and an alternative technique based on image registration is studied.

III.3.2 Using deformable image registration

Current method is based on the 4D CT images only. Fiducial coordinates are read from the CT images corresponding to the different breathing phases. As demonstrated in paragraph III.1 the GTV is not

always clearly visible on the CT images. To address this problem we copied the PET GTV on the reference CT image (e.g. phase 1) and used deformable image registration (ITK/Elastix) to propagate the GTV contours to the other breathing phases. This provided GTV contours and fiducial positions on the same 4D CT images.

III.3.2.1 Evaluation of deformable registration

To ensure an accurate propagation of the GTV contours from the reference phase the quality of the deformation vector fields (DVFs) was determined by applying the deformation on the fiducials and comparing their "deformed position" to their actual positions. Results are shown in figure 7, illustrating a 1 mm precision of the deformation field for this intra-modality deformable registration. During the determination of the DVFs the fiducials were erased from the images to avoid any bias. Furthermore we compared the dice values of the actual and deformed kidney contour as this structure is clearly visible on CT and allows accurate delineation on all breathing phases. Dice coefficients > 0.9 were systematically obtained.

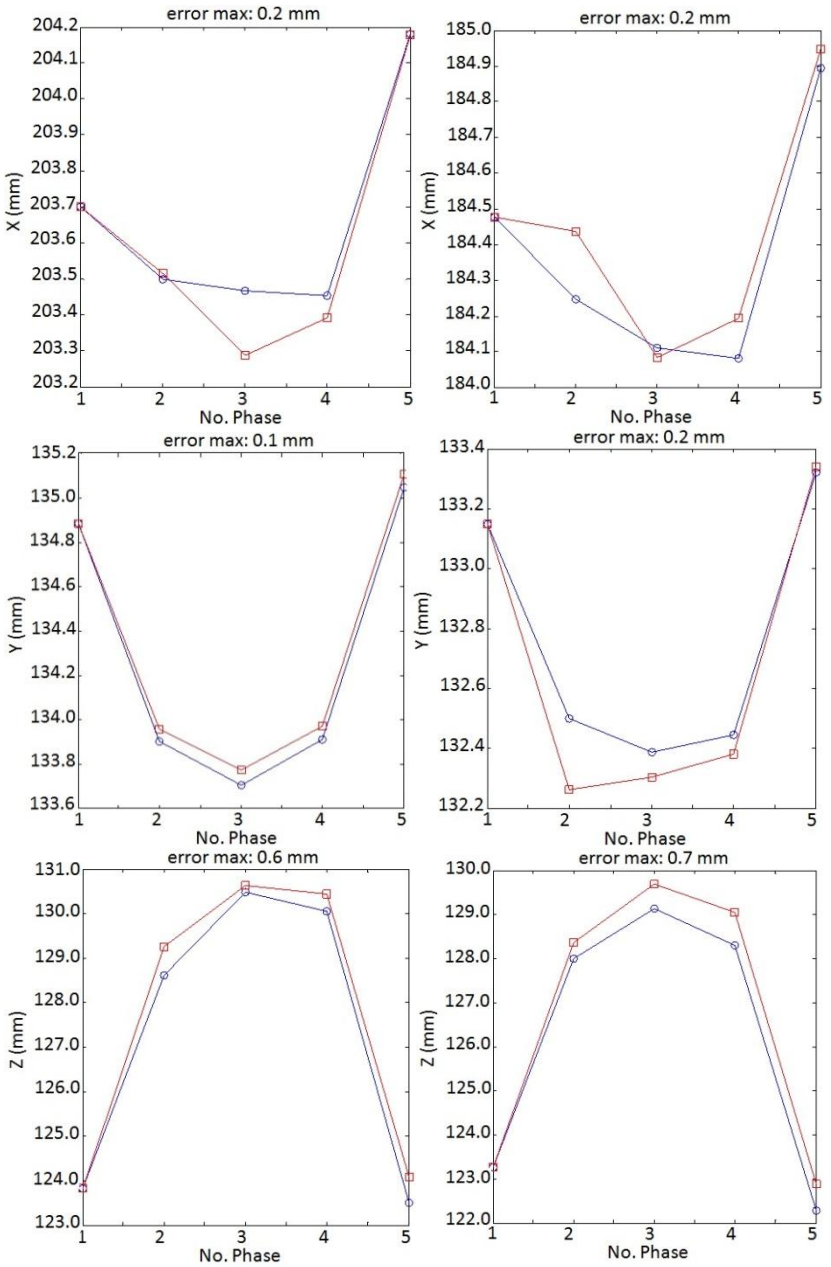


Figure 7 : Original fiducial coordinates of all phases (squares) and “deformed” fiducial coordinates obtained after DVFs (circles) (results for two fiducials as an example). A 1 mm precision is obtained, proving the accuracy of propagating the GTV contour.

III.3.2.2 Correlation evaluation

As the accuracy of the deformable image registration method has been demonstrated, the actual correlation between the movements of the target and the center of mass of the fiducials can be determined. The blue diamonds in figure 8 provide the dice coefficient of the GTV on phase i and that of the reference phase (phase 1). This simulated a Cyberknife treatment without Synchrony tracking. The yellow circles are obtained by shifting the GTV contour from the reference phase to breathing phase I using the translation determined by the shift of the center of mass of the fiducials on the corresponding phase. This simulated a Cyberknife treatment using Synchrony tracking (the Synchrony tracking ignores rotations and deformations).

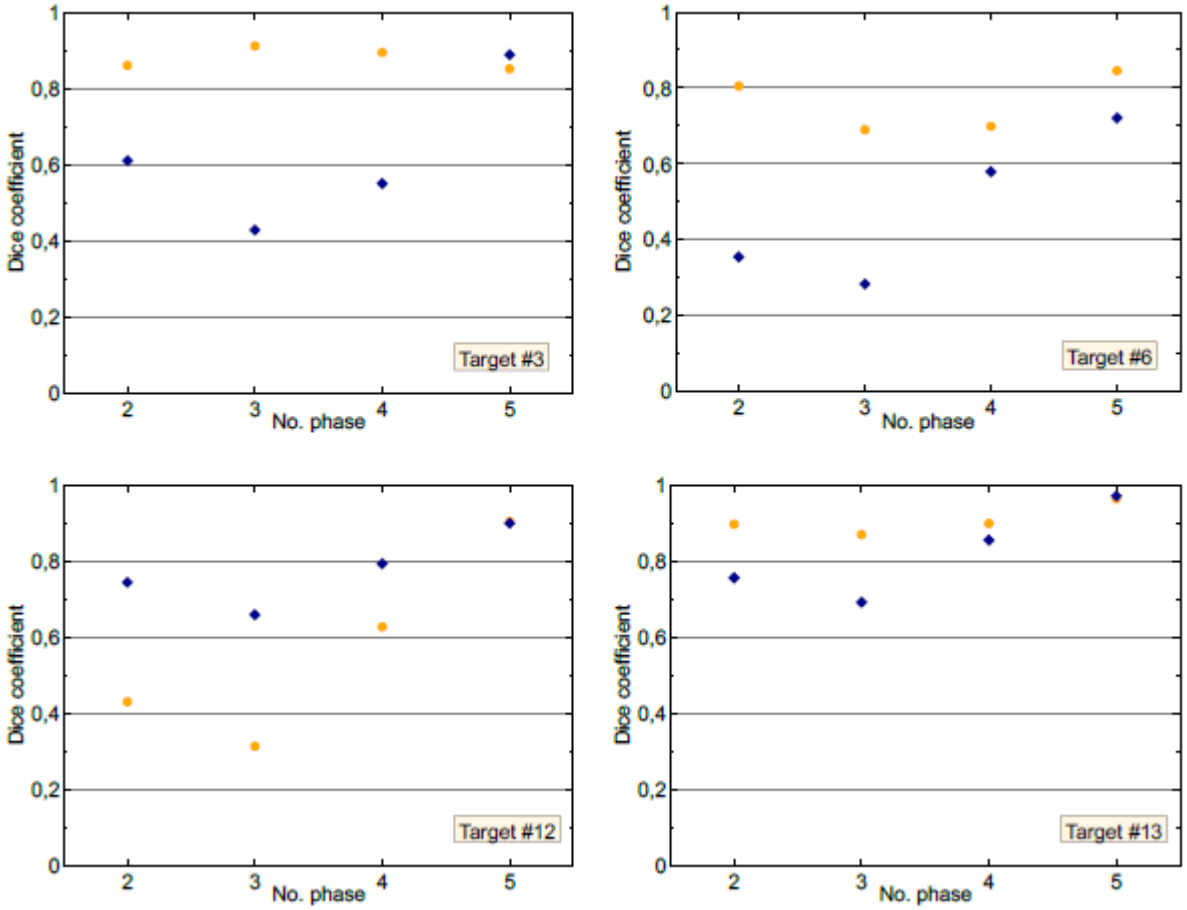


Figure 8 : Dice coefficients between the targets of each phase and of the reference phase (Ph1), for three patients. Targets 12 and 13 correspond to the same patient. Target 12 is not surrounded optimally by the fiducials. Blue diamonds refer to the results without registration, orange circles refer to the registration with fiducials COM translation.

For most patients, Synchrony tracking systematically improves the dice coefficient between the GTV contours, demonstrating the added value of real-time tracking for these patients. For target 12 though, the opposite result is obtained. Synchrony tracking decreases the dice value on all phases. This is because of a non-optimal positioning of the fiducials. This patient was originally treated on Target 13 and fiducials were positioned around this target. Afterwards a second lesion was detected on MRI. Our method demonstrated that additional fiducials needed to be introduced for this second target. When 4D CT data are available for these liver patients, our method can systematically be used to determine the added value of Synchrony tracking.

The impact of taking into account rotations of the fiducials was studied in figure 9. Introducing rotations in the Synchrony tracking system does not increase the dice coefficients. And as the dice values obtained, when only considering the translations, are close to 1 for Target #2 (adequate fiducial positioning), even deformations can be ignored. In the case of Target #12 where fiducial placing was not optimal, results show that adding rotations would not improve target tracking.

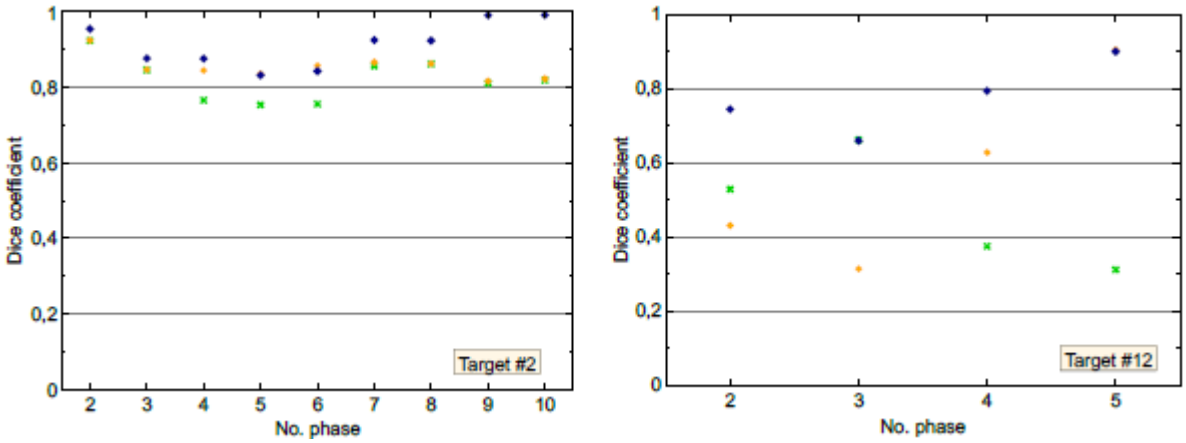


Figure 9 : Dice coefficients between the targets of each phase and of the reference phase (Ph1), for targets #2 and #12. Blue diamonds refer to the results without registration, orange circles refer to the registration with fiducials COM translation and green crosses refer to the registration with translation and rotation of the fiducials

III.4 Monte Carlo dose calculations evaluating the impact of uncertainties

III.4.1 Impact of delineation uncertainties

To evaluate the impact of delineation uncertainties, two GTV positions were selected based on the two delineation sessions and dose was calculated, using the same plan, optimized on one of the positions. This test was performed with a 1 cm GTV to CTV margin, which was the margin used in clinical routine. The dose distribution was calculated using our in-house Monte Carlo platform Moderato. The results are shown in figure 10.

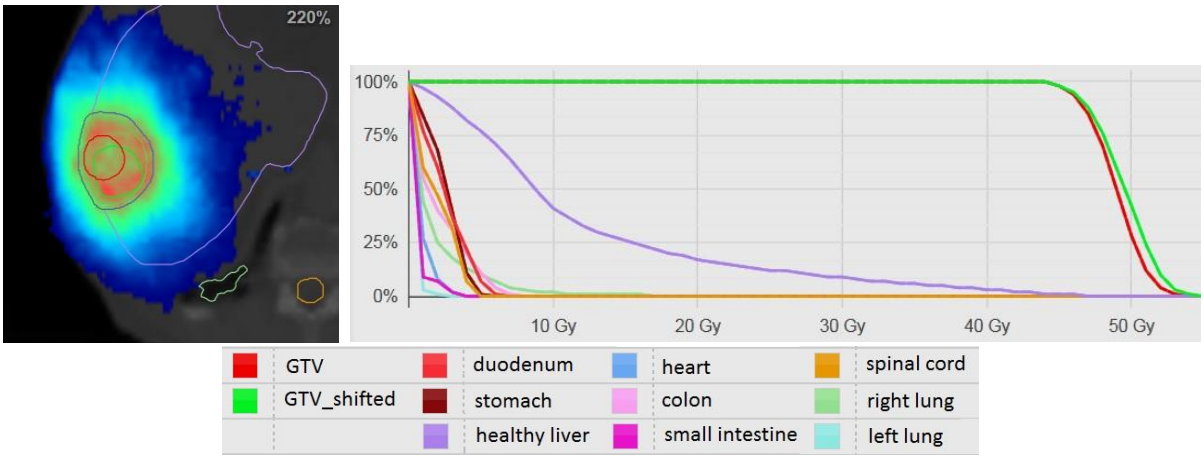


Figure 10 : DVHs for 3D distribution calculated by Moderato, targeting the PTV.

III.4.2 Dosimetric effect of tracking motion on OARs

The Synchrony tracking only takes into account the motion of the fiducials and thus that of the GTV. As the treatment plan is not optimized in 4D, deformations can lead to an over-irradiation of organs at risk. This was studied in detail in figure 11, comparing the 4D dose distribution with a static dose distribution as optimized on the 3D planning CT. The GTV to CTV margin was delineated as a specific structure allowing distinguishing the differences in GTV and CTV. The impact on the organs at risk is rather limited for the cases studied in the current work.

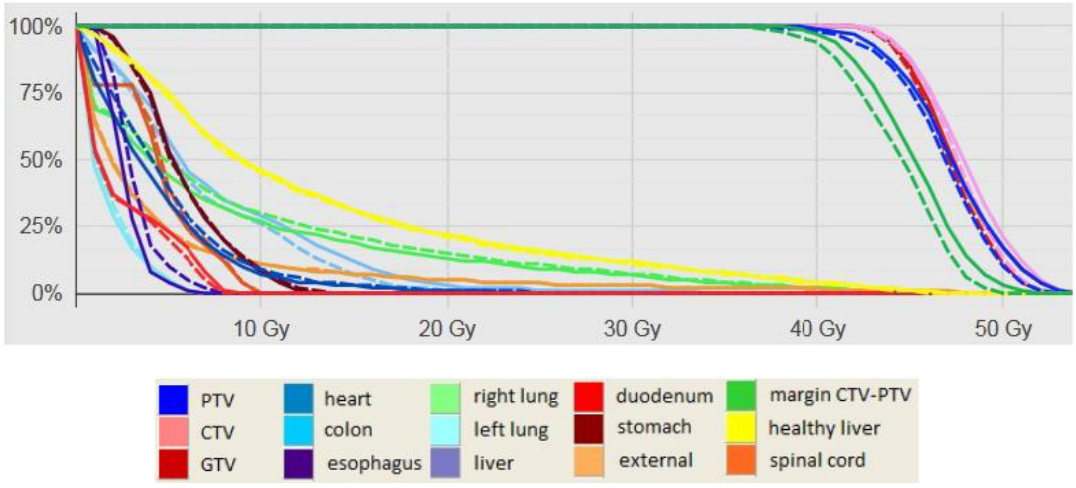


Figure 11 : DVHs for 3D dose distribution calculated by Moderato on the planning CT images (solid line) and 4D dose distribution calculated by Moderato with fiducial tracking (dashed line), for one patient.

III.4.3 Accuracy of fiducial tracking

The accuracy of fiducial tracking can be compared by calculating the 4D dose distribution that would have been obtained by tracking the center of mass of the actual GTV (target tracking). Only one example is shown in figure 12, but comparable results are obtained for all patients. The dosimetric impact of tracking the fiducials instead of the target directly is negligible.

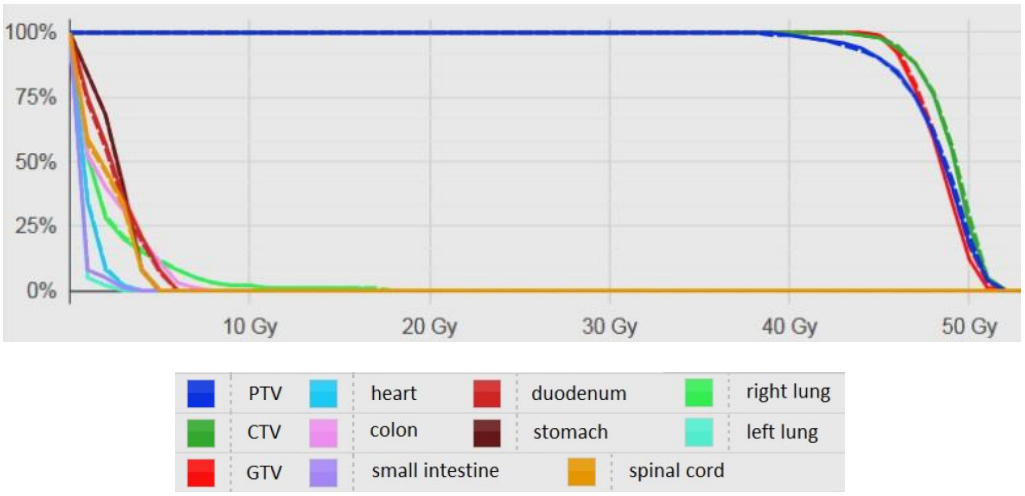
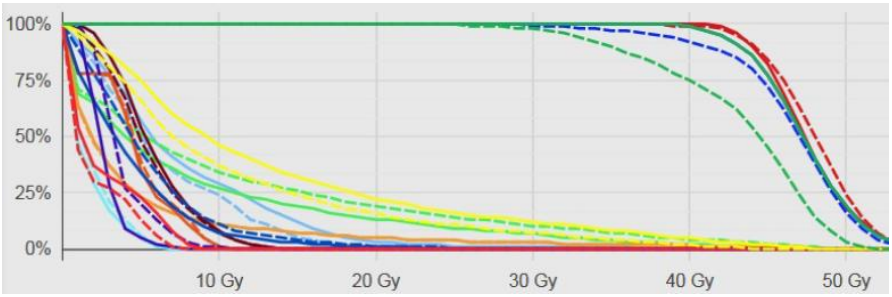


Figure 12 : DVHs for 4D dose distribution calculated by Moderato with target tracking (solid line) and 4D dose distribution calculated by Moderato with fiducials tracking (dashed line), for patient 3

III.4.4 Utility of respiratory tracking

Moderato can calculate dose without using the actual robot positions available in the log files. This simulates a treatment without tracking and allows determining its added value. Results are shown in figure 13. Although the impact of tracking might be important when considering a single phase (especially in the GTV to CTV margin), after summation over the complete breathing cycle, the impact becomes almost negligible. Two additional patients, illustrating a slightly larger impact, are shown in figure 14. Remark that, although the PTV-DVH is actually shown, only the CTV (and GTV) DVHs should be considered for this 4D comparison.



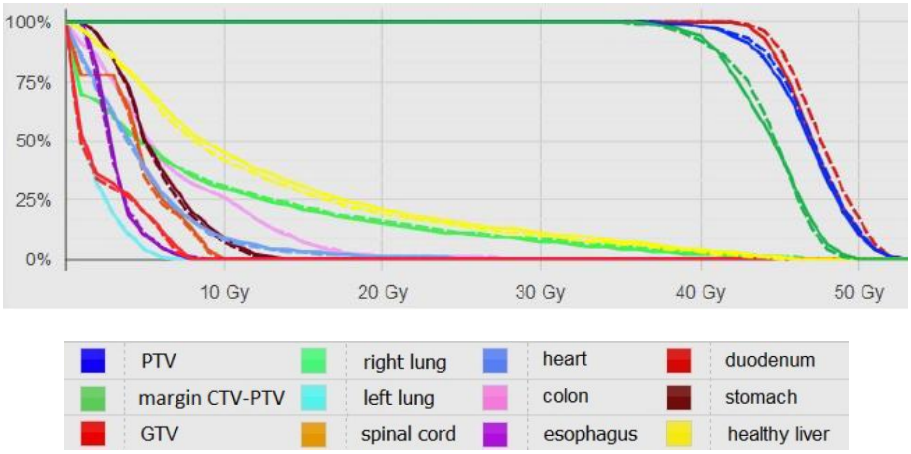


Figure 13 : DVHs comparing an individual phase with the planning CT (up) and the cumulated 4D dose distribution calculated by Moderato with (solid line) and without fiducial tracking (dashed line)

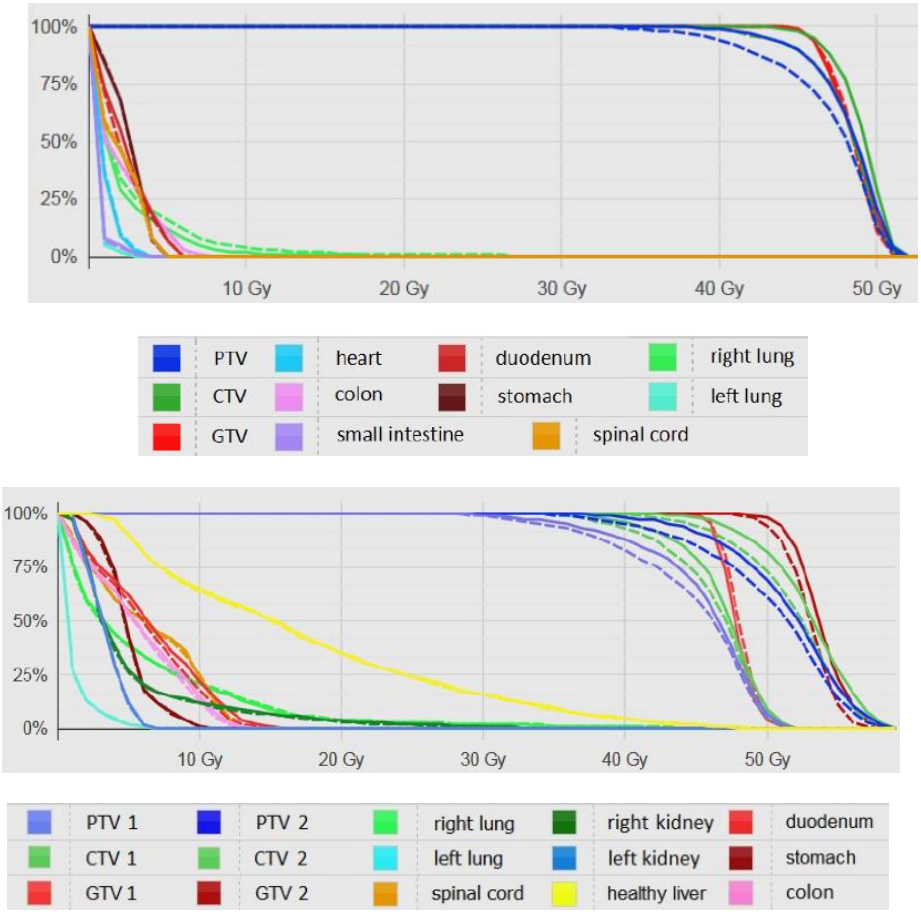


Figure 14 : DVHs for 4D dose distribution calculated by Moderato with fiducials tracking (solid line) and 4D dose distribution calculated by Moderato without tracking (dashed line), for two patients where the impact was the most important in the current study.

III.5 Impact of CTV margin

The results discussed above, seem to illustrate that the dosimetric impact of using Synchrony tracking is rather limited. This is partly thanks to the relatively large GTV to CTV margin applied in our clinic. As this is not generally the case in all departments we evaluated the impact of this margin on the influence of all considered uncertainties.

III.5.1 Impact of CTV margin on delineation uncertainty

When using a CTV margin, which is the current clinical procedure in our department, the impact of the uncertainty on the actual GTV contour is limited, as shown in figure 10. This is no longer the case when no CTV margin is applied, which is often the case in other departments, as shown in figure 15.

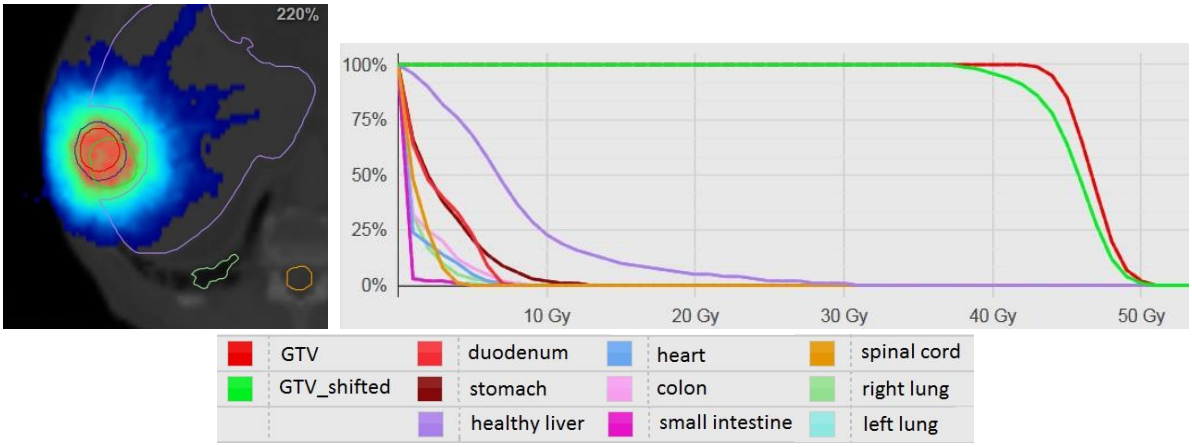


Figure 15: Dosimetric impact of delineation uncertainty when no GTV to CTV margin is applied. The GTV coverage becomes sub-optimal.

III.5.2 Impact of CTV margin on tracking

As illustrated in figure 16, even when no CTV margin is applied the accuracy of fiducial tracking is still adequate. Actual target tracking does not further improve target coverage.

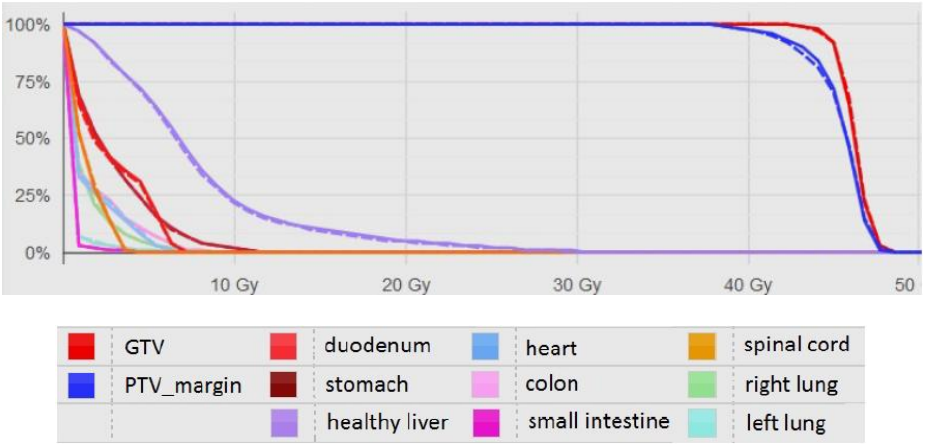


Figure 16 : DVHs for 4D dose distribution calculated by Moderato with target tracking (solid line) and 4D dose distribution calculated by Moderato with fiducials tracking (dashed line), for a patient without CTV margin.

As shown in figure 17, the conclusion regarding the utility of target tracking changes dramatically when removing the CTV margin. The GTV coverage is clearly improved by performing real-time tracking.

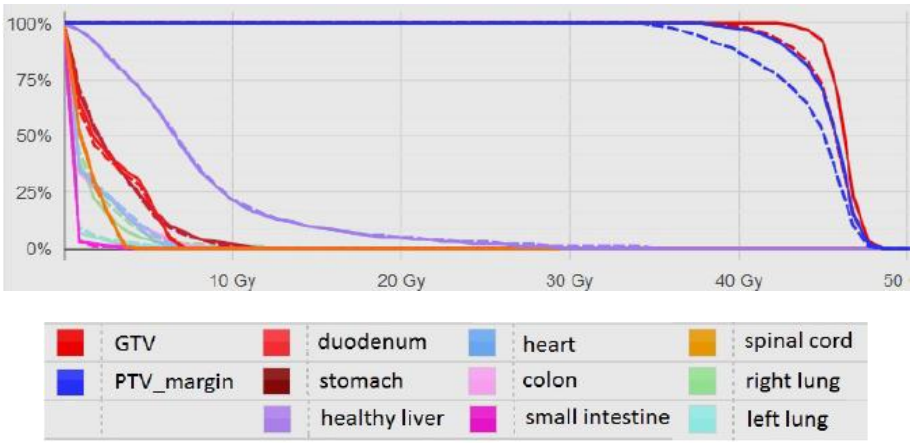


Figure 17 : DVHs for 4D dose distribution calculated by Moderato with fiducials tracking (solid line) and 4D dose distribution calculated by Moderato without tracking (dashed line), for a patient without CTV margin.

IV. Discussion

The main question when initiating the research described in the current paper, was the added-value of Synchrony tracking when treating liver patients using Cyberknife. As, in our department, a belt is systematically used to limit liver motion related to the respiration of the patient, we did not observe large motions and this for all studied cases. Tracking the tumor without using 4D plan optimization does not seem a robust treatment with respect to the organs at risk. In case of deformations, the actual beam can be oriented towards organs at risk while tracking the tumor. So the potential gain in target coverage should be compared to the increased risk of over irradiation of the organs at risk. Especially for liver treatments, where the lesion is close to the duodenum, the stomach or the small intestine, care should be taken. In current study, we did not find any patient where this deformation effect led to an important increase of the dose in these organs.

On the other hand the added value regarding target coverage was rather limited as well. But this conclusion is largely influenced by the GTV to CTV margin. In our clinic, a large margin is systematically applied (5 mm). Per definition the CTV margin should only be applied to include microscopic disease. In clinical practice though, the CTV margin often compensated for uncertainties in GTV delineation in general. The impact of tracking is removed by the large margin. When removing this margin, as is clinical practice in several other departments, the added value of tracking the tumor becomes much more important.

We also wanted to put the relevance of tracking into perspective to other uncertainties in the treatment chain. For these specific patients, the GTV delineation step has proven to be the most important contribution to this uncertainty. For a large number of patients the delineation is solely based on CT (with or without contrast). As demonstrated in the current paper, delineation of liver tumors on CT is a very difficult task. Intra and inter-operator reproducibility is limited as shown by the obtained dice coefficients and the large distances between the centers of mass. Delineation on MRI certainly improved the reproducibility although this demands a profound knowledge of the anatomy of the liver and of the specific MRI sequences used in clinical practice. This also requires a close collaboration between radiotherapist and radiologists and a standardization of MRI sequences in terms of contrast and geometrical precision. In practice, several sequences should be combined, although this complicates the procedure even further. Often CT and MRI are combined (visually or by applying rigid or deformable image registration). But again, the liver is a complicated region and the patient needs to be scanned on MRI and CT in comparable conditions. Even when considering all above, the final uncertainty on the GTV delineation is more than 2 mm. Again the usage of a large CTV margin will remove the dosimetric impact of this uncertainty. MRI-only radiotherapy might provide a solution for this problem, as the registration uncertainty is completely removed, leading to an increased geometrical precision (maybe at the price of a decreased dose precision, due to the generation of pseudo-CT).

Another important result of the current work, is the fact that fiducial tracking, only taking into account the translation of the center of mass of the fiducials seems to be adequate. Taking into account rotations and deformations does not improve target coverage, and this result does not depend on the CTV margin. It might be related to the systematic usage of the belt though.

Although the 4D PET/CT data were not really usable for our study, the results were included anyway, as the clinical impact can be very large. The fact that a 4D PET and 4D CT information is not consistent is not only relevant for the combination of the two imaging modalities, but has also very general consequences. It provides again an illustration that 4D CT is probably not the optimal tool for treatment planning as it only provides a snapshot of the motion related to breathing, which is not necessarily representative for the motion during the actual treatment delivery. As discussed by other authors (PSI e.g. refs), 4D MRI might provide an increased accuracy. 4D MRI allows averaging over an important number of breathing cycles, but also evaluating the impact of variations in the breathing motion.

When the CTV margin is only considering microscopic disease, the comparisons should be focusing on CTV coverage (instead of that of the GTV). This is not clinical reality though and was thus not considered in the current paper.

V. Conclusion

In the current work it has been shown that real-time tracking for liver tumors is only relevant when the GTV to CTV margin is relatively small and when care is taken regarding the precision of GTV delineation (combining CT and MRI information). Rotations and deformations have no impact on the precision of tracking and the impact on the organs at risk is negligible. It is possible that the results are largely influenced by the usage of a belt that limits the liver movements during treatment delivery. The accuracy of GTV delineation of liver lesions should be further improved, and this is not possible without MRI. This demands a close collaboration between the radiology and radiotherapy departments, and an optimization of the MRI sequences used. MRI-only radiotherapy might provide an important step towards a more accurate GTV definition. The utility of 4D CT in general should be questioned as well.

Acknowledgements

The first author of the current paper has worked on a PhD thesis sponsored by Accuray.

References

[1] J. M. Balter, R. K. Ten Haken, T. S. Lawrence, K. L. Lam, and J. M. Robertson, “Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing,” Int. J. Radiat. Oncol. Biol. Phys., vol. 36, no. 1, pp. 167–174, 1996.

[2] P.-Y. Y. Bondiau, K. Bénézery, V. Beckendorf, D. Peiffert, J.-P. P. Gérard, X. Mirabel, A. Noël, V. Marchesi, T. Lacornerie, F. Dubus, T. Sarrazin, J. Herault, S. Marcié, G. Angellier, and E. Lartigau, “[CyberKnife robotic stereotactic radiotherapy: technical aspects and medical indications],” Cancer Radiother., vol. 11, no. 6–7, p. 338—344, 2007.

[3] E. Lartigau, X. Mirabel, B. Prevost, T. Lacornerie, F. Dubus, and T. Sarrazin, “Extracranial stereotactic radiotherapy: Preliminary results with the cyberknife??,” Onkologie, vol. 32, no. 4. pp. 209–215, 2009.

[4] A. S. Beddar, K. Kainz, T. M. Briere, Y. Tsunashima, T. Pan, K. Prado, R. Mohan, M. Gillin, and S. Krishnan, “Correlation between internal fiducial tumor motion and external marker motion for liver tumors imaged with 4D-CT,” Int. J. Radiat. Oncol. Biol. Phys., vol. 67, no. 2, pp. 630–638, 2007.

[5] L. A. Dawson, K. K. Brock, S. Kazanjian, D. Fitch, C. J. McGinn, T. S. Lawrence, R. K. Ten Haken, and J. Balter, “The reproducibility of organ position using active breathing control (ABC) during liver radiotherapy,” Int. J. Radiat. Oncol. Biol. Phys., vol. 51, no. 5, pp. 1410–1421, 2001.

[6] H. Shirato, Y. Seppenwoolde, K. Kitamura, R. Onimura, and S. Shimizu, “Intrafractional Tumor Motion: Lung and Liver,” Seminars in Radiation Oncology, vol. 14, no. 1. pp. 10–18, 2004.

[7] E. D. Brandner, A. Wu, H. Chen, D. Heron, S. Kalnicki, K. Komanduri, K. Gerszten, S. Burton, I. Ahmed, and Z. Shou, “Abdominal organ motion measured using 4D CT,” Int. J. Radiat. Oncol. Biol. Phys., vol. 65, no. 2, pp. 554–560, 2006.

[8] D. P. Gierga, J. Brewer, G. C. Sharp, M. Betke, C. G. Willett, and G. T. Y. Chen, “The correlation between internal and external markers for abdominal tumors: Implications for respiratory gating,” Int. J. Radiat. Oncol. Biol. Phys., vol. 61, no. 5, pp. 1551–1558, 2005.

[9] E. W. Pepin, H. Wu, Y. Zhang, and B. Lord, “Correlation and prediction uncertainties in the CyberKnife Synchrony respiratory tracking system,” Med. Phys., vol. 38, no. 7, p. 4036, 2011.

[10] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, “The ITK Software Guide,” ITK Softw. Guid., vol. Second, no. May, p. 804, 2005.

[11] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, “Elastix: A toolbox for intensity-based medical image registration,” IEEE Trans. Med. Imaging, vol. 29, no. 1, pp. 196–205, 2010.

[12] ICRU, Prescribing, Recording and Reporting Photon-Beam Intensity Modulated Radiation Therapy (IMRT) (ICRU Report 83), vol. 10, no. 2. 2010.

[13] N. Reynaert, B. Demol, M. Charoy, S. Bouchoucha, F. Crop, A. Wagner, T. Lacornerie, F. Dubus, E. Rault, P. Comte, R. Cayez, C. Boydev, D. Pasquier, X. Mirabel, E. Lartigau, T. Sarrazin, « Clinical implementation of a Monte Carlo based treatment plan QA platform for validation of Cyberknife and Tomotherapy treatments, », *Physica Medica*, In Press, http://dx.doi.org/10.1016/j.ejmp.2016.09.009

4. La start-up autour de Moderato

Le but de la start-up est d'offrir un service de double calcul Monte Carlo à distance. La collaboration avec un client commence avec la modélisation de son Linac (un ou plusieurs Linacs) et l'étalonnage stœchiométrique de son scanner CT (pour un ou plusieurs protocoles). Pour un type de Linac déjà implémenté dans Moderato, la modélisation d'un Linac spécifique revient simplement à optimiser l'énergie et la taille du spot des électrons primaires. Cette optimisation peut même se faire automatiquement. Pour l'étalonnage du CT, un fantôme contenant des inserts de différentes densités doit être scanné, et les images analysées afin de déterminer l'étalonnage en termes de densité et de composition des tissus. Une fois que cette phase initiale est réalisée, Moderato est configuré pour la validation des plans de traitement du client. Le client reçoit l'accès au système grâce à une interface web (Javascript), où il charge les fichiers dicom du patient (CT, RTStruct, RTDose, RTPlan) dans le « dépôt » du logiciel (les données sont automatiquement anonymisées). Le calcul est alors lancé sur la batterie de serveurs de calcul localisée dans la start-up (ou un service sur le « cloud »). Une fois que le double calcul est finalisé, le client reçoit un émail qui contient un lien vers les résultats. Le résultat final (la distribution de dose recalculée en Monte Carlo) est disponible dans le GUI de l'interface Web. Le client peut imprimer le rapport de double calcul en pdf et l'importer dans son système de Record&Verify. Le choix de centraliser les serveurs de calculs va faciliter la maintenance du système. De cette façon le client ne doit pas installer de serveurs de calcul dans son hôpital. Il faut, par contre, augmenter dynamiquement le nombre de serveurs en fonction du nombre de clients au sein de la start-up. L'investissement initial (maintenant que le logiciel est déjà développé) est minimal. Il faut prévoir un support informatique : un informaticien toujours disponible pour garantir le fonctionnement du système (le logiciel et les serveurs Linux). Les aspects commerciaux peuvent être gérés par la société Aquilab, qui a déjà son réseau de clientèle en radiothérapie. Un « business plan » est déjà créé avec EuraSanté. La SATT est également en train d'étudier le dossier afin de nous supporter concrètement pendant la création de la start-up. L'obtention du marquage CE reste une étape importante à prévoir.

Chapitre 3 : Radiothérapie en IRM seule

1. Introduction

Aujourd'hui, un scanner (CT-simulateur) du patient est utilisé pour optimiser les traitements de radiothérapie grâce à un système de planification de traitement (TPS). Les niveaux de gris du scanner (unités Hounsfield) sont convertis en densités électroniques, un paramètre indispensable pour le calcul de l'absorption du faisceau et de la distribution de dose. L'imagerie par IRM permet une meilleure différenciation des tissus mous et la délinéation de façon optimale du volume cible et des organes à risque. Actuellement, les données IRM et CT doivent être recalées, ce qui augmente l'incertitude. Même en utilisant des algorithmes récents de recalage déformable, l'incertitude est supérieure à 2 mm en dehors du crâne [48]. Cette incertitude liée au recalage peut être réduite en calculant la dose directement sur l'IRM, ce qui simplifie le processus de planification de traitement et limite la dose associée avec le scan [48][49].

Plusieurs facteurs ont freiné l'utilisation de l'IRM-seule en radiothérapie : impossibilité de positionner les patients dans la position de traitement (avec les systèmes de contention comme le matelas, le masque thermoformé, …), trop de distorsions, et bien sûr le fait que l'information concernant la densité électronique ne soit pas disponible. Entretemps, les fournisseurs IRM ont développé des machines dédiées à la radiothérapie, offrant la possibilité de positionner le patient comme pendant le traitement [50][51]. Les distorsions sont globalement corrigées avec une précision millimétrique [49] et plusieurs méthodes sont décrites dans la littérature scientifique afin de résoudre le problème concernant l'absence des densités électroniques dans les images IRM. Ces méthodes sont surtout développées dans les pays nordiques [52]. En Suède par exemple, il y a plusieurs groupes spécialisés sur ce sujet et ils viennent même de lancer un projet national afin de stimuler l'utilisation de l'IRM en radiothérapie (« Gentle Radiotherapy » : http://www.radiotherapytestbed.com/project/mri-based-radiotherapy/).

La France n'est pas à la pointe dans ce domaine très prometteur et l'obtention du label PhysiCancer a été un signal très encourageant (Aviesan, INCA 2015). Le projet, coordonné par le Centre Oscar Lambret est fait en collaboration avec l'ICO de Nantes et le LTSI de Rennes [53][54].

Une première étude sur ce sujet a déjà été entreprise au Centre Oscar Lambret en 2006, dans le cadre du cancer de la prostate en radiothérapie conformationnelle. Pour chaque organe contouré, une densité constante était définie (« bulk densities » [55]). Les résultats ont montré la faisabilité de cette méthode pour la radiothérapie conformationnelle sans modulation d'intensité. Par contre, cette approximation a des conséquences importantes, si on considère des traitements avancés comme la stéréotaxie (CyberKnife) ou la RCMI (Tomothérapie, RapidArc) [49]. Ces deux modalités utilisent de petits faisceaux « beamlets » qui traversent des géométries hétérogènes (par exemple les poumons, tête et cou). L'évaluation de la distribution de dose demande des algorithmes de calcul très précis et une description détaillée du patient, concernant les densités et la composition de ses tissus (tumeur et organes à risque). Il est donc nécessaire de chercher des séquences spécifiques qui sont appropriées pour déterminer la composition et la densité des tissus (sans distorsions). On peut même envisager de combiner plusieurs séquences afin d'en extraire les informations requises pour convertir l'IRM en TDM virtuel (pseudo-CT) voxel par voxel pour le calcul précis des doses.

Le problème principal est le faible signal de l'os cortical en raison de sa faible teneur en hydrogène. Un certain nombre de groupes ont introduit des séquences de double écho ultracourt (DUTE) pour distinguer l'os de l'air [56][57][58]. Ces méthodes fournissent des cartes d'atténuation excellentes pour le système TEP / IRM, mais l'utilisation de trois densités/matériaux (air, tissus, os) n'est pas suffisante pour la planification de traitement. Elle conduit à des déviations pouvant aller jusqu'à 9% dans les organes à risque [59], ce qui n'est pas souhaitable pour les techniques les plus évoluées. L'association de ces séquences DUTE avec des images de densité de spin devrait permettre d'affiner la définition des tissus (composition et densité). Johansson [57] combine les deux solutions, à l'aide des images T_2 et de deux séquences DUTE utilisant un algorithme pour convertir les informations IRM en TDM virtuel. Leur méthode a été appliquée avec succès au niveau cérébral, mais des artefacts importants sont obtenus pour le bassin et le thorax (grand champ de vue). L'incertitude sur les unités Hounsfield calculées dans les images pseudo-CT a été mesurée, permettant d'estimer l'erreur sur le calcul de dose [52]. Pour le moment les séquences UTE ne sont appliquées que dans le cerveau avec des développements en cours (Nyholm 2014) [48]. Il faut encore recaler les images UTE avec d'autres images IRM (par exemple T_2), sur lesquelles on peut faire la délinéation des organes, parce que les veines sont noires sur des images UTE. Ce recalage intra-modalité est moins problématique au niveau des incertitudes. Une autre méthode consiste à utiliser un atlas avec des images de référence (CT et IRM pour un certain nombre de patients) ; l'IRM du nouveau patient est déformée sur l'IRM de référence, et les vecteurs de déformations sont utilisés pour déformer le CT de référence afin d'obtenir le pseudo-CT du patient [60].

Une alternative originale, étudiée au Centre Oscar Lambret, est de relier (d'abord empiriquement puis plus théoriquement) la densité électronique au contenu en hydrogène. Le contenu en hydrogène peut alors être déterminé en utilisant l'IRM. Cela demande une reprogrammation de la physique des calculs de dose dans la plate-forme Monte Carlo utilisée [17][18].

En général l'IRM peut jouer un rôle important dans chaque étape de la chaîne de traitement en radiothérapie :

- la délinéation des organes (meilleur contraste des tissus mous en IRM)
- calcul de dose directement sur les images IRM
- positionnement pendant chaque séance de traitement
- IRM 4D pendant chaque séance : reconstruction de la dose délivrée
- suivi de l'évolution de la tumeur entre les séances de traitement
- suivi (efficacité/toxicité) après le traitement
- IRM-Omics

Dans ce chapitre, le focus est mis sur les calculs de dose sur les images pseudo-CT (pCT) générées à partir de l'IRM pour les traitements du crâne, de la tête et du cou). Les articles introduits dans ce chapitre ont été écrits dans le cadre d'une bourse Cifre entre Aquilab, l'IEMN de Lille1 et le Centre Oscar Lambret ; et dans le cadre d'un projet PhysiCancer, démarré en septembre 2015 (INCa). Ces projets ont deux volets :

- Un volet pratique ou clinique: en utilisant des méthodes de génération de pseudo-CT existantes, on peut vite tester et appliquer des images pCT en routine clinique pour plusieurs indications. Même si on sait que la solution ne sera pas parfaite.
- Un volet fondamental : en collaboration le LTSI de Rennes, on va chercher et optimiser des séquences IRM (DUTE, ZTE, …) afin de pouvoir convertir directement les images IRM en pCT sans passer par un atlas.

Le deuxième volet semble plus intéressant, mais on ne peut pas attendre cette solution, sinon le projet serait fini avant qu'on ait testé le premier pCT. De plus, il est possible que la solution parfaite (la séquence parfaite) n'existe pas. C'est pourquoi on a choisi de combiner les deux volets.

2. Article 3 : L'importance de l'hydrogène dans les calculs de dose

2.1 Introduction

Dans le premier article (article 3), nous avons démontré que seul le contenu hydrogène des tissus est nécessaire pour obtenir une précision de 1 à 2 % dans les calculs de dose Monte Carlo. Cette étude fait partie du volet fondamental. Cette hypothèse est basée sur la figure 20 où on voit que l'hydrogène domine tous les facteurs d'interactions comme les pouvoirs d'arrêts et les facteurs d'atténuation.

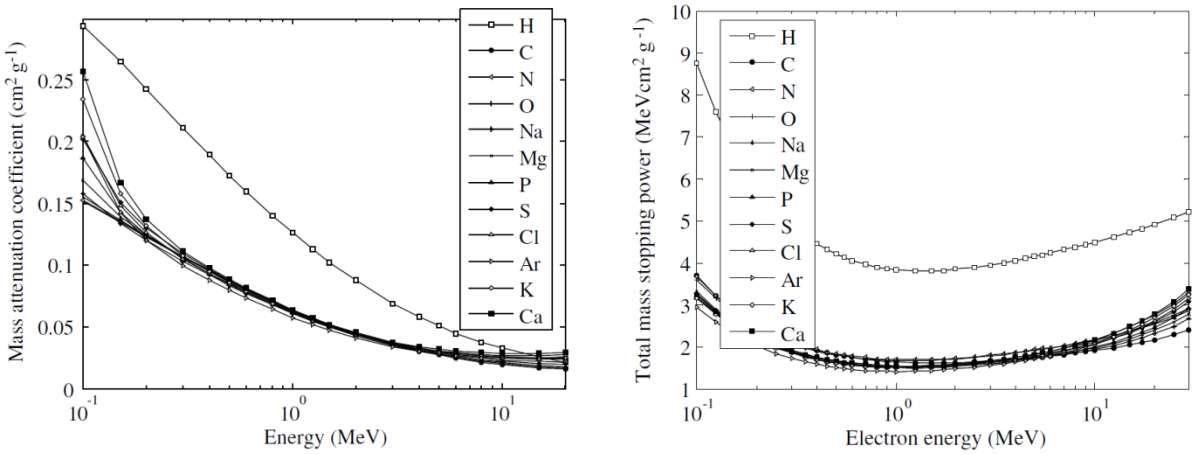


Figure 20 : Pouvoir d'arrêts et coefficients d'atténuation pour les différents éléments. On voit clairement que dans la gamme d'énergie correspondant à la radiothérapie moderne (des énergies photons dans le patient entre 0.5 et 3 MeV et des électrons d'une énergie moyenne de 0.3 MeV) les facteurs sont plus importants pour l'hydrogène et que pour les autres éléments les courbes se superposent.

Dans cette étude, on part d'un étalonnage stœchiométrie du CT comme référence [61][17]. La composition des voxels est ensuite simplifiée : tous les éléments, sauf l'hydrogène, sont remplacés par l'oxygène, qui est considéré comme un élément neutre. En utilisant Moderato, nous avons démontré que cette simplification ne modifie pas du tout la distribution de dose. Par contre, dès le moment où on modifie légèrement le contenu hydrogène des tissus, la dose change. Une fois démontré qu'il ne faut que l'hydrogène, il faut encore trouver une séquence qui offre cette information. En théorie, ça semble possible. On devrait être capable de sortir le contenu hydrogène avec l'IRM car le signal (cf. premier chapitre) dépend de la densité de protons. Par contre, le

contraste de l'image IRM est largement influencé par les effets de relaxation (T_1 et T_2). Il faut alors chercher des séquences qui donnent des images qui ne sont pas ou peu pondérées par ces effets, par conséquent des séquences avec un temps d'écho le plus court possible. Le signal de l'os est faible, pas parce qu'il n'y pas de signal du tout, mais plutôt parce-que la relaxation se fait tellement vite que le signal est déjà perdu avant qu'on commence à le collecter (on attend l'écho). Dans l'air par contre, la raison de la faible intensité du signal est due au fait qu'il n'y pas de signal du tout. C'est pour cette raison que les séquences DUTE offrent la possibilité de distinguer l'air de l'os. Une autre alternative est de ne pas attendre l'écho du tout et de collecter le signal tout de suite après l'excitation (séquence ZTE). Ces deux séquences sont installées sur l'IRM préclinique du LTSI de Rennes et peuvent être testées et comparées. Entretemps, on a commencé à installer des séquences spécifiques dans les services de radiothérapie de Rennes (IRM de GE avec séquence Silenz™ : ZTE) et de Nantes (IRM de Siemens avec séquence PETRA™).

2.2 Article 3

Monte Carlo calculation based on hydrogen composition of the tissue for MV photon radiotherapy

Benjamin Demol [1,2,3], Romain Viard [2], Nick Reynaert [1]
[1] *Department of radiotherapy, Centre Oscar Lambret, Lille, France*
[2] *AQUILAB SAS, Loos Les Lille, France*
[3] *IEMN, UMR CNRS 8520, Villeneuve d'Ascq, France*

Journal of Applied Clinical Medical Physics, 2015; 16(5); 117-130

Abstract. The purpose of this study was to demonstrate that Monte Carlo treatment planning systems require tissue characterization (density and composition) as a function of CT number. A discrete set of tissue classes with a specific composition is introduced. In the current work we demonstrate that, for megavoltage photon radiotherapy, only the hydrogen content of the different tissues is of interest. This conclusion might have an impact on MRI-based dose calculations and on MVCT calibration using tissue substitutes. A stoichiometric calibration was performed, grouping tissues with similar atomic composition into 15 dosimetrically equivalent subsets. To demonstrate the importance of hydrogen, a new scheme was derived, with correct hydrogen content, complemented by oxygen (all elements differing from hydrogen are replaced by oxygen). Mass attenuation coefficients and mass stopping powers for this scheme were calculated and compared to the original scheme. Twenty-five CyberKnife treatment plans were recalculated by an in-house developed Monte Carlo system using tissue density and hydrogen content derived from the CT images. The results were compared to Monte Carlo simulations using the original stoichiometric calibration. Between 300 keV and 3 MeV, the relative difference of mass attenuation coefficients is under 1% within all subsets. Between 10 keV and 20 MeV, the relative difference of mass stopping powers goes up to 5% in hard bone and remains below 2% for all other tissue subsets. Dose-volume histograms (DVHs) of the treatment plans present no visual difference between the two schemes. Relative differences of dose indexes D_{98}, D_{95}, D_{50}, D_{05}, D_{02}, and D_{mean} were analyzed and a distribution centered around zero and of standard deviation below 2% (3 σ) was established. On the other hand, once the hydrogen content is slightly modified, important dose differences are obtained. Monte Carlo dose planning in the field of megavoltage photon radiotherapy is fully achievable using only hydrogen content of tissues, a conclusion that might impact MRI dose calculation, but can also help selecting the optimal tissue substitutes when calibrating MVCT devices.

Keywords. Monte Carlo, hydrogen content, stoichiometric calibration, mega-voltage photon radiotherapy, Magnetic Resonance Imaging

I. Introduction

In Monte Carlo treatment planning systems, tissue composition represented by density and elemental composition, is required as a function of Computed Tomography (CT) numbers. Some current common practices to establish tissue characterization can introduce systematic errors in dose planning. Although density is often interpolated in a continuous way, the CT number scale is always divided into a finite number of subsets to link CT number to chemical elemental composition[1,2]. Usually, six or less media of average composition are defined, e.g. air, lung, fat, water, muscle and bone. In this way, some media are averaged on a large range of CT numbers (tissue compositions), e.g. between soft bone and high density cortical bone the composition changes dramatically[3].It has been shown by Verhaegen and Devic that large dose errors (up to 10%) for MV photon beams can occur from inaccurate assignment of media[4]. Moreover tissue equivalent substitutes are often used to calibrate CT scanners just by matching their density to the density of real tissues. As their elemental composition diverges from that of real tissues, calibration curves of CT number to density can differ from reality.

In their work, Vanderstraeten *et al*[5] intended to deal with these two pitfalls using a stoichiometric CT calibration method, previously described by Schneider *et al*[6]. This resulted in ten different bone subsets that were selected by dosimetric properties. From this work, a general inverse proportionality between hydrogen content and bone density of the obtained tissue composition subsets can be deduced. Furthermore, using the database of the National Institute of Standards and Technology (NIST) XCOM[7] and ESTAR[8], one can easily notice that hydrogen differs from other human body elements in both mass attenuation coefficient and total mass stopping power. Mass attenuation coefficients for a mixture can be determined by the weighted sum of the constituent elements[9]. Also for the mass stopping powers of compounds this method can be used as a close approximation[10]. As these two quantities govern dose deposition in a Monte Carlo algorithm, it is reasonable to assume that only the hydrogen content could be sufficient for dose calculation in human tissues and it is considered as the main hypothesis of this study.

In current work we demonstrate that for mega-voltage photon radiotherapy, only the hydrogen content of the different tissues is of interest. Treatment plans on the CyberKnife® accelerator have been used in this study to estimate the accuracy of such a hypothesis. Current observations might have an impact on MRI based dose calculations and on MVCT calibration using tissue substitutes.

II. Methods

A. Stoichiometric model photon attenuation and energy deposition characteristics of dosimetric tissue subsets

The cornerstone of our work is based on the stoichiometric calibration method described by Vanderstraeten *et al*[5,6], applied to our Aquilion 16LB CT scanner (Toshiba Medical Systems, Otawara, Japan). The aim of this method is to divide tissue materials into more dosimetric subsets than usually described in order to depict tissue diversity in a better way. The first step of this method is to determine the empirical parameters of the total attenuation coefficient (in the CT beam) from the contribution of each individual chemical element within the following system of equations:

$$\mu = \rho N_A \sum_{i=1}^{n} \left(\frac{w_i}{A_i} \left(K^{ph} Z_i^{4.62} + K^{coh} Z_i^{2.86} + K^{KN} Z_i \right) \right) \quad (1)$$

$$H = 1000 \left(\frac{\mu}{\mu_{water}} - 1 \right) \quad (2)$$

with ρ the mass density, N_A the Avogadro constant, w_i the elemental weight of the element i, A_i its atomic mass, Z_i its atomic number, and H the CT number. K^{ph}, K^{coh} and K^{KN} are the parameters to be determined experimentally for our CT beam (they are CT scanner dependant). They characterize respectively the abundance of the photoelectric effect, of the Rayleigh scattering, and of the Compton scattering, in terms of cross section. For this step, the so-called cheese phantom (Accuray Inc., Sunnyvale, CA) with 13 inserts of known density and elemental composition was scanned (120 kVp beam, field of view of 70 cm, 1.0x1.37x1.37 mm^3 voxel size) and the corresponding CT numbers were measured. The parameters were determined by a multiple linear regression and they are given here for information: $K^{PH}/K^{KN} = 3.51 \times 10^{-5}$; $K^{coh}/K^{KN} = -1.82 \times 10^{-3}$. Once the parameters are defined, the equations (1) and (2) can be used to calculate the CT number of any tissue of known density and elemental composition. A table of more of 70 tissues classified by Woodard and White[3] was first used in this way to establish the mass density calibration curve.

The second step of the stoichiometric method in Vanderstraeten *et al*[5] consists of gathering these tissues into discrete dosimetric tissue subsets with similar properties in function of CT number. Composition of theses subsets is obtained by the mean composition of the tissues in the subset. A general relation between hydrogen content and density is observable (figure 1).The hydrogen weight fraction of a subset is close to that of the tissues composing this subset (small standard deviation). The number and the range of the subsets in term of CT number depend on a dosimetric criterion which states that the Monte Carlo simulated depth dose ratios in a large cubic homogenous phantom from two adjacent subsets (with mass density set to 1 $g.cm^{-3}$) do not differ by more than 1 %. In current work, dosimetrically equivalent tissue subsets are generated using Monte Carlo calculations in homogeneous phantoms and from preconceived results of Vanderstraeten *et al*[5] regarding the approximate number of intervals and the boundaries between them. Simulations were performed with DOSXYZnrc[11] using a phase space of a CyberKnife® (Accuray Inc., Sunnyvale, CA) unit, previously modeled[12]. The result is an almost continuously changing tissue composition in consecutive subsets and thus a continuously changing hydrogen content, with a continuously changing density (from the mass density calibration curve). This specific calibration scheme will be considered as the reference scheme.

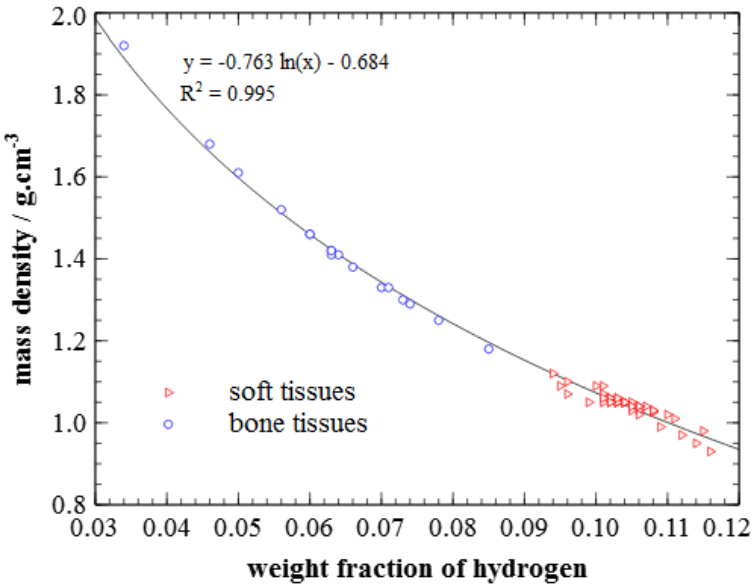


Figure 1 Correlation between hydrogen content and mass density of the human tissues provided by Woodard and White[3].

To demonstrate the importance of hydrogen, a dosimetric subset scheme based on the original was generated by keeping the correct hydrogen weight fraction in all subsets and by allocating the remaining weight fraction to oxygen (replacing all non-hydrogen elements by oxygen). Only the hydrogen content is considered this way. This scheme will be denoted the HO scheme (Hydrogen Oxygen).

Alternatively, a third scheme was derived from the original by keeping a constant hydrogen weight fraction of 11 % within each subset. The difference between the actual hydrogen weight and current value is transferred to the oxygen fraction. This scheme was named the HA scheme (Hydrogen Adipose) as adipose has a hydrogen weight fraction of 11%. In this scheme all non-hydrogen elements are modeled correctly (except oxygen). This scheme is expected to provide noticeable deviations from the reference scheme under the assumption that the hydrogen content mostly governs particle interactions.

The mass attenuation coefficients and total mass stopping power have been calculated using the NIST databases and have been compared for the three schemes in order to provide primary indications of the validity of the main hypothesis.

B. Monte Carlo doses calculations

An in-house developed Monte Carlo system based on MCDE[13] is currently used in our center as a dose verification tool for quality assurance for all CyberKnife treatment plans performed in our radiotherapy department. This system is based on the EGSnrc[14] code and its simulation engines BEAMnrc[15] for the accelerator phase space and DOSXYZnrc[11] for the particle transport in the patient geometry. This required a professional reprogramming of MCDE to obtain a more robust user-friendly system using a JavaScript GUI that allows launching the calculations on different servers and evaluating the obtained results. Different Monte Carlo engines (EGS, Geant4[16]) can be selected

for different parts of the simulation (e.g. Geant4 can be used to determine the phase-space and EGS to perform the dose calculation in the patient geometry or vice versa).

The CyberKnife accelerator has been modeled using BEAMnrc (and the egs++ geometry package to model the Iris collimator), generating phase space files at the exit plane of the secondary collimator[12]. The system allows selecting a treatment plan established on the Multiplan® (Accuray, Inc., Sunnyvale, CA) Treatment Planning System (TPS) and to compare the dose distribution and Dose Volume Histograms (DVH) between Multiplan and the Monte Carlo dose engine.

The different dosimetric subset schemes were retrospectively applied to patient treatment plans using the Monte Carlo system. The simulations were realized on 25 CyberKnife treatments divided in anatomical locations according to table 1. Different anatomical areas were used to cover a large panel of tumor tissues and tissues near the tumor, as for example the bone metastasis which allow testing the modelization of the bone subsets for the HO scheme.

Localization	CyberKnife
lung	6
head	10
pelvis	2
pancreas	1
ENT (ear, nose, throat)	1
bone metastasis	5
Total	**25**

Table 1 Anatomical repartition of the treatment plans. PTVs in the head are localized in soft tissues or close to bone of the skull. The pelvis group gathers prostate, cervix, endometrium, and anus. The bone metastasis group gathers vertebra, femoral head, pelvic bone, and femur.

CyberKnife treatment plans are composed of more than one hundred beams with diameters between 7.5 mm and 60 mm. CTs were scanned at a resolution of 1x1x1 mm^3 and the resolution of the dose voxels in the Monte Carlo system was set at 2x2x2 mm^3. The number of histories simulated was fixed to 10^9 to get a statistical uncertainty below 1 % within at least 95 % of the voxels in the Planning Target Volume (PTV). Dose calculations using the Monte Carlo system were performed using the three dosimetric subset schemes for each treatment plan.

Comparison analyses of dose indexes from the DVH were performed. It was considered as the optimal tool to compare dose results since optimization and evaluation of treatment plans are routinely based on DVHs. DVH, dose repartition and dose profiles along the PTV are presented for one CyberKnife treatment plan allowing a more detailed analysis.

III. Results

A. Stoichiometric model photon attenuation and energy deposition characteristics of dosimetric tissue subsets

Discrepancies between the CT numbers calculated from the stoichiometric model and the measured values are presented in table 2 allowing a qualitative appreciation of the coherence between the model and the measures. There are globally no significant discrepancies and the results are in better

agreement than in Vanderstraeten *et al*[5]. The density calibration is presented in figure 2. The points represent the CT number of the tissues calculated using the stoichiometric model and the Woodard and White tables[3]. The correlation coefficients confirm the validity of the linear fit. As also the case for the more conventional calibration model (using tissue substitutes) the curve is not continuous at 100 HU, but the difference of mass density which could occur from a misattribution of tissue around this point is only 0.03 $g.cm^{-3}$, which can be considered negligible.

insert denomination	mass density ($g.cm^{-3}$)	$H_{measured}$	$H_{calculated}$	residual error $H_{measured} - H_{calculated}$
LN300	0.28	-713	-727	14
LN450	0.48	-540	-532	-8
adipose	0.945	-95	-91	-4
breast	0.984	-55	-39	-16
solid water	1.019	-4	7	-11
brain	1.051	13	35	-22
liver	1.094	76	90	-14
inner bone	1.15	223	196	27
B-200 bone	1.157	222	200	22
CB2-30%	1.335	476	461	15
CB2-50%	1.561	856	835	21
cortical bone	1.824	1253	1273	-20

Table 2 Differences in terms of CT numbers between measures ($H_{measured}$) and calculations ($H_{calculated}$) using the stoichiometric calibration of the inserts of the cheese phantom. CT numbers are expressed in HU.

The application of the second part of the method of Vanderstraeten *et al*[5] led to the dosimetric tissue subsets shown in table 3 constituting the reference scheme, consistent with their results. The weight fraction of hydrogen and the density within subsets agree with the previous observation of figure 1 that there is an inverse proportionality between these quantities, except for the lung subset which is particular in its definition. Lung tissue in the Woodard and White table has a density of 1.05 and a weight fraction of hydrogen of 10.3%, comparable to muscle tissue for example. To simulate the partial volume effect of the voxelisation of lung with air due to the small structure of pulmonary alveolus and pulmonary bronchus compared to the voxel size, density is artificially lowered within the lung subset without changing the composition. Because of that the inverse proportionality is not highlighted within table 3 although it actually exists.

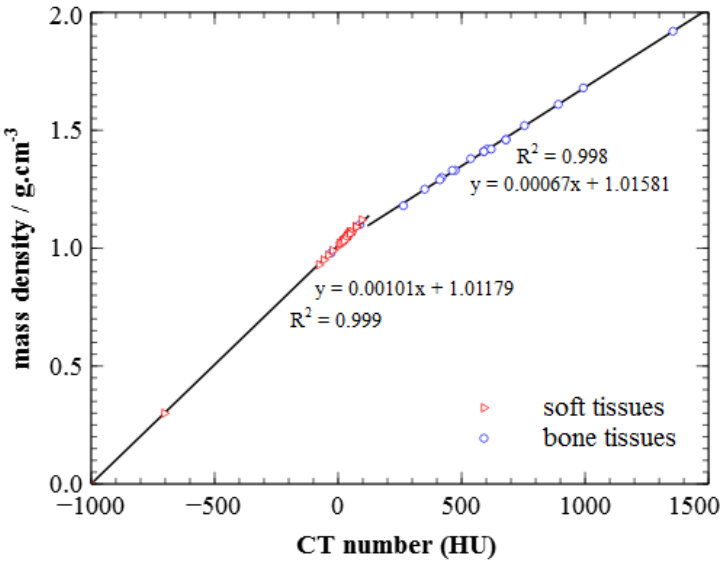


Figure 2 Mass density calibration curve. The equations of linear regression are displayed on the graph. R^2 is the correlation coefficient. Tissue CT numbers are calculated using the stoichiometric model.

subset name	H_l	H_u	w_H	w_C	w_N	w_O	w_{Na}	w_{Mg}	w_P	w_S	w_{Cl}	w_K	w_{Ca}	density
lung	-900	-100	0.103	0.101	0.029	0.755	0.002	0	0.002	0.003	0.003	0.002	0	0.299
adipose	-100	20	0.11	0.345	0.015	0.523	0.002	0	0.001	0.001	0.002	0.001	0	1.001
muscle	20	80	0.103	0.143	0.031	0.713	0.002	0	0.002	0.002	0.002	0.002	0	1.054
bone1	80	200	0.095	0.153	0.042	0.683	0.0055	0	0.011	0.0075	0.003	0	0	1.105
bone2	200	320	0.085	0.405	0.028	0.367	0.001	0.001	0.034	0.002	0.002	0.001	0.074	1.192
bone3	320	440	0.075	0.294	0.037	0.45	0.001	0.001	0.044	0.002	0.001	0.001	0.094	1.278
bone4	440	560	0.069	0.264	0.037	0.46	0.001	0.001	0.053	0.001	0.001	0.001	0.112	1.343
bone5	560	680	0.062	0.291	0.035	0.402	0.001	0.001	0.065	0.002	0.001	0	0.14	1.433
bone6	680	800	0.056	0.235	0.04	0.434	0.001	0.001	0.072	0.003	0.001	0.001	0.156	1.485
bone7	800	920	0.05	0.212	0.04	0.435	0.001	0.002	0.081	0.003	0	0	0.176	1.610
bone8	920	1040	0.046	0.199	0.041	0.435	0.001	0.002	0.086	0.003	0	0	0.187	1.677
bone9	1040	1160	0.042	0.184	0.041	0.435	0.001	0.002	0.092	0.003	0	0	0.2	1.748
bone10	1160	1300	0.038	0.169	0.041	0.435	0.001	0.002	0.098	0.003	0	0	0.213	1.835
bone11	1300		0.034	0.155	0.042	0.435	0.001	0.002	0.103	0.003	0	0	0.225	1.919

Table 3 Elemental composition of dosimetric tissue subsets of the reference scheme. Subsets are represented by a name, a CT number lower limit H_l, a CT number upper limit H_u, an atomic weight fraction w_i and a continuous mass density calculated from the mass density calibration curve. The density provided in the last column is the mean density of the tissues within the subset and is given to show the inverse proportionality between density and hydrogen content.

To illustrate the generation of the two additional designed schemes from the reference scheme, the construction of the bone6 subset of the HO scheme and the HA scheme is given in table 4 as an example. In this subset, the error induced on the weight fraction of hydrogen in the HA scheme is almost twice, to regard with the actual value of 5.6% in the HO scheme.

scheme	w_H	w_C	w_N	w_O	w_{Na}	w_{Mg}	w_P	w_S	w_{Cl}	w_K	w_{Ca}
reference	0.056	0.235	0.04	0.434	0.001	0.001	0.072	0.003	0.001	0.001	0.156
HO	0.056			0.944							
HA	0.11	0.235	0.04	0.380	0.001	0.001	0.072	0.003	0.001	0.001	0.156

Table 4 Composition of the bone6 subset for the three schemes under consideration

The following representations of the mass attenuation coefficients and the mass stopping powers of both the HO and the HA schemes were normalized to those of the reference stoichiometric calibration scheme allowing better visualization of the impact of the applied modifications in tissue composition. The mass attenuation coefficients of several subsets of the HO and HA schemes are represented in figure 3. Not all subsets have been displayed for clarity. The hidden bone curves follow the same trend as displayed. Three main parts can be distinguished for both the HO and the HA schemes.

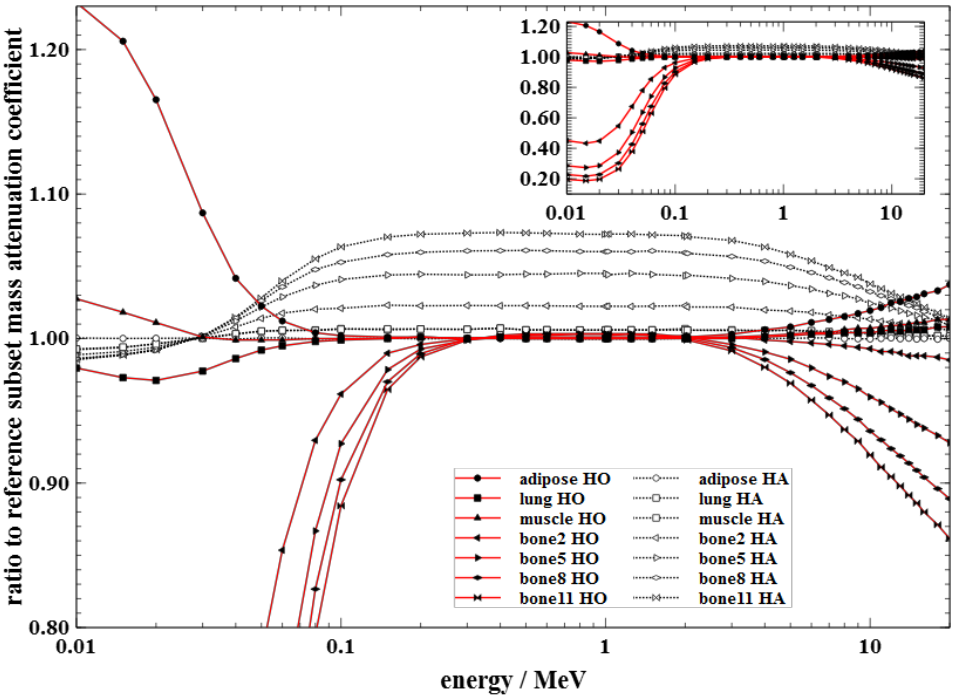


Figure 3 Ratio to reference subset mass attenuation coefficients of some subsets for the HO scheme and the HA scheme. Not all subset curves have been displayed to improve clarity. The top right window shows the same curves on a higher vertical range.

The most important part is the energy range between 200 keV and 3 MeV where the reference scheme and the HO scheme present almost the same attenuation throughout all subsets. Maximum difference in this range is about 1 % for bone11 namely the subset with the lowest hydrogen content. This result is consistent with the results obtained by Seco and Evans[17] who have demonstrated that in this energy range the attenuation is proportional to electron density. Electron density is proportional to the weighted ratio of atomic number to atomic weight Z_i/A_i, which roughly equals 0.5 for all chemical elements of the human body except for hydrogen having a ratio of 1. Replacing elements (except hydrogen) by oxygen does not modify the weighted ratio of Z_i/A_i nor the attenuation.

Deviations arise above 3 MeV within all subsets, with larger deviations for bone tissues as shown by the example of bone11 with a ratio of 0.86 at 20 MeV. Seco *et al*[17] have also shown pair production becomes important above 10 MeV especially for high Z elements such as calcium. A certain part of the pair production is lost when calcium is not considered within the HO scheme and leads to a decrease of attenuation. Nevertheless the use of the ratio virtually intensifies these discrepancies above 10 MeV: in this range, discrepancies ranging from 5 % to 15 % between mass attenuation coefficients of bone can be noticed. As the curve of the mass attenuation coefficient has a negative slope for these subsets and is on the order of 10^{-2} beyond 10 MeV, the relative difference for instance between bone11 (reference scheme) and bone11 (HO scheme) of absolute attenuation is less than 2.5 % for a distance of 5 cm. Such differences are difficult to translate to dosimetric impact but they seem negligible. Furthermore the energy spectrum of the CyberKnife[18] has a small part of photons between 3 MeV and about 6 MeV.

Deviations also occur below 300 keV within all subsets, especially for the mass attenuation coefficient of bone tissues whose ratio is a factor 5 at 20 keV. Calcium content within the stoichiometric scheme is the origin of such deviations. Indeed the high atomic number of calcium provides more attenuation by the photoelectric effect in this energy range. Although the energy spectrum of the CyberKnife is centered around 1 MeV and does not contain a lot of photons of these energies[18], a proportion of photons could reach such energies from scattering or Bremsstrahlung. This could cause some deviations in the dose calculation, but the effects of these discrepancies on the dose calculation are difficult to predict and may be weak.

Contrarily, even if the HA scheme curves are more faithful within the energy range extremities, they differ from the reference within the main energy range from 2.2 % for the bone2 subset, 4.5 % for the bone5 subset, 6.0 % for the bone8 subset and 7.3 % for the bone11 subset at 1 MeV. These discrepancies remain almost constant between 100 keV and 5 MeV. The adipose subset ratio is unity along the energy range because the hydrogen weight fraction of the adipose subset is 11%. These discrepancies were expected as the hydrogen weight fraction difference with the reference scheme increases. Calcium content has only an effect for low energies and energies above 10 MeV as shown before, and has almost no impact on the mass attenuation coefficient and the mass stopping power between these thresholds. The HA scheme should introduce noticeable deviations from the stoichiometric scheme within bony targets.

The total mass stopping powers for both the HO and HA schemes are displayed in figure 4. The mass stopping power curves show larger differences among the subsets and energies. The soft tissue subsets from both the HO and the HA schemes have approximately the same mass stopping power as the reference calibration scheme throughout the energy range; the deviations are well below 2%.

The bone2 subset from the HO scheme is consistent within 0.6 % throughout the energy range. The bone5 subset from the HO scheme has a mass stopping power that differs by less than 1.3 % from 10 keV to 12 MeV and reaches a 2.5 % difference at 20 MeV. The difference in mass stopping powers within the bone8 subset from the HO scheme is between 3 % and 2 % from 10 keV to 100 keV, near 1.8 % up to 1 MeV, progressively decreases to 0% at 4 MeV and to -4 % at 20 MeV. The differences in mass stopping powers are 4 % at 10 keV, 2 % at 1 MeV, near 0 % at 4 MeV, -1 % at 7 MeV and 5 % at 20 MeV within the bone11 subset from the HO scheme. This spread of discrepancies in stopping powers is more difficult to analyze, and the prediction of the impact on the dose calculation is more

challenging. The energy range from about 6 to 20 MeV should not be taken into account for the CyberKnife. The differences are however small for the energies of interest. These observations on the attenuation coefficient and the stopping power will lead to an overall result which will be studied in the following section.

For the HA scheme, the bone subset curves deviate as a function of the difference in hydrogen content. Even the soft bone subsets suffer from this difference: 2.7 % difference at 1 MeV for the bone2 subset for instance. For the bone11 subset the discrepancy goes up to 8.7 % at 1 MeV. These discrepancies in the HA scheme would lead most likely to erroneous results in bony structures and point out the importance of the hydrogen content.

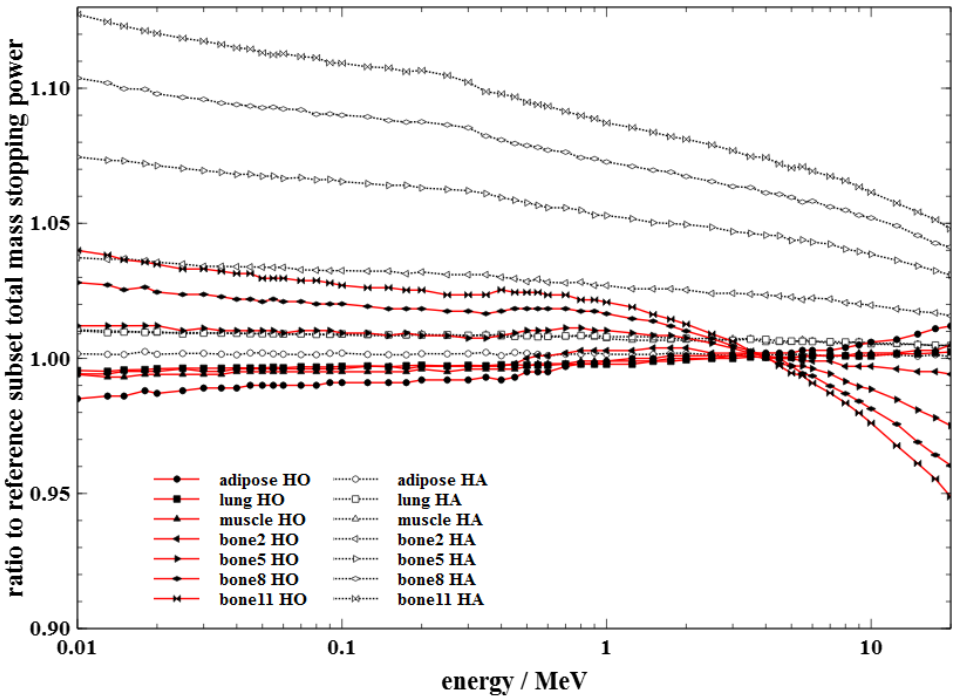


Figure 4 Ratio to reference unrestricted mass stopping powers of some subsets for the HO scheme and the HA scheme.

B. Monte Carlo doses calculations

D_{02}, D_{05}, D_{50}, D_{95}, D_{98} and D_{mean} of the PTV were evaluated for each treatment plan using the three dosimetric tissue subset schemes. A distribution of the sample has been established for each dose index. The main hypothesis is that there is no difference between the HO scheme and the reference scheme, i.e. the null hypothesis corresponds to the mean of the population being zero. A threshold of 0.05 (p-value) is used for the t-test for significant divergence of the null hypothesis. Complementary the standard deviation of the population should stay below a defined threshold, defined as: 3σ should be inferior to 2% (to have 99% of the values below 2%). For simplicity, the standard deviation of the population has been estimated as the standard deviation of the sample. Relative difference of these indexes between the reference and the HO scheme are presented in figure 5. The red windows present the summary of the characteristics of the sample. The mean of the difference is comprised between 0% and 0.1%, which is consistent with the null hypothesis. All p-values are above the defined threshold, which further validates the null hypothesis. The low spread

of the differences for each dose index explains the low standard deviation. This complies with our defined criterion of 2% (3σ) except in D_{98}. However a systematic error in small targets seems to be present in the dose index because of the small number of voxels inside the PTV. In spite of this eventual error, the value of 2.1% (3σ) has been accepted. The positive extreme value of the D_{98} distribution (1.9%) arises from a bone metastasis case (vertebra) which is presented as an example in figures 7 and 8. This small discrepancy can be attributed to a lack of calcium within the bone region when using the HO scheme. However this difference in D_{98} is hardly visible in the DVH and is not clinically significant as it will not influence the validation or the optimization of a treatment plan. The DVHs obtained for the HO and the reference schema are hard to distinguish, indicating a good agreement. In general all the DVHs compared present the same trend of quasi-superposition between the two curves and are therefore clinically equivalent. There is a concordance between the two schemes even in the bone regions and in regions close to the skin, and the behavior of attenuation at low energies did not introduce a significant bias. These overall results suggest that modeling of tissue by only density and hydrogen weight fraction is well suited for accurate dose calculations within any type of tissue for the CyberKnife treatments.

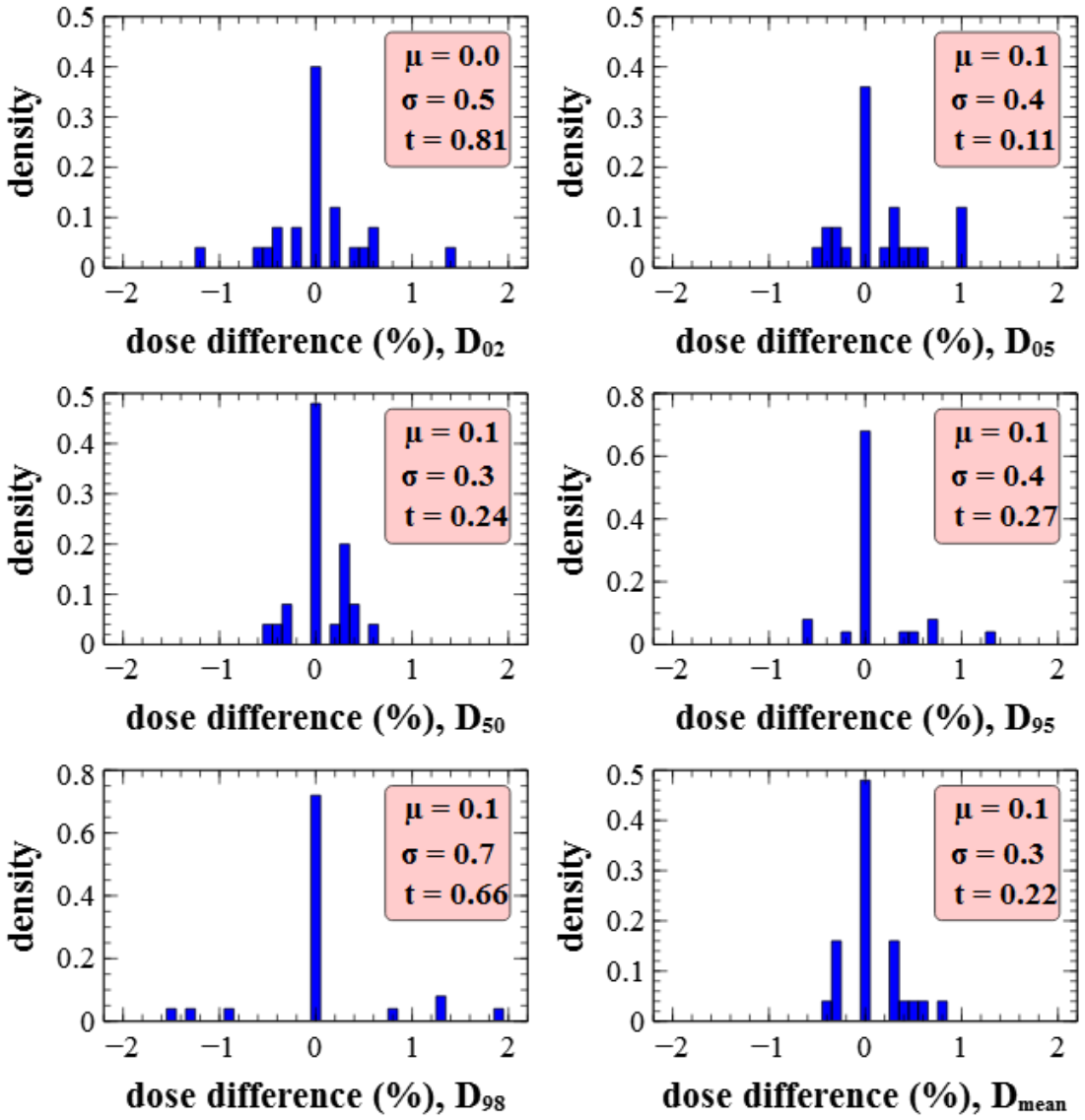


Figure 5 Density of relative dose difference of several dose indexes between the reference scheme and the HO scheme for all treatments. Dose difference is expressed as $(D_{X,ref} - D_{X,HO})/D_{X,ref}$. Density is normalized to unity. μ is the mean of the distribution, σ is the standard deviation of the distribution, and t is the p-value of the distribution for the null hypothesis (mean equal zero).

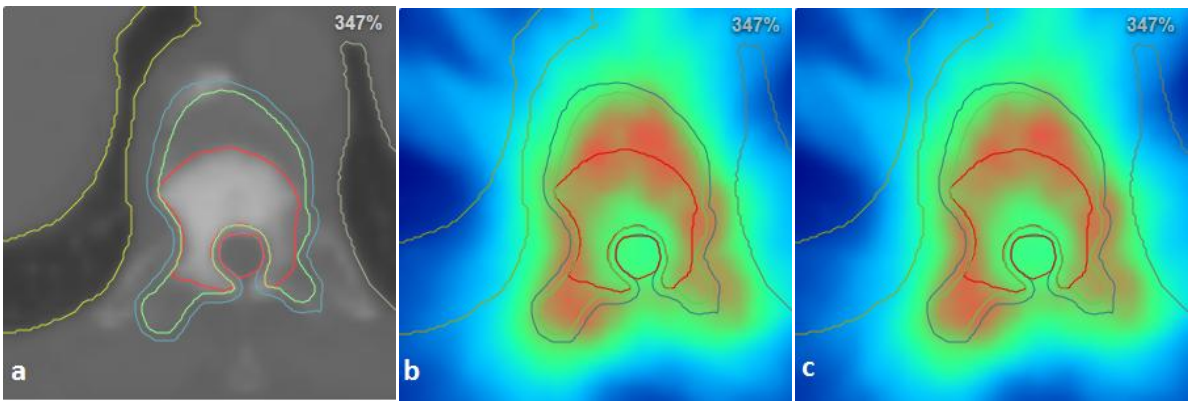


Figure 7 Delineation of the GTV (red contour), the CTV (green contour) and the PTV (blue contour) comprising the bone marrow (brown contour) of a T11 vertebra palliative treatment (a). Dose repartition using the reference scheme (b) and

using the HO scheme (c). Color scale goes to dark blue (0 Gy) to green (23 Gy -30 Gy) up to intense red (42 Gy). D_{mean} to PTV: 32.1 Gy. Maximum dose to bone marrow is under 30 Gy.

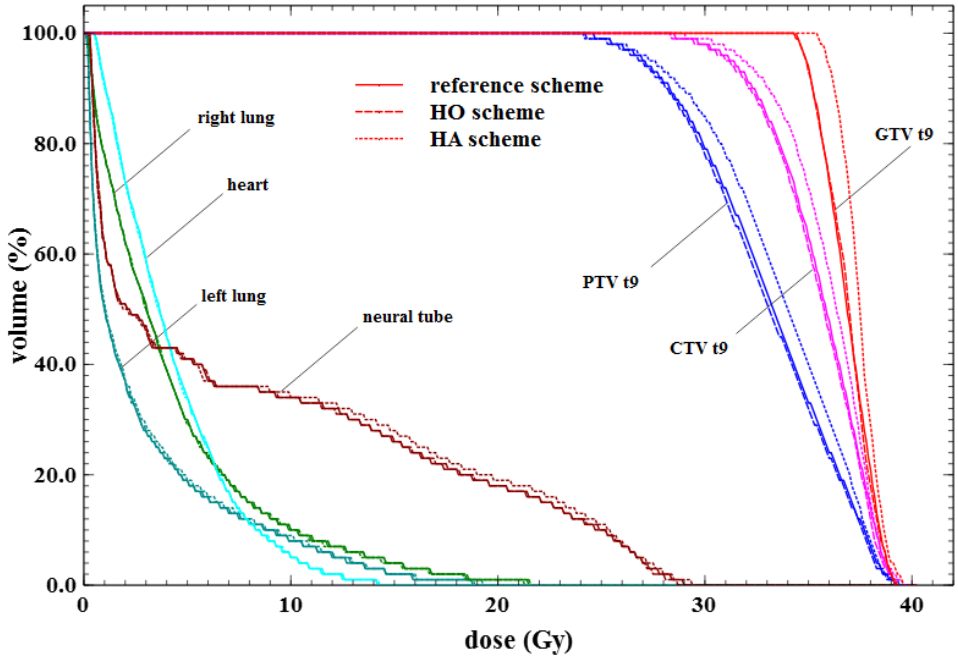


Figure 8 DVH of a vertebra target case. One color (online version) corresponds to one volume. Full line corresponds to the reference scheme, dashed to the HO scheme, dotted to the HA scheme.

Dose repartition can be evaluated qualitatively in figure 7 for the same vertebra case. Local higher dose (in intense red) are linked into the two repartitions, and there is no visual discrepancy. Dose profiles were drawn through the PTV and are presented in figures 9 and 10. The profiles of the two schemes follow the same trend with some local discrepancies which are statistical. These local differences are however averaged on the entire PTV as the previous results indicated.

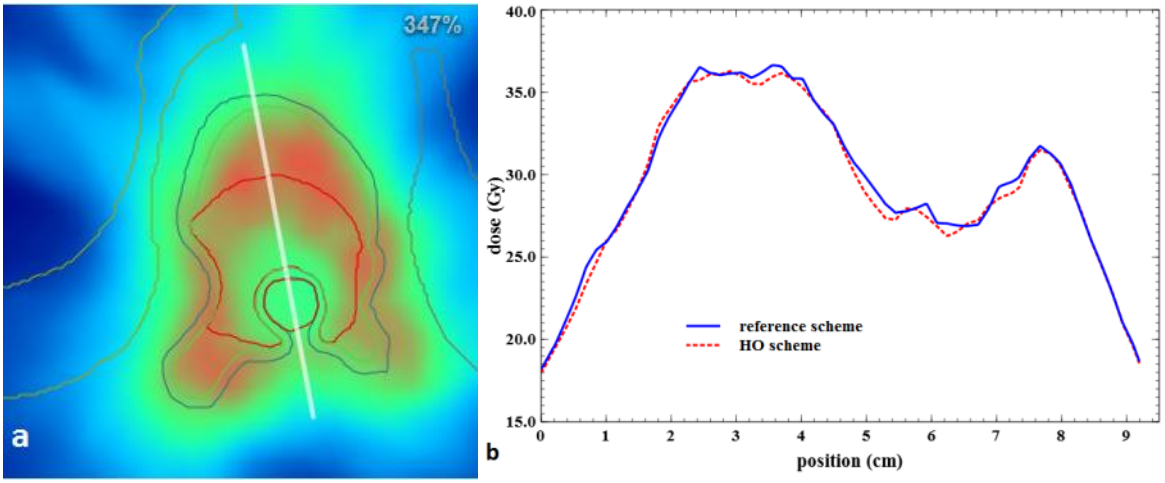


Figure 9 Line profile (top to bottom) crossing the tumor and the bone marrow of the same treatment plan than on figure 7 (a) and dose value along this line for the reference scheme and the HO scheme (b).

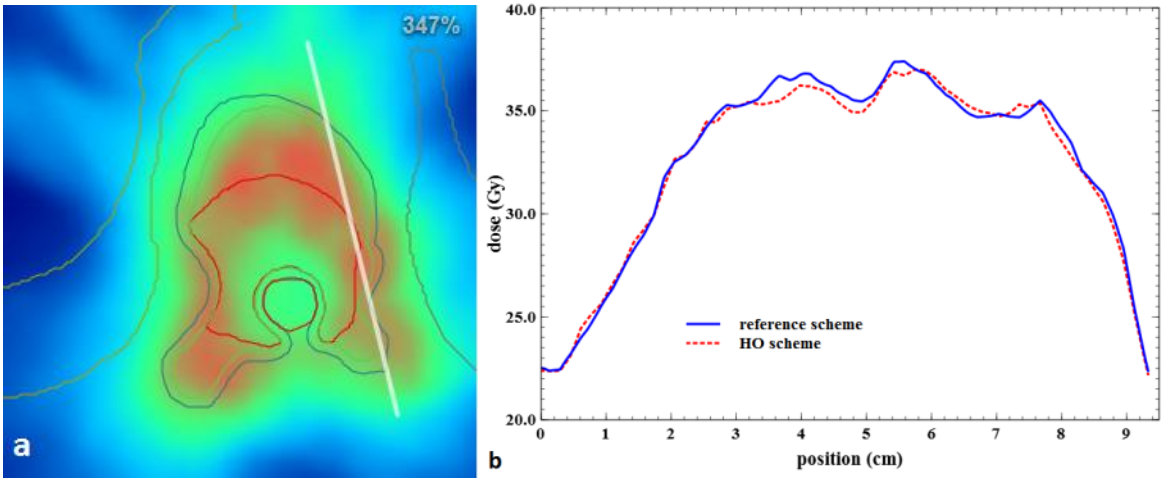


Figure 10 Line profile (top to bottom) crossing the tumor of the same treatment plan than on figure 7 (a) and dose value along this line for the reference scheme and the HO scheme (b).

Relative difference of dose indexes between the reference and the HA scheme are presented in figure 6. The same comparisons were made as described above, although only a few results are needed to prove that the HA scheme is not equivalent to the reference scheme. For example, the p-value of D02, D05, D50 and Dmean are inferior to 0.01 indicating that the mean of their respective population is not zero. The second argument is the too high standard deviation (superior to 0.9% for all dose indexes) which cannot be accepted. The discrepancies obtained with the HA scheme occur for bone targets or targets close to bone tissues, which demonstrates that the hydrogen content is an important factor in dose calculation whereas the calcium content has little impact. Such a deviation is shown in figure 8 where the HA scheme DVH systematically overestimates the dose inside the PTV. Discrepancies inferior to ±1% in dose indexes are obtained within targets of soft tissue when there is no bone in close proximity. This was predictable and due to the intrinsic constitution of the HA scheme in terms of hydrogen content.

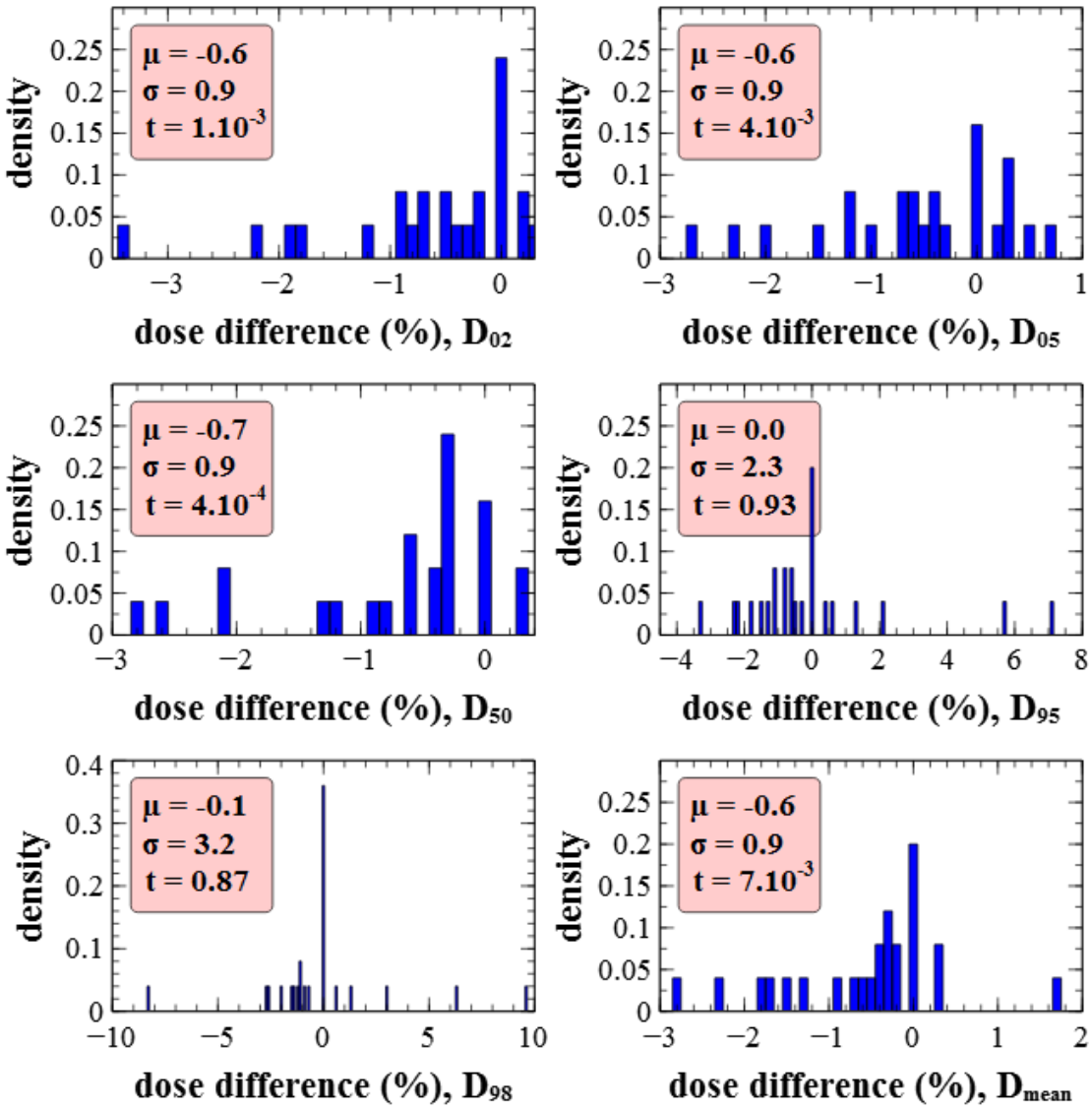


Figure 6 Density of relative dose difference of several dose indexes between the reference scheme and the HA scheme for all treatments. Dose difference is expressed as $(D_{X,ref} - D_{X,HA})/D_{X,ref}$. Density is normalized to unity. μ is the mean of the distribution, σ is the standard deviation of the distribution, and t is the p-value of the distribution for the null hypothesis (mean equal zero).

IV. Discussion

The main hypothesis postulated after studying the attenuation coefficients and the stopping powers has been verified by Monte Carlo dose calculations using the HO scheme. There are small differences in dose indexes and DVHs between the stoichiometric scheme and the HO scheme even within the bone structures. The quasi-superposition of the DVH curves for these two schemes in all studied cases is an important result since the DVH is the main tool for evaluation and optimization of a treatment plan. These overall results prove that modeling of tissue by only hydrogen weight fraction and density is suitable for accurate dose calculations within any tissue type of the target for the CyberKnife treatments. As the CyberKnife beam is produced without any hardening filter, the energy spectrum contains a higher component of photons with energy around 0.6-1 MeV compared to a

conventional 6 MV spectrum. As mass attenuation coefficients and mass stopping powers for the HO scheme are in good agreement in this energy range as seen in figures 3 and 4, using a conventional 6 MV spectrum will not produce any alteration in the dose results. Previous conclusion can thus be applied more generally in the mega-voltage photon radiotherapy field. However no study on higher nominal energy like 18 MV or 20 MV was initiated due to the trend to use less and less this type of beams particularly in the fields of IMRT and stereotaxy.

This result can be of prime interest in MRI-only based Monte Carlo dose calculation, which is of great interest taking into account the high tissue contrast of MRI for target and organ delineation and removing the uncertainty induced by the CT to MRI registration. The main obstacle hampering the usage of MRI as the unique modality in treatment planning is the lack of information about electron density. Current methods to derive electron density from MRI data by generating a pseudo-CT are divided into two approaches: the anatomy based and the voxel based methods. In the anatomy based method, deformable registration of reference paired MRI/CT scan (atlas or reference patient) to the MRI scan of the patient under consideration is used to build a pseudo CT from the warped reference CT scan[19,20,21]. Although mathematical or statistical operations could be used to smooth registration errors due to inter-patient anatomical differences[19,20], these method fail for patients presenting high anatomy dissimilarities. The voxel based methods convert MRI intensity into pseudo CT number using either a bulk density assignment after segmenting tissues or using a regression model to compute continuous values of CT number from a set of MRI data[22,23,24,25]. Ultra short echo time sequences[26] are usually used in voxel based methods to distinguish the signal of bone from air. As the signal is acquired very close to the end of the excitation in UTE imaging, almost no supplemental weighting of the signal is added to the proton density weighting, a link could be identified between UTE data and hydrogen density[27]. This may open the way to a voxel based method to convert MRI data to hydrogen content which could then be used for treatment planning by the method presented in this study. This method would rely on a physical basis whereas all the current methods rely on mathematical models. This work is however outside the scope of this study. Another field of interest for this study is the use of tissue substitutes to calibrate MVCT devices for adaptive treatment planning. For adaptive Tomotherapy for example, MVCT is required to calculate dose on the patient geometry during treatment delivery. A calibration curve is needed to link MVCT intensities to tissue density. Phantom inserts that are considered tissue equivalent for kVCT calibration (determined by the high Z component) are not necessarily optimal for MVCT calibration. As demonstrated in current paper, the degree of tissue equivalence is determined by the hydrogen content, while the high Z elements have no impact. However, sometimes discrepancies in hydrogen content can be compensated by changing the density, as e.g. for Solid Water® (RMI 457, Gammex, Middleton, WI). This issue has to be investigated in more detail.

V. Conclusion

Our results are in agreement with our hypothesis that hydrogen weight fraction and density suffice to perform accurate Monte Carlo dose calculations within any tissue type in mega-voltage photon radiotherapy conditions. The calcium content of bone tissues is not required for Monte Carlo simulation of a 6 MV beam. We have linked these results to mass attenuation coefficient and mass stopping power properties of designed subset schemes which explains the success of this method.

This outcome may be of a high importance in the context of an MRI-only radiotherapy workflow. Indeed MRI is based on the concentration of protons inside tissues so a method could be designed to extract hydrogen content from one or several MRI sequences. This result can also be used in the field of calibration of MVCT devices to select optimal tissue substitutes. This selection should be based on the hydrogen content instead of the high z-elements that are important for KVCT calibration.

References

1. Ma C-M, Mok E, Kapur A, et al. Clinical implementation of a Monte Carlo treatment planning system. *Med Phys.* 1999;26(10):2133.
2. Wang L, Chui C-S, Lovelock M. A patient-specific Monte Carlo dose-calculation method for photon beams. *Med Phys.* 1998;25(6):867.
3. Woodard HQ, White DR. The composition of body tissues. *Br J Radiol.* 1986;59(708):1209-1218.
4. Verhaegen F, Devic S. Sensitivity study for CT image use in Monte Carlo treatment planning. *Phys Med Biol.* 2005;50(5):937-946.
5. Vanderstraeten B, Chin PW, Fix M, et al. Conversion of CT numbers into tissue parameters for Monte Carlo dose calculations: a multi-centre study. *Phys Med Biol.* 2007;52(3):539-562.
6. Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol.* 1996;41(1):111-124.
7. Berger MJ, Hubbell JH, Seltzer SM, et al. XCOM: Photon Cross Section Database (version 1.5). 2010.
8. Berger MJ, Coursey JS, Zucker MA, Chang J. ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3). 2005.
9. Jackson DF, Hawkes DJ. X-ray attenuation coefficients of elements and mixtures. *Phys Rep.* 1981;70(3):169-233.
10. Bragg WH, Kleeman R. XXXIX. On the α particles of radium, and their loss of range in passing through various atoms and molecules. *Philos Mag Ser* 6. 1905;10(57):318-340.
11. Walters B, Kawrakow I, Rogers DWO. DOSXYZnrc Users Manual. NRCC Report PIRS-794 (revB). 2011.
12. Wagner A, Crop F, Lacornerie T, Vandevelde F, Reynaert N. Use of a liquid ionization chamber for stereotactic radiotherapy dosimetry. *Phys Med Biol.* 2013;58(8):2445-2459.
13. Reynaert N, Smedt B De, Coghe M, et al. MCDE: a new Monte Carlo dose engine for IMRT. *Phys Med Biol.* 2004;49(14):N235-N241.
14. Kawrakow I, Mainegra-Hing E, Rogers DWO, Tessier F, Walters BRB. The EGSnrc Code System: Monte Carlo Simulation of Electron and Photon Transport. NRCC Report PIRS-701; 2013.
15. Rogers DWO. BEAM: A Monte Carlo code to simulate radiotherapy treatment units. *Med Phys.* 1995;22(5):503.
16. Agostinelli S, Allison J, Amako K, et al. Geant4—a simulation toolkit. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip.* 2003;506(3):250-303.
17. Seco J, Evans PM. Assessing the effect of electron density in photon dose calculations. *Med Phys.* 2006;33(2):540.
18. Ma C-M, Li JS, Deng J, Fan J. Implementation of Monte Carlo Dose calculation for CyberKnife treatment planning. *J Phys Conf Ser.* 2008;102:012016.

19. Andreasen D, Van Leemput K, Hansen RH, Andersen JAL, Edmund JM. Patch-based generation of a pseudo CT from conventional MRI sequences for MRI- only radiotherapy of the brain Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain. *Med Phys.* 2015;42(4):1596-1606.

20. Gudur MSR, Hara W, Le Q-T, Wang L, Xing L, Li R. A unifying probabilistic Bayesian approach to derive electron density from MRI for radiation therapy treatment planning. *Phys Med Biol.* 2014;59(21):6595-6606.

21. Dowling J, Lambert J, Parker J, et al. An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy. *Int J Radiat Oncol Biol Phys.* 2012;83(1):e5-e11.

22. Kapanen M, Tenhunen M. T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning. *Acta Oncol.* 2013;52(3):612-618.

23. Korhonen J, Kapanen M, Keyriläinen J, Seppälä T, Tenhunen M. A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer. *Med Phys.* 2014;41(1):011704.

24. Johansson A, Karlsson M, Nyholm T. CT substitute derived from MRI sequences with ultrashort echo time. *Med Phys.* 2011;38(5):2708.

25. Edmund JM, Kjer HM, Van Leemput K, Hansen RH, Andersen JA, Andreasen D. A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times. *Phys Med Biol.* 2014;59(23):7501-7519.

26. Robson MD, Gatehouse PD, Bydder M, Bydder GM. Magnetic resonance: an introduction to ultrashort TE (UTE) imaging. *J Comput Assist Tomogr.* 2003;27(6):825-846.

27. Ma W, Sheikh K, Svenningsen S, et al. Ultra-short echo-time pulmonary MRI: Evaluation and reproducibility in COPD subjects with and without bronchiectasis. *J Magn Reson Imaging*. 2014;00.

2.3 Limitations des images ZTE couramment disponibles en routine clinique

Il n'y a pas de séquences ZTE disponible sur l'IRM du Centre Oscar Lambret à Lille (pas compatible). Deux options étaient possibles : attendre l'installation du deuxième IRM prévu en 2017, ou collaborer avec des hôpitaux où ces séquences sont disponibles. L'hôpital Leonard de Vinci de Douai nous a permis d'acquérir ces séries d'images (ZTE, T_1 et CT) sur 4 de leurs patients. . Notre thésard Benjamin Demol a passé quelques après-midis à Douai pour optimiser la séquence ZTE qui n'était pas réellement utilisée en clinique. Malgré ce nombre de patients limité, on est déjà capable d'évaluer le lien entre les UH sur le CT et les niveaux de gris dans les images ZTE. L'épaisseur de l'os du crâne est déjà beaucoup plus réaliste que sur les images des séquences classiques. Par contre, il reste compliqué de distinguer l'os de l'air. Wiesinger et al [62][63] ont récemment publié la courbe montrée en figure 21.

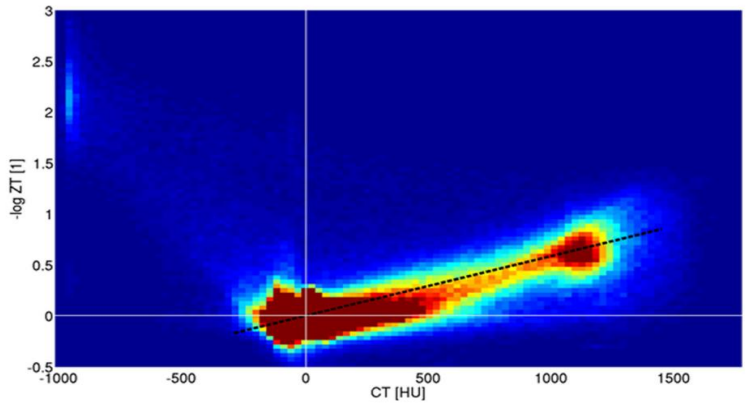
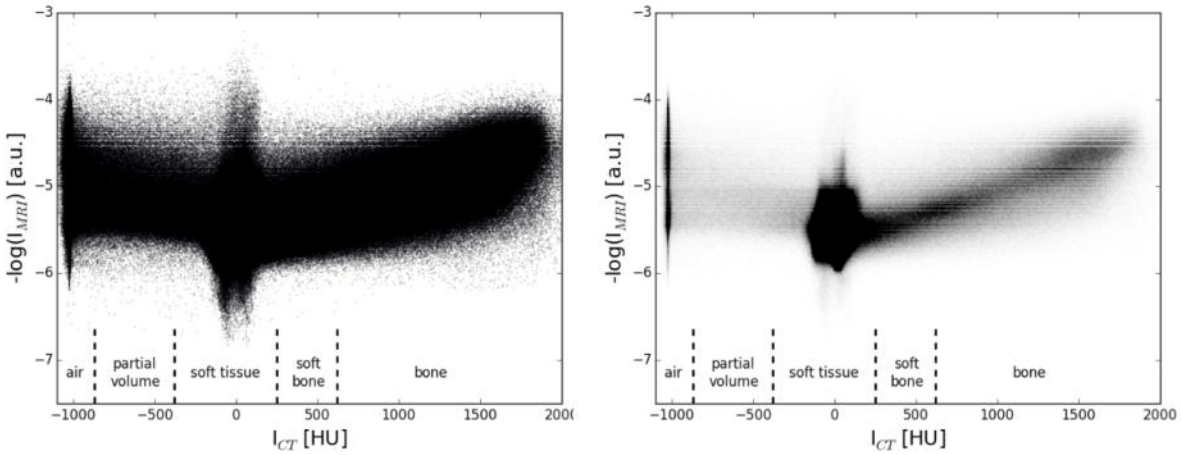


Figure 21 : Lien entre les niveaux de gris de la séquence ZTE avec les UH du CT du même patient (Wiesinger et al 2016 [63]).

Même si à première vue on a l'impression qu'il y a effectivement une relation entre les intensités de l'IRM et celles du CT, il y a quand-même des limitations. Tout d'abord, on voit qu'il y a beaucoup de voxels IRM avec des valeurs élevées (la fonction –log(ZTE)) pour des UH négatives. En plus, la pente de la fonction n'est pas assez élevée et on voit le chevauchement entre l'os et l'air. Nous avons reproduit cette courbe avec nos données en figure 22.



(a) ZTE - high marker size (b) ZTE - low marker size

Figure 22 : Même courbe que la figure 21 pour les données de Douai.

On obtient le même résultat. La conclusion de cette étude est que même la séquence ZTE n'est pas tout de suite applicable pour sortir l'information hydrogène ni pour faire une conversion directe entre IRM et CT. Ce qui démontre l'importance de notre collaboration avec le LTSI de Rennes, afin d'optimiser des séquences ZTE et/ou DUTE, car les séquences couramment disponibles en routine clinique ne sont pas encore adéquates.

2.4 Evaluation des incertitudes dosimétriques associés à l'étalonnage du CT

Dans la littérature scientifique la qualité des images pCT est souvent évaluée utilisant des courbes MAE (« Mean absolute Error »), défini par l'équation suivante :

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left| nCT_{pCT}(i) - nCT_{tCT}(i) \right|$$

Alors pour chaque voxel du pCT on cherche la valeur correspondante dans le vraie CT, et on prend la valeur absolue de la différence entre les deux voxels. Ces différences absolus sont moyenné utilisant des bins spécifiques (de par exemple 30 UH). Ces courbes donnent l'impression que le pCT n'est pas du tout équivalent au vraie CT et qu'il y des différences systématiques. Les images pCT et vraie CT sont recalés, alors il y a systématiquement des shifts entres les images qui vont causer des différences positives et négatives entre les voxels. Le fait de prendre les valeurs absolues donne l'impression que les déviations sont énormes. Alors une amélioration serait de calculer des courbes ME (« Mean Error »), donné par l'équation :

$$ME = \frac{1}{N}\sum_{i=1}^{N} nCT_{pCT}(i) - nCT_{tCT}(i)$$

Mêmes ces courbes sont trop influencées par la qualité de recalage. Afin de l'illustrer nous avons déterminé ces courbes pour des images CT (vraie CT) prises utilisant différents protocoles.

Dans un premier temps, afin d'avoir une idée de l'incertitude dosimétrique associée avec l'étalonnage du scanner, la courbe IVDT (« image value to density table ») est déterminée en variant plusieurs paramètres, comme l'énergie, le diamètre de reconstruction, le FOV, la taille du fantôme, la position des inserts dans le fantôme, … C'était le travail de stage d'un étudiant de deuxième année de la formation DQPRM. L'idée de déterminer cette incertitude est bien cadrée dans le projet PhysiCancer. Le but est de comparer les incertitudes des images pCT avec celles du vrai CT.

Les résultats démontrent qu'il y a 4 paramètres du scan qui peuvent vraiment impacter les UHs du scan :

- L'énergie (100 kV – 135 kV)
- Le noyau de reconstruction (FC64 – FC13)
- L'épaisseur du patient (utilisant un fantôme consistent de plusieurs couches, figure 23)
- Le FOV (surtout le FOV « Large »)

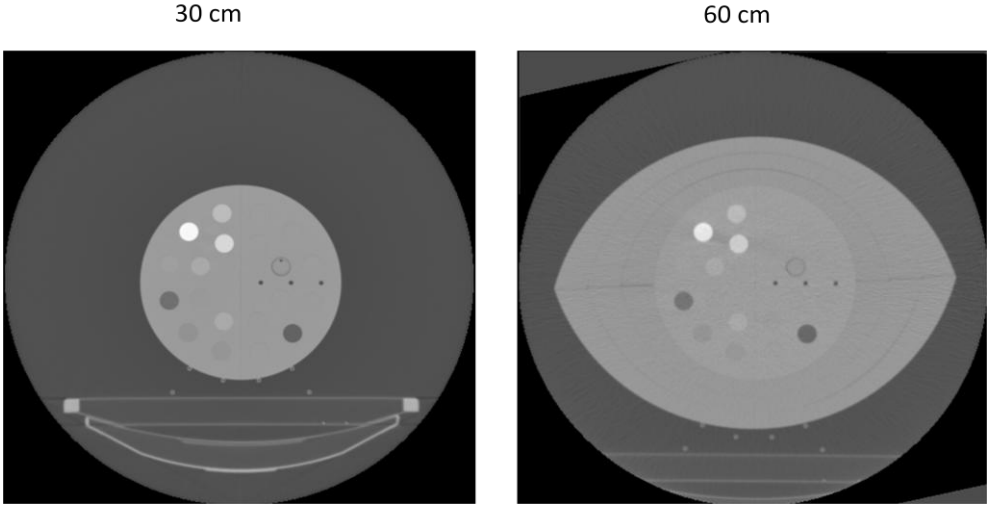


Figure 23 : Fantome « maison » pour étudier la dépendance de l'IVDT sur l'épaisseur du patient.

Les 3 premiers paramètres causent des différences > 300 UH dans les inserts os-équivalents, des écarts comparable avec les pCT (comme montré plus loin dans les papiers). Le FOV par contre donne des différences qui sont plus modéré mais qui sont systématique pour tous les UHs. Afin de tester l'impact dosimétrique, une dizaine de patients sont recalculés utilisant différent des images scanner avec les différents paramètres. Comme on a pour les pCT (où on voit les mêmes écarts en UH par rapport au vraie CT), les différences dosimétrique sont presque toujours négligeable (< 2 %), si on fait varier les paramètres individuellement. L'impact est montré dans les courbes MAE et ME en figure 24.

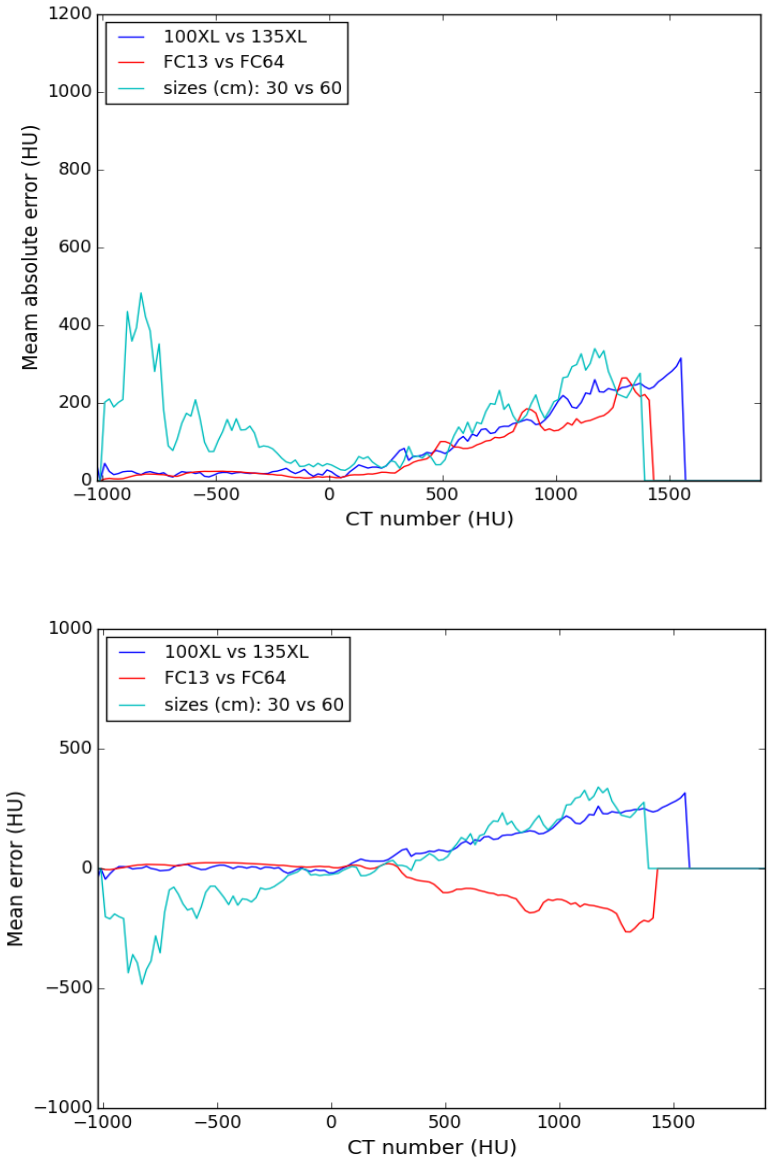


Figure 24 : Courbes MAE et ME pour l'énergie, le noyau de reconstruction et le diamètre du fantôme.

Par contre, l'énergie et le diamètre du fantôme sont souvent corrélés (c'est souvent pour des petits patients qu'on va diminuer l'énergie) et on voit des écarts UH dans le même sens. Alors on peut facilement imaginer qu'on va tomber sur des cas en routine clinique où les écarts entre deux patients vont jusqu'à 900 UH, avec un impact dosimétrique non-négligeable (> 4 %, voir figure 25). Evidemment en pratique, on peut prévoir des IVDTs spécifique à chaque énergie et chaque noyau de reconstruction. Par contre ça devient plus compliquer de prendre en compte l'impact du diamètre du patient. En plus, dans le TPS qu'on utilise au COL pour les traitements conformationnelles (Masterplan, Nucletron), il n'y qu'un IVDT utilisé (défini par Nucletron).

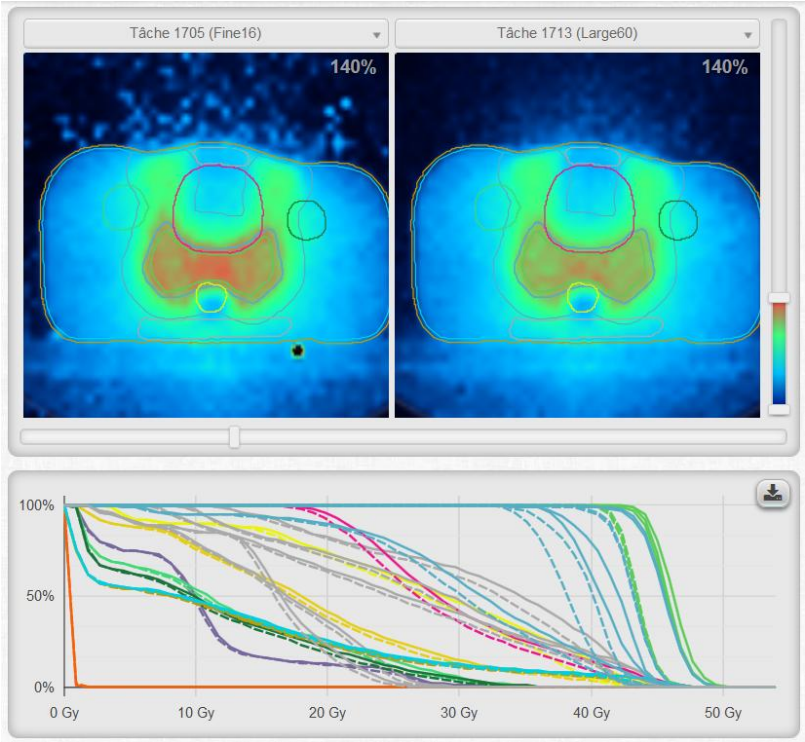


Figure 25 : Comparaison dosimétrique (utilisant Moderato) entre deux IVDT. Fine16 correspond à l'étalonnage utilisant le petit fantôme (16 cm) en 100 kV reconstruit avec le noyau FC64. Large60 correspond à l'étalonnage utilisant un grand fantôme (60 cm) en 135 kV reconstruit avec le noyau FC13. Les différences sont > 5 % dans les PTVs et certains OARs.

On résumé, on va voir que les pCT décrits dans les papiers 4 et 5 ne correspondent jamais parfaitement avec le vraie scan, mais que les déviations sont très comparables avec les incertitudes qu'on a sur le vrai CT, même dans le cas qu'on peut générer des IVDTs spécifiques pour l'énergie, le noyau de reconstruction, le FOV, ….

3. Article 4 : Génération pCT utilisant la méthode atlas

3.1 Introduction

Dans le cadre du volet clinique du projet PhysiCancer, nous avons implémenté une méthode atlas en combinaison avec des recalages déformables (ITK, Elastix) [45][46] pour la génération des images pCT à partir d'une séquence T_1, déjà utilisée en routine clinique pour la définition des volumes. Par contre, dès que la géométrie du patient devient trop spécifique ou qu'elle diffère trop de l'atlas, on ne peut jamais déformer correctement les images de l'atlas pour obtenir celles du nouveau patient. Ce principe est bien connu et illustré dans figure 26.

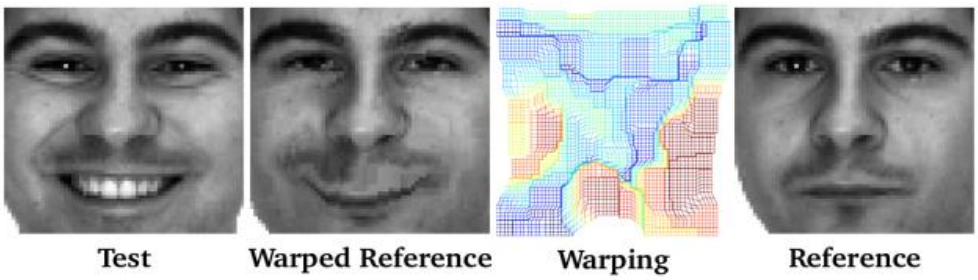


Figure 26 : Illustration que la méthode de recalage déformable ne fonctionne pas dès le moment qu'il y de l'information dans la nouvelle image (l'image « Test » dans la figure) qui n'est pas disponible dans les images de l'atlas (l'image « Reference » dans la figure).

C'est pour cette raison que nous avons implémenté une méthode hybride qui va combiner la méthode atlas tout en utilisant les niveaux de gris de l'image IRM du nouveau patient. L'idée de combiner ces deux informations est basée sur un article de Gudur et al [64], mais implémenté complètement différemment. Dans l'article introduit dans ce paragraphe, on compare notre nouvelle méthode avec la méthode classique utilisant juste la déformation avec un filtre gaussien pour diminuer l'impact du bruit. On démontre que la géométrie est beaucoup mieux reproduite, surtout les cavités d'air dans la région nasale et les os avec notre méthode. La différence entre les deux méthodes est que les intensités de l'IRM du nouveau patient sont utilisées afin de déterminer quels voxels peuvent être moyennés en appliquant le filtre sur l'image déformée. La définition de la taille de la matrice du filtre et un seuil arbitraire vont déterminer la pondération des deux informations dans la définition du pCT. Plus de détails sont disponibles dans l'article 4. Dans cet article, Raystation^{TM} de RaySearch Laboratories est utilisé pour déterminer les vecteurs de déformation.

3.2 Article 4

Dosimetric characterization of MRI-only treatment planning for brain tumors in atlas-based pseudo-CT images generated from standard T1-weighted MR images

Benjamin Demol[1,2,3], Christine Boydev[1], Juha Korhonen[4,5], Nick Reynaert[1]

[1] *Department of radiotherapy, Centre Oscar Lambret, Lille, France*

[2] *AQUILAB SAS, Loos Les Lille, France*

[3] *IEMN, UMR CNRS 8520, Villeneuve d'Ascq, France*

[4] *Department of Radiation Oncology, Comprehensive Cancer Center, Helsinki University Central Hospital, Helsinki, Finland*

[5] *Department of Radiology, Helsinki University Central Hospital, Helsinki, Finland*

Article soumis à *Medical Physics* le 08.06.16, accepté sous réserve de modifications le 03.08.16. Deuxième soumission le 30.08.16

Abstract. Purpose: MRI-only radiotherapy treatment planning requires accurate pseudo-CT (pCT) images for precise dose calculation. Current work introduces an atlas-based method combined with the usage of MR intensity information. pCT analysis and Monte Carlo dose calculations for intracranial stereotactic treatments were performed.

Methods: 22 patients, representing 35 tumor targets, were scanned using a 3D T_1 weighted MRI sequence according to the clinical protocol. The MR atlas image was registered to the MR patient image using a deformable algorithm, and the deformation was then applied to the atlas CT. Two methods were applied. The first method (MRdef) was based on deformations only while the second (MRint) also used the actual MR intensities. pCT analysis was performed using the mean (absolute) error , as well as an in-house tool based on a gamma index. Dose differences between pCT and true CT were analyzed using dose volume histogram (DVH) parameters, statistical tests, gamma index and probability density functions. An unusual case where the patient had been surgically operated (part of the skull bone was removed) was studied in detail.

Results: soft tissues presented a mean error inferior to 50 HUs while low-density tissues and bones presented discrepancies up to 600 HUs for hard bone. The MRdef method led to significant dose differences compared to the true CT (p-value < 0.05, Wilcoxon-signed-rank test). The MRint method performed better. The DVH parameter differences compared to CT were between -2.9% and 3.1% except for two cases, where tumors were located within the sphenoid bone. For these cases the dose errors were up to 6.6% and 5.4% (D98 and D95). Furthermore, for 85% of the tested patients the mean dose to the PTV was within 2% agreement when comparing to a calculation using the actual CT. Fictitious bone was generated in the unusual case when using atlas-based methods.

Conclusions: the atlas-based method led in general to acceptable dose distributions. Use of common T_1 sequences allows implementing the method into clinical routine. However unusual patient anatomy may produce large dose calculation errors. Detection of large anatomic discrepancies by using MR image subtraction can be realized, but an alternative way of producing synthetic CT numbers in these regions is still required

Keywords. MRI treatment planning, pseudo-CT, atlas-based method, hybrid method, T_1-weighted sequence

I. Introduction

Radiotherapy treatment planning (RTP) based on magnetic resonance imaging (MRI) only is identified as an advance in modern radiotherapy[1-2]. MRI offers superior and multiple tissue contrasts compared to CT and is often used for tumor delineation using a CT-MRI registration step. One of the benefits of MR-only RTP would be to suppress the introduced registration uncertainty of approximately 0.5-3.5 mm in the head region[3] and in the prostate[4]. The other benefits are non-exhaustively a reduction in terms of examination time and cost (CT no longer needed), non-exposure of x-ray radiation, and a coherent way of producing data when using an MRI-linac.

Previous studies have suggested several methods to generate pseudo-CT (pCT) images from one or more MR sequence data. Methods can be classified into two families. The first methods are intensity-based[5-7] generating pCT using a direct conversion of the intensity of the MR data. The second class consists of the so-called atlas-based methods[8-11] that synthesize pCT from a deformed reference CT atlas using the deformation between the MR patient image and the MR atlas. Both methods can include a prior segmentation. The intensity-based methods can provide confusing results because of the fact that no clearly bijective relation between MR and CT intensities has been established at the moment. In the pelvis, previous studies have shown possibility for conversion from MR intensities to CT intensities even with standard MRI sequence[5,12]. This approach has been adopted successfully also for other body parts such as abdomen and head[13-14]. However, the differentiation between air and bone intensities in MR images still remains difficult even with ultrashort echo time (UTE) sequences[15]. Atlas-based methods could provide a solution to separate bone and air even with standard MRI sequences. On the other hand, atlas-based methods suffer from registration errors due to the non-consistency between patient and atlas[16]. Post-processing steps after registration is a key step for increasing the accuracy of the pCT and remains an important field of investigation.

Bulk density methods have been reported to provide a good agreement in prostate cancer[17] but not for spinal bone metastasis for example[18]. For head and neck treatments, Chin *et al*.[19] stated that the bulk density method is feasible for intensity modulated radiotherapy (IMRT) treatments although having several limitations. Moreover, Korhonen *et al*.[20] showed dosimetric deviations of over 2% behind the bones in a pelvic phantom when using bulk densities. Dosimetric precision of MR-only RTP has been assessed in most of the previously introduced heterogeneous pCTs for conventional[6], IMRT[5,8] or volumetric modulated arc therapy (VMAT)[5,10-11] treatments. For conventional treatments, deviations of the median dose to the planning target volume (PTV) from a calculation based on pCT compared to a calculation of the patient CT go up to 3% depending on the selected

method[6]. IMRT treatments present deviations below 1% in prostate cancers[5] but large deviations up to 12.5 % in minimum dose has been reported in intracranial lesions[8]. Mean dose deviation of 0% (±0.2%, 1σ) and dose to half the volume of the PTV of 0.3% (±0.2%, 1σ) were reported for VMAT treatments in prostate cancer[5,10]. Most of the dose calculations were realized using an anisotropic analytical algorithm (AAA).

MR treatment planning would be of prime interest for Cyberknife stereotactic treatments to ensure a correct coverage of the target as delineated on the MR images. Some characteristics of a Cyberknife treatment plan, such as the high number of small beams with diameters ranging from 0.5 cm to 6 cm and the lack of posterior beams due to technical properties of the robot, make MR-only RTP challenging. All beams therefore come from anterior and lateral directions, passing with higher probability through bone and air cavity regions within the nasopharyngeal region compared to a conventional, IMRT, or VMAT treatment.

Current work aims to develop an intensity information-linked atlas-based set of two methods inspired by Gudur *et al*[11] following the same ideology but calculating CT numbers (nCT) differently. Our approach of the pCT generation is based on a more straightforward computation of CT numbers. Instead of estimating a probability density function inside a voxel we aimed to assign a weighted mean of neighboring voxels of the deformed CT atlas. These methods are applied in the head requiring a single conventional 3D T_1 weighted MR sequence dataset of the patient which is already included in the current workflow for target delineation. This study evaluated two pCT generation methods that rely on the atlas-based strategy. The first method used a conventional deformation-based algorithm, while the second method implicitly added the actual MR intensities, while generating the pCT, in order to find a compromise between the deformation and intensity based methods.

The pCTs were tested for Cyberknife stereotactic treatments on a set of 22 patients representing 35 targets in the head region. Moreover a Monte Carlo algorithm was used in order to get a high level of confidence in the planned dose. By applying the presented methods to extreme cases, the limits of current and all atlas-based methods in general are demonstrated.

II. Materials and Methods

A. CT and MR Imaging

22 patients were scanned with a 3D T_1 weighted sequence with injection of Gadolinium on a GE Signa MR750 3.0T (GE Healthcare, Milwaukee, WI) according to the clinical protocols. The presence of Gadolinium was not required, but in case of this retrospective study all data have been collected after injection. Patients were imaged in a diagnostic position without radiotherapy treatment positioning systems. TE was set to 2.3 ms and TR to 7 ms with a flip angle of 15°. The 2D image grid is composed of 512 x512 pixels and has an original resolution of 0.47 x 0.47 mm2. The number of slices along the axial direction is variable but is always superior to 300 slices spaced by 0.5 mm. This sequence was chosen as it is used to delineate tumors by the radio-oncologists in clinical routine. The data were re-sampled using Matlab in a 256 x 256 pixel grid and the number of slices was divided by a factor of two to get a voxel size of 0.94 x 0.94 x 1.0 mm3. This reduces the processing time to synthesize the pCT using voxel sizes similar to CT.

CT images were acquired by a Toshiba Aquilion 16LB at 120 kV (Tokyo, Japan), using 512 x 512 pixels of 0.68 x 0.68 mm2 size, and a slice thickness of 1 mm (required for Cyberknife (Accuray, Inc., Sunnyvale, CA) treatments), according to the clinical protocols. The data were re-sampled using a 256 x 256 pixel grid of 1.36 x 1.36 mm, without modifying the slice thickness.

B. Atlas generation

Two atlas-based methods were used in current work: 1) a conventional MR deformation based (MRdef) method and 2) a method combining MR intensity levels with deformation fields (MRint). The atlas, or the model as we refer in current paper, consists of a matched pair of MR and CT data selected from our patient list and was left out of the study. The model consists of an "average" patient, having an average bone density without extreme geometrical features or image artifacts. The model CT dataset was aligned with the model MR image frame using a rigid registration determined by MasterPlan (Nucletron B.V., Veenendaal, The Netherlands) using the mutual information algorithm. For each patient, the model MR was non-rigidly registered to the patient MR using the hybrid deformable algorithm[21] of RayStation software (RaySearch Laboratories, Stockholm, Sweden) and subsequently, the deformation was applied to the model CT using Matlab (The Math-Works, Inc., United States). After transformation, model MR and model CT were re-sampled within the MR patient frame and have consequently the same coordinate grid. Intensity dynamic adjustment was applied between model MR and patient MR to spread intensities into the same range. Mathematical operations on each voxel of the patient MR image were applied following one of the two methods described below. Figure 1 summarizes this atlas-based strategy.

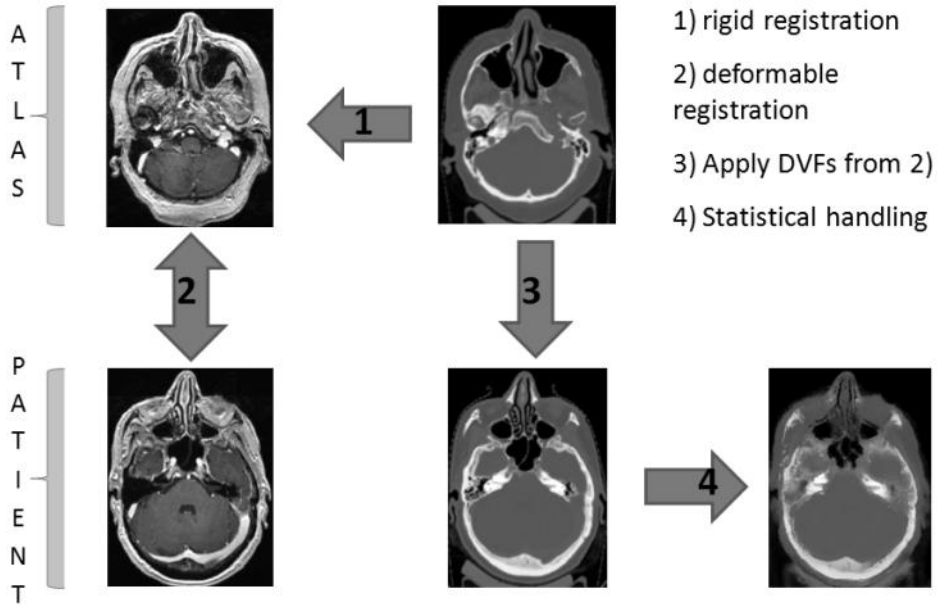


Figure 1: Overall strategy of the proposed atlas-based methods.

C. Generation of pCTs

The MRdef method does not take into account any information from the patient MR image and simply acts as a low pass filter. A 5x5x5 voxel box is centered on the voxel under consideration, and the corresponding model CT voxels inside the box are averaged using inverse-distance-weighting (the voxel under consideration has an introduced weight of twice the inverse distance of the nearest

voxel in order to have a coherent weighting). This method is thus purely based on deformable image registration.

The MRint method relies on a voxel intensity comparison between the patient MR image and the model MR image. An intensity search inside a 9x9x9 voxel box centered on each voxel of the patient MR image is performed using a 10% threshold to select a restricted group of voxels. The pCT value is calculated by averaging the corresponding deformed model CT numbers of the restricted group of voxels in an inverse-distance-weighted fashion (if the central voxel under consideration is selected, its weight was defined as twice the inverse distance to the nearest voxel). During this step, if the restricted group of voxels is empty, a global intensity search is performed on the entire slice using the same threshold and the first five nearest voxels are averaged without weighting. This specific selection of voxels is introduced to force applying the actual MR voxel intensities. The MRint method is schematically described in figure 2.

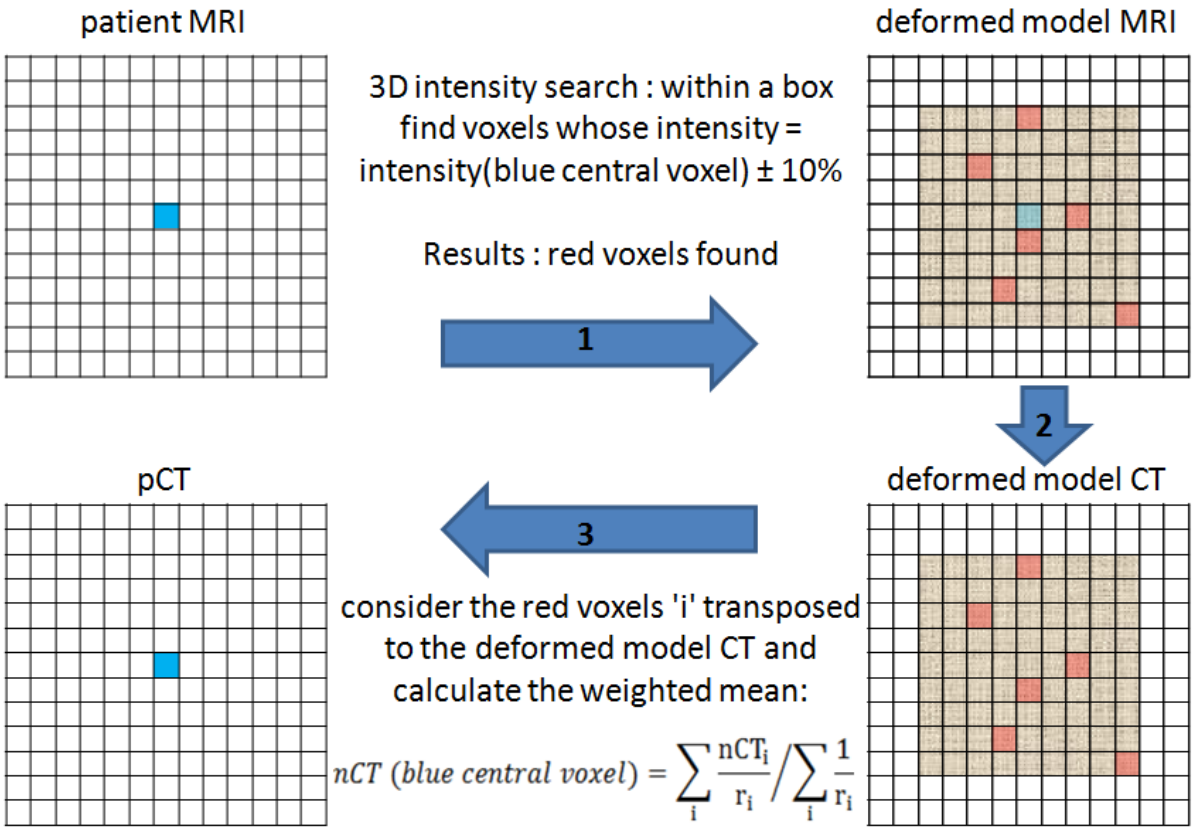


Figure 2: Simplified scheme of pCT generation of the MRint method. nCT is the CT number and *r* the distance of a voxel to blue central voxel. If the central voxel is selected, its weighting is twice the inverse distance to the nearest voxel. If no voxel is found in the box, the intensity search is performed on the entire slice and the CT numbers of the five nearest voxels are averaged.

D. pCT accuracy

A semi-quantitative analysis under the form of a revised gamma index was performed on the pCT for several cases to assess overall reliability of this method. This in-house tool is able to show where discrepancies between pCT and true CT occur. The gamma index is formulated as:

$$gammaCT = min\left(\sqrt{\frac{\Delta r^2}{\Delta r_{max}{}^2} + \frac{\Delta nCT^2}{\Delta nCT_{max}{}^2}}\right)$$

where *ΔnCT* is the difference of nCT between the voxel under consideration in the pCT and a voxel of the true CT at the distance *Δr* from the voxel under consideration. Δr_{max} and ΔnCT_{max} are the imposed criteria representing the superior limit of success of the test. They were set to 2 mm and 50 HU respectively.

The mean absolute error (MAE) between voxels of the pCT and voxels of the true CT was calculated as well to conform with previous publications such as in [10,22-23]. Images were previously re-sampled with a low-pass 27-voxels neighborhood filter to decrease the influence of the noise. A graphical representation of the MAE was used in the conventional range of CT numbers by collecting groups of 20 HUs as in [22]. The mean error (ME) was also calculated using the same settings. MAE and ME are defined by the following formulas:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|nCT_{pCT}(i) - nCT_{tCT}(i)\right|$$

$$ME = \frac{1}{N}\sum_{i=1}^{N} nCT_{pCT}(i) - nCT_{tCT}(i)$$

where *N* is the total number of voxels, *i* the voxel index, nCT_{pCT} the CT number of the pCT, and nCT_{tCT} the CT number of the true CT.

E. Dose analysis

The feasibility of pCTs for dose calculation was assessed in terms of dose to target volume compared to the true CT. A retrospective study was conducted on a dataset composed of 22 cases representing 35 tumor sites (10 patients with two targets, 2 patients with 3 targets) localized in the head region. MR data, CT data, structures and Cyberknife treatment plans were used to produce pCTs and duplicate the plans on pCTs. Targets were composed of 29 cerebral tumors located in the inner brain, close to cranial bone, in the cerebellum, or near the eyes, including 7 targets treated with beams passing through bone and air cavity regions of the nasopharynx anatomy. The remaining cases consist of 2 tumors covering the sphenoid bone, of 1 skin melanoma and of 3 choroid tumors. pCTs were registered within the true CT frame using the patient MR image to patient CT image registration and a dose calculation was performed on the Moderato Monte Carlo platform[24] using both the true CT and the pCTs. No pCT-to-true-CT registration was used to avoid introducing a bias in the results. Statistical uncertainties inside dose voxels were below 1% (2σ) in the Monte Carlo simulations.

The study evaluated dose volume histograms (DVHs) as these are routinely used for optimization and evaluation of treatment plans. Mean dose D_{mean} and several DVH parameters D_X concerning the PTV were used, indicating the dose to x% of the volume: D_{98}, D_{95}, D_{50}, D_{05} and D_{02}. For each PTV, these parameters were measured on the true CT and the two pCTs. Relative dose differences between the true CT and each pCT were calculated and an empiric density distribution was plotted for each index. The mean of the sample of the 35 tumor sites was calculated and the mean's uncertainty was determined assuming that the mean of the sample followed a normal law when the sample size is superior to 30 (central-limit theorem). From these results and following the recommendations of

Chaikh *et al*.[25], the normality test of Shapiro-Wilk was applied to each index for each pCT – true CT pair to evaluate the normality of the global population. In case many p-values were below 0.05, a non-parametric test was used instead for all DVH parameters. The Wilcoxon signed rank test was performed on each index using the null hypothesis as the global population has a mean of zero. All statistical tests were realized using R[26].

The best of the two methods based on the result of the Wilcoxon test and on the statistical analysis of the samples was selected and an estimation of the probability density function of each index was performed using the nonparametric kernel density estimation method[27] in Matlab. In this way we were able to assess the percentage of patients receiving MRI-only treatments with a 2% or 3% dose agreement in PTV compared to the true CT using cumulative density functions, indicating the reliability of the applied pCT methods.

The Gamma index, defined by

$$\gamma = min\left(\sqrt{\frac{\Delta r^2}{\Delta r_{max}{}^2} + \frac{\Delta D^2}{\Delta D_{max}{}^2}}\right)$$

was also computed to compare dose discrepancies against nCT discrepancies, where *ΔD* is the difference of dose between the voxel under consideration in the pCT and a voxel of the true CT at the distance *Δr* from the voxel under consideration. Δr_{max} and ΔD_{max} are the imposed criteria representing the superior limit of success of the test. They were set respectively to 2 mm and 2% of the mean dose to the PTV.

F. Limitations of atlas-based methods

Atlas-based methods are well known to be sensitive to differences between atlas and new patients. In order to evaluate how our methods handle this problem a specific case, namely a patient who has been surgically operated, was studied more intensely. In this case, a part of the skull bone was removed (see figures 9 and 10). In such a case, an atlas-based method is expected to produce large discrepancies in this area because of the non-consistency between patient and model images. For this specific case, the pCT was generated and evaluated in detail. In order to isolate this type of atypical area and possibly use the patient MRI intensity levels directly, a voxel-to-voxel relative ratio between the patient MR image and the deformed model MR image was used.

III. Results

A. pCT analysis

Figure 3 presents the overall appearance of the two pCTs. Figure 4 shows an example of a gammaCT image in a challenging anatomical region including air cavities and bone regions, for both methods.

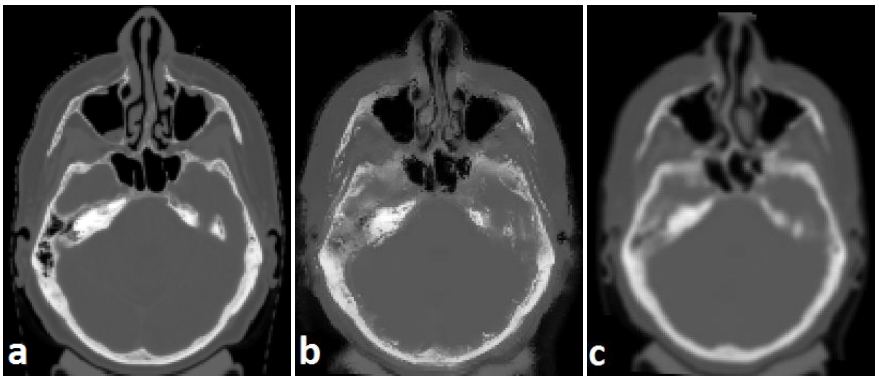


Figure 3: Examples of the generated pCTs compared to the corresponding CT slices in a). b) presents pCT by the MRint method and c) presents pCT by the MRdef method.

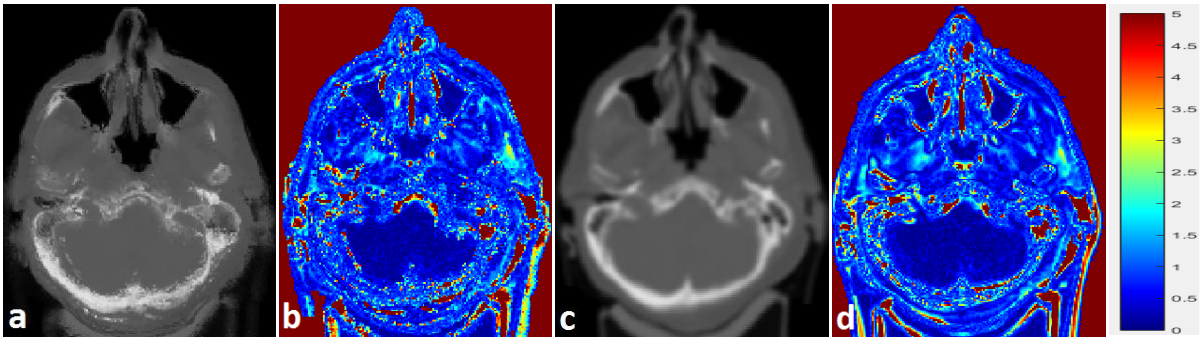


Figure 4: pCTs and their associated gammaCT calculated with the true CT as reference. a) pCT from the MRint method; b) gammaCT of a); c) pCT from the MRdef method; d) gammaCT of c). In the air outside the internal contour of the head, the value of gammaCT (red color) has no meaning because a nCT value of -2048 HU was artificially introduced, which is considered as air in the dose calculation. The color map of the gammaCT is always the same through the entire paper.

The gammaCT of the MRint method is more anarchical but this is due to the original appearance of the pCT. Homogeneous areas are more easily reproducible than heterogeneous areas while generating the pCT. More quantitative results are needed for selecting the optimal method though.

Figure 5 shows the MAE averaged on all patients for both methods using the external contour to restrict data to patient voxels only. For both methods, soft tissue characterized by CT numbers between approximately -100 and +100 HUs are better assessed than air-tissue mixtures and bones. Cerebral tissues including white and gray matter (nCT ≈ 30 HUs), representing the major proportion of cranial tissues, present a MAE of 50 HUs. Low density tissues (nCT < -100), representing soft tissues suffering from partial volume effect, show deviations between 100 HUs and 650 HUs. The peak near -900 HUs is probably due to the partial volume effect for voxels located at the skin. Low and medium density bone (nCT from 100 to 1200) present an almost invariable MAE of 280 HUs, while high density bones (nCT > 1200 HUs) are more difficult to quantify correctly with discrepancies of almost 650 HUs at 1800 HUs for the MRdef method, and 550 HUs for the MRint method. The MRint method deals better with high density bones but gives rise to more discrepancies within the [-600 HUs, -200 HUs] and [800 HUs, 1200 HUs] ranges compared to the MRdef method.

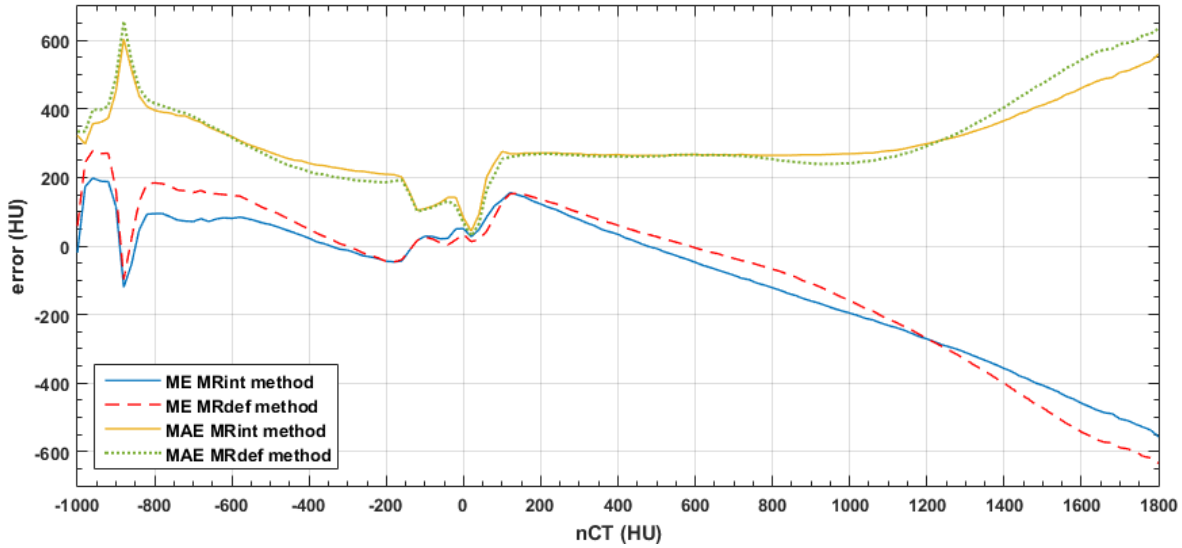


Figure 5: mean error and mean absolute error, both calculated in bins of 20 HUs and averaged across all patients for the two methods.

The mean error (ME) is shown using the same voxel selection procedure as described for the MAE. CT numbers of the MRint method are closer to reality for the range [-1000;-200] and for the range above 1200 HUs than the MRdef method. Both methods provide similar results between these intervals and a mean error inferior to 50 HUs for soft tissues.

B. Dose analysis

Table 1 shows the series of statistical parameters regarding the relative dose differences for different DVH parameters for both pCT methods.

For almost all DVH parameters, the MRint method has a smaller data range except for D_{05} and D_{02} regarding the minimum value of the dataset. The mean of dose differences for the MRint method is closer to zero compared to that of the MRdef method. As the standard deviation of this mean is approximately the same for the two methods, the mean of the global population is closer to zero for the MRint method, which is confirmed by the analysis of the p-value of the Wilcoxon signed-rank test. As the p-value is inferior to 0.05 for all DVH parameters except D_{02} in the MRdef method, we can say that this method provides significant dose differences compared to the true CT. This is not the case for the MRint method as all p-values are superior to the 0.05 threshold. Following these results, the MRint method was considered as providing the best results and a deeper statistical analysis was conducted only for this method.

	MRint method		MRdef method	
DVH parameter	mean ± stdev [min;max]	p-value	mean ± stdev [min;max]	p-value
D_{98}	-0.1 ± 0.35 [-2.9;6.6]	0.24	-0.5 ± 0.38 [-4.7;6.6]	0.02
D_{95}	-0.1 ± 0.30 [-2.5;5.4]	0.22	-0.6 ± 0.33 [-3.6;6.0]	0.01
D_{50}	-0.4 ± 0.21 [-2.3;2.1]	0.09	-0.7 ± 0.22 [-2.7;2.6]	0.00
D_{05}	-0.3 ± 0.22 [-2.4;2.5]	0.13	-0.5 ± 0.22 [-2.7;2.1]	0.04
D_{02}	-0.3 ± 0.24 [-2.4;3.1]	0.30	-0.4 ± 0.23 [-3.2;2.2]	0.20
D_{mean}	-0.4 ± 0.21 [-2.0;2.3]	0.14	-0.7 ± 0.21 [-2.7;2.6]	0.00

Table 1 : Statistical indicators of relative dose differences of different DVH parameters calculated from the dataset of 35 tumor sites. Mean, min, max are respectively the mean value, the lowest value, and the highest value across the dataset. stdev is the standard deviation of the mean across the dataset calculated under the assumption of the central limit theorem. p-value is the p-value resulting from the Wilcoxon signed-rank test.

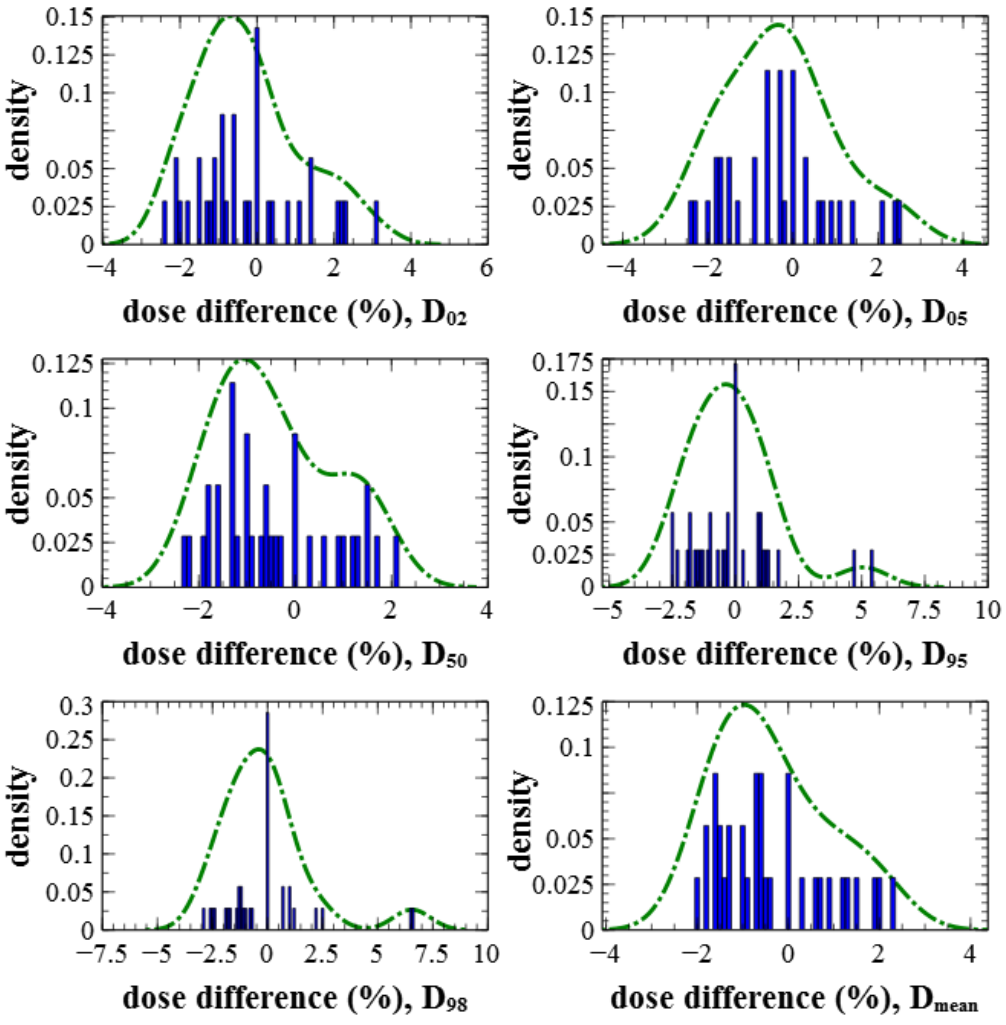


Figure 6: Density distribution of the dose difference between the pCT from the MRint method and the true CT for several DVH parameters. Bars (blue): data distribution; Line (green): probability density function based on these data. The height of the green curve was reduced for clarity. The width of the bars is equal to 0.1%.

The empirical distribution and the estimated density function of all DVH parameters of the MRint method are shown in figure 6.

The estimated probability density function indicates that for D_{50}, D_{05}, D_{02}, and D_{mean}, the absolute dose difference of the global population is less than 4%. For the D_{98} and D_{95} indices, there are two empirical values which pull out the curve from this range. These two values are associated with a target localized within a sphenoid bone area. Figure 7 displays the anatomical region, the dose distribution into a slice, and the corresponding DVH for one of these two cases. A dose profile is plotted as an example to complete the observation.

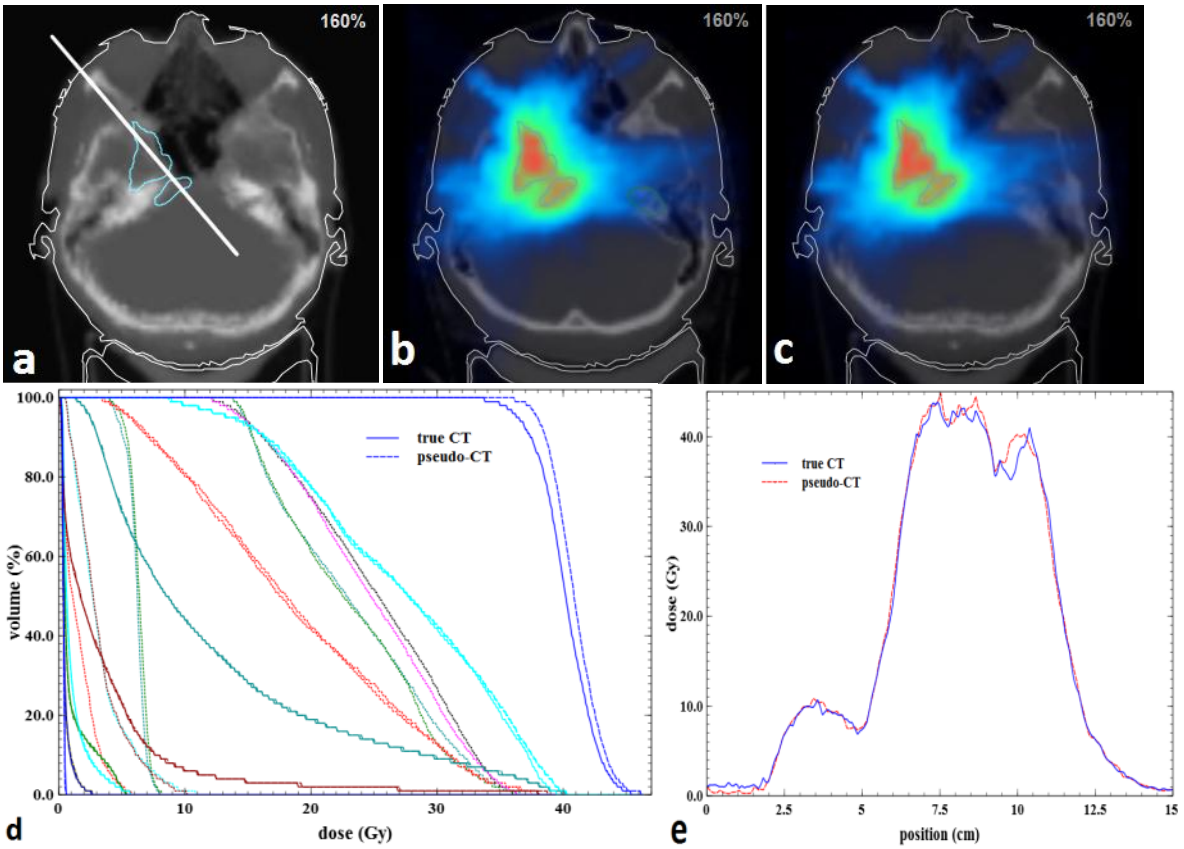


Figure 7: Sphenoid case where target is represented by the blue contour on image a) (true CT). b) and c) are respectively the dose distribution in the true CT and in the pCT. The DVH of this case is plotted in d) where the PTV curves are in dark blue at the most right of the window. Other curves concern organs at risk. A profile following the white line starting from the upper left corner to bottom right in a) is plotted in e).

The target is clearly localized in a challenging region encompassing soft tissues, bone and air. Furthermore, beams pass through this diversity of tissues. The 2D dose distribution presents no noticeably discrepancies between the dose in pCT and true CT, but the dose profile highlights some differences, in particular near the smallest part of the PTV. Discrepancies may be the consequence of an incorrect assignment of CT numbers in the pCT in the area of the target or on the beam path, which are revealed by the gammaCT in figure 8.b).

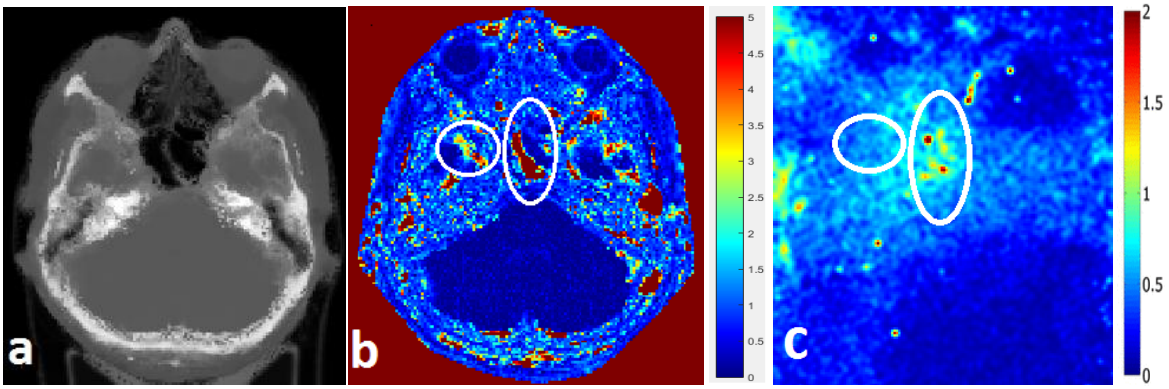


Figure 8: a) pCT, b) gammaCT, and c) gamma index of the slice selected in figure 7. Two misleading areas are white circled on the gammaCT image and copied on the gamma index image.

An important inconsistency can be visualized within the main air cavity, where some artificial soft tissue has been introduced in the pCT. As several beams cross this region, attenuation is modified and more secondary electrons reach the target at the edge of the air cavity, inducing a higher dose around this region. This can, at least partially, explain the results of the DVH and of the dose profile. Other differences of CT numbers are revealed including in the target area and can affect the global result. However, this kind of disagreement seems to affect more the minimum dose in the target, as D_{98} and D_{95} reveal in figures 6 and 7.d), as other DVH parameters are in majority below 2% (except for Dmean which equals 2.3% for the other sphenoid case). The Gamma index distribution of the same slice is shown in figure 8.c). Discrepancies between the gamma index and the gammaCT present some agreements within some regions, but no clear connection was noticed on the entire volume.

Results of cumulative density function are presented in table 2.

	method	D_{98}	D_{95}	D_{50}	D_{05}	D_{02}	D_{mean}
[-2% ; 2%]	MRint	73.6%	75.4%	87.2%	81.3%	82.7%	84.8%
	MRdef	69.1%	71.5%	80.4%	80.5%	81.1%	83.8%
[-3% ; 3%]	MRint	88.9%	90.4%	99.3%	96.9%	97.2%	98.4%
	MRdef	85.2%	86.2%	97.4%	96.9%	96.8%	98.1%

Table 2: Estimation of the percentage of patients within the global population passing 2% and 3% dose agreement criteria (DVH parameters) in pCT compared to CT.

Although the MRdef method was dismissed by the Wilcoxon test, its estimations are displayed to reinforce the better results obtained using the MRint method which indeed surpasses the MRdef method for all DVH parameters. Agreement is less for D_{98} and D_{95} for the reasons already mentioned. To illustrate the results, the mean dose received by the PTV would be planned within a 2% agreement using the MRint method for 84.8% of cases.

C. Limitations of atlas-based methods

As shown in figure 9, large geometrical differences between atlas and new patient lead to the appearance of fictitious bone in the pCT, although the patient MR image presents no bone (see figure 10.a). This virtual bone is generated due to the existence of similar MR values near the area of concern in the deformed model.

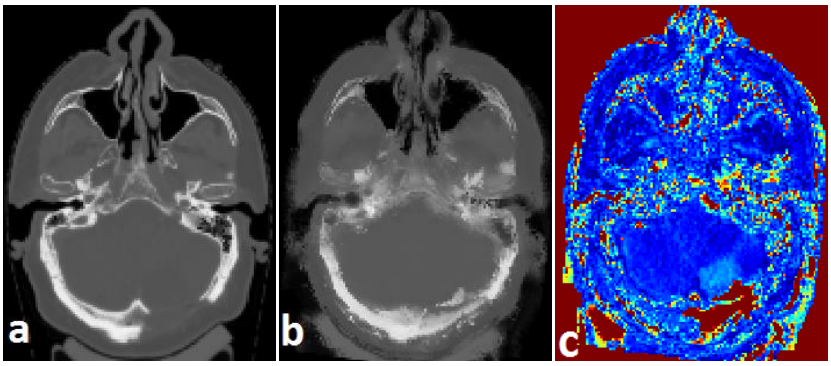


Figure 9: Set of images showing a patient with atypical anatomy. a) true CT; b) pCT from the MRint method; c) gammaCT between b) and a)

Figure 10.c) shows that it is not straightforward to determine a unique threshold through the entire image in order to delimit regions of high discrepancies. Some regions appearing in red are also targeted by this operation, so the technique should be adaptive in function of localization (context-based). Moreover, bone thickness is not representative on these T_1 weighted images as we can observe the extremely thin dark line in the first row of images which corresponds approximately with the localization of the slice in figure 9, where the bone thickness is shown to be larger. Discrepancies were however detected on further slices, but the bone width was still inferior to the real value. An alternative value could be set for the isolated regions for example by using a global relation between MR intensity and CT number. However, as no clear relation could be defined because of the spread of soft tissues on the entire range of MR intensity, and the non-distinction between bone and air regions, this correction method was not applicable when using conventional MRI sequences.

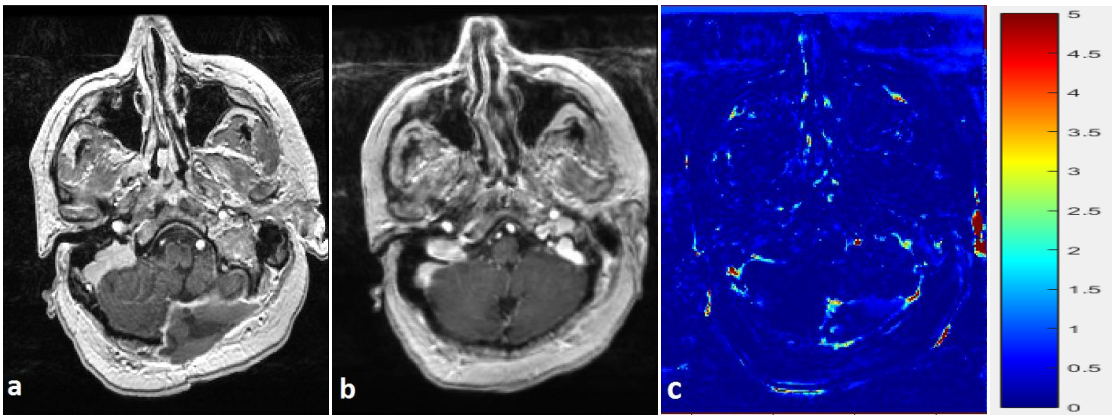


Figure 10: Set of images showing a patient with atypical anatomy. a) patient MR image; b) deformed model MR image; c) absolute relative ratio of the two former.

From a dosimetric point of view, as the tumor was outside this anatomical particularity location, the case was excluded from the dataset. Nevertheless, dose calculation was similar between the true CT and the pCT as no beam entered from this side because of the Cyberknife configuration (no beams coming from the back).

IV. Discussion

As illustrated in figure 3, the pCT of the MRint method contains a lot of noise near bones and air cavities, whereas pCT of the MRdef method clearly introduces blurring and geometry errors. The geometry is accurately reproduced in case of the MRint method regarding air cavities and bony

regions, illustrating the importance of taking into account patient MRI intensities. These patient specific regions can currently never be reconstructed correctly, when only using atlas-based deformations fields. Moreover, these air cavities have a large impact on the particle transport for small beams traversing these regions, which is, as explained above, happening regularly when using Cyberknife. Thus the impact on the final dose distribution can be relevant. On the other hand, for these conventional MRI sequences, one cannot rely completely on patient MR intensity levels because bone and air cannot be distinguished. Combining deformation fields and MRI intensity levels seems to provide a very good compromise for the majority of cases handled in current work. The current limitation of our method combining MRI intensity levels with deformation fields is that for approximately 10% of patient MRI voxels, no surrounding voxels with corresponding grey value (within 10 % as defined in the MRint method) were found. For these voxels we tried to dynamically increase the search region around the voxel of interest. This method can be considered as an MRI-to-CT calibration method using only the local context. However, this method regularly introduced incoherent CT numbers, such as negative values for a voxel in dense bone. Therefore, this method was replaced by using the CT number of the deformed model CT instead. All occurrences in air were due to noise, and did not introduce any bias in the dosimetric results, as those voxels were still considered as air. Occurrences into soft tissue should not have introduced important differences in terms of CT number either, as cerebral tissues are relatively homogeneous. As the substitution by the CT number of the CT model occurred only a limited number of times within the bone region, no bias was expected in general.

In general, as illustrated by the gammaCT, soft tissue regions are well quantified by the two methods, just as large air areas in the oral cavity. CT numbers of the skin are pretty well reproduced except near the ears and the posterior part of the head due to deformable registration errors. Some differences are due to the contention material (pillow and mask) which is not in the same configuration between the model image and the patient image. The remaining differences are caused by the generation of the pCT itself. More important differences are visible in small or linear air cavities and for small or mixed-composition bones.

As confirmed by the MAE and ME results, soft tissues are well characterized by the two methods, whereas bone still remains difficult to synthesize. Bone density is underestimated by about 500 HUs, and the divergence increases with higher densities. Large MAE values up to 600 HUs were obtained. It is stressed that this comparison of voxel intensities relies on a rigid registration which is not exempt from error. So the voxel-to-voxel comparison is not the optimal way to characterize a method but can assess an overall trend. Results are similar to Andreasen[(22)] except for CT numbers inferior to -200 HUs where the results don't match. The reason might be a different handling of voxels outside the patient or a real difference originating from the methods. Moreover, the absolute error does not indicate its distribution. Values of ME lower than MAE in the range [-1000;-100] demonstrate that air-tissue mixture CT numbers present high variance with a lot of negative and positive values compared to the true CT. This observation can be applied to a lesser extent up to 1200 HU. The use in conjunction of MAE and ME allows a general evaluation of a pCT but does not set single-handedly the reliability of the pCT generation method.

Dosimetric results clearly demonstrated that the MRint method reproduces the patient anatomy more accurately and also confirmed that the MRdef method depends too much on the model anatomy and has difficulties dealing with air cavities, as shown in figure 3. The dosimetric precision

of the MRint method is well characterized by table 2. In this table the number of cases, falling into a threshold criterion of 2% or 3% is reported, indicating a certain confidence level when implementing the method in clinical practice. If we focus on the mean dose D_{mean}, results indicate that near 85% of patients having an MR-only treatment will receive a similar mean dose (±2%) compared to a conventional CT-based treatment, and 98.4% of patients at ±3%. Consequently, 1.6% of pCT patients will receive an absolute dose difference superior to 3%. One often considers the threshold of 2% to be the highest incertitude[28] on the planned dose. Here this threshold is not respected in about 20% of cases when combining all DVH parameters, but it has to be emphasized that dose will be effectively planned to the localization of the target because the main advantage of the MR-only treatment is the removal of the registration error between MRI and CT: small dose discrepancies (with known limits) are obtained, but the target is correctly covered, which is not always the case with MR-contour CT-based treatments. Although registration errors can occur during the deformable registration step of this method, these errors will lead solely to an incorrect CT number and not to an incorrect targeting. The main advantages of this method remain that classical T_1 weighted sequence are used which are already available in clinical practice, and that no segmentation is required, so no time is lost to produce or verify additional contours not required for the treatment prescription. This method could easily be set up in clinical routine. Use of a multi-atlas could also enhance the registration process and is currently under development.

Some cases are nevertheless more challenging such as the case for the sphenoid target. For such a case, the method shows its limits. The difference in the target DVH of figure 7.d) is relatively important. Furthermore, for very particular cases where the patient contains some anatomic anomaly like missing tissue due to surgery for instance, the proposed method can potentially provide high divergences compared to a CT-based treatment. This is a characteristic of the atlas-based method in general which is currently limited to some spatial constraints. Bone is artificially created by the deformation vector field of the atlas. This problem cannot be solved even when using a multi-atlas where some specific situations could have been added to the atlas database. For the current case, detection was in particular limited by the excessively high diversified tissue contrast and the underestimation of the bone width due to the T_1 weight. This difficulty could be resolved by using UTE sequences within our method for example, where bone width has been demonstrated to be assessed correctly[29]. The excessively high tissue contrast is also most probably at the origin of the failure of the defined alternative way of producing restricted data when no similar voxels are found in the bounding box.

A way to deal with unusual cases would be to combine current method with an error detection engine such as the gammaCT tool applied to MR images and an alternative way to produce pCT numbers in regions where the main method is detected to fail. For instance we could imagine reproducing this method with zero time echo (ZTE) sequence images, analyzing the intensity differences between patient MR image and the deformed model MR image to detect large areas of deviation coming from a hypothetic anatomy anomaly such as crane holes, and applying an alternative method such as a direct conversion from MR intensity to CT number. Wiesinger *et al*[30] have demonstrated the potential linearity between ZTE intensity and CT number. Although we have reproduced and tested the atlas-based method for some cases using this type of images with good agreement, we were unfortunately not able to observe the same relation in our ZTE images.

This argument could also be integrated to the core of the method to propose a hybrid method. For all points of patient MR image where no similar values could be found in neighboring voxels, such an intensity conversion could be applied *on-the-fly* with ZTE images. Not finding any voxels near the voxel of interest would justify using an empiric relation, even if this relation is not robust (once a good correlation is obtained between MR intensity and CT number the atlas methods are no longer needed). It could help to assign a more reliable CT number where anatomy between patient and reference has some discrepancies not necessarily as large as with the error detection engine.

V. Conclusion

This study presented a new atlas-based method to produce pCT from MR data using a simple conversion model dealing with MR intensity. The MRint method combined deformation fields and the MRI intensities. MAE and ME indicated that soft tissues were well represented while low-density tissues and bones presented discrepancies up to 600 HUs for hard bone. The MRint method did not provide significant dose differences compared to the true CT as all p-values were superior to 0.05 using a Wilcoxon-signed-rank test. The DVH parameter differences compared to true CT were between -2.9% and 3.1% except for two cases, where tumors were located within the sphenoid bone. For these cases the dose errors went up to 6.6% and 5.4% (D_{98} and D_{95}). Use of cumulative density functions set the percentage of patients that will receive a mean dose to the PTV in a 2% and 3% agreement with true CT using this method to 84.8% and 98.4%. Fictitious bone was generated in the unusual case of missing bone when using our atlas-based method. A hybrid method was designed to detect and to deal with this kind of particular anatomy, but correlation between MR intensities and CT numbers has not been achieved yet. Use of common T_1 sequences allows implementing straightforwardly the method in clinical routine. Nevertheless, unusual anatomy in patient data may result in low-quality pCT and thus in unacceptable dose calculation errors.

References

1. M. A. Schmidt and G. S. Payne, “Radiotherapy planning using MRI,” Phys. Med. Biol., vol. 60, no. 22, pp. R323–R361, 2015.

2. J. J. W. Lagendijk, B. W. Raaymakers, C. a T. Van den Berg, M. a Moerland, M. E. Philippens, and M. van Vulpen, “MR guidance in radiotherapy,” Phys. Med. Biol., vol. 59, no. 21, pp. R349–R369, Nov. 2014.

3. K. Ulin, M. M. Urie, and J. M. Cherlow, “Results of a multi-institutional benchmark test for cranial CT/MR image registration.,” Int. J. Radiat. Oncol. Biol. Phys., vol. 77, no. 5, pp. 1584–9, Aug. 2010.

4. T. Nyholm, M. Nyberg, M. G. Karlsson, and M. Karlsson, “Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments.,” Radiat. Oncol., vol. 4, no. 1, p. 54, Jan. 2009.

5. J. Korhonen, M. Kapanen, J. Keyriläinen, T. Seppälä, and M. Tenhunen, “A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer.,” Med. Phys., vol. 41, no. 1, p. 011704, Jan. 2014.

6. J. M. Edmund, H. M. Kjer, K. Van Leemput, R. H. Hansen, J. A. Andersen, and D. Andreasen, “A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times.,” Phys. Med. Biol., vol. 59, no. 23, pp. 7501–19, Dec. 2014.

7. A. Johansson, M. Karlsson, and T. Nyholm, “CT substitute derived from MRI sequences with ultrashort echo time,” Med. Phys., vol. 38, no. 5, p. 2708, 2011.

8. T. Stanescu, H.-S. Jans, N. Pervez, P. Stavrev, and B. G. Fallone, “A study on the magnetic resonance imaging (MRI)-based radiation treatment planning of intracranial lesions.,” Phys. Med. Biol., vol. 53, no. 13, pp. 3579–93, Jul. 2008.

9. J. Dowling, J. Lambert, J. Parker, O. Salvado, J. Fripp, A. Capp, C. Wratten, J. W. Denham, and P. B. Greer, “An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy,” Int. J. Radiat. Oncol. Biol. Phys., vol. 83, no. 1, pp. e5–e11, 2012.

10. C. Siversson, F. Nordström, T. Nilsson, T. Nyholm, J. Jonsson, L. E. Olsson, and T. Nilsson, “Technical Note : MRI only prostate radiotherapy planning using the statistical decomposition algorithm,” Med. Phys., vol. 42, no. 10, pp. 6090–6096, 2015.

11. M. S. R. Gudur, W. Hara, Q.-T. Le, L. Wang, L. Xing, and R. Li, “A unifying probabilistic Bayesian approach to derive electron density from MRI for radiation therapy treatment planning,” Phys. Med. Biol., vol. 59, no. 21, pp. 6595–6606, Nov. 2014.

12. M. Kapanen and M. Tenhunen, “T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning.,” Acta Oncol., vol. 52, no. 3, pp. 612–618, Apr. 2013.

13. L. Koivula, L. Wee, and J. Korhonen, “Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: Dose calculation accuracy in substitute CT images,” Med. Phys., vol. 43, no. 8, pp. 4634–4642, 2016.

14. J. Korhonen, L. Koivula, T. Seppälä, M. Kapanen, and M. Tenhunen, “PO-0912: MRI-only based RT: adopting HU conversion technique for pseudo-CT construction in various body parts,” Radiother. Oncol., vol. 119, p. S440, 2016.

15. S. Hsu, Y. Cao, K. Huang, M. Feng, and J. M. Balter, “Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy,” Phys. Med. Biol., vol. 58, pp. 8419–8435, 2013.

16. M. Holden, “Nonrigid Body Registration OF PA VS WM,” IEEE Trans. Med. Imaging, vol. 27, no. 1, pp. 111–128, 2008.

17. D. Pasquier, N. Betrouni, M. Vermandel, T. Lacornerie, E. Lartigau, and J. Rousseau, “MRI alone simulation for conformal radiation therapy of prostate cancer: technical aspects.,” Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 1, pp. 160–3, Jan. 2006.

18. S. J. Hoogcarspel, J. M. Van der Velden, J. J. W. Lagendijk, M. van Vulpen, and B. W. Raaymakers, “The feasibility of utilizing pseudo CT-data for online MRI based treatment plan adaptation for a stereotactic radiotherapy treatment of spinal bone metastases.,” Phys. Med. Biol., vol. 59, no. 23, pp. 7383–7391, Nov. 2014.

19. A. L. Chin, A. Lin, S. Anamalayil, and B. K. Teo, “Feasibility and limitations of bulk density assignment in MRI for head and neck IMRT treatment planning,” J. Appl. Clin. Med. Phys., vol. 15, no. 5, pp. 100–111, 2014.

20. J. Korhonen, M. Kapanen, J. Keyrilainen, T. Seppala, L. Tuomikoski, and M. Tenhunen, “Absorbed doses behind bones with MR image-based dose calculations for radiotherapy treatment planning.,” Med. Phys., vol. 40, no. 1, p. 011701, 2013.

21. O. Weistrand and S. Svensson, “The ANACONDA algorithm for deformable image registration in radiotherapy.,” Med. Phys., vol. 42, no. 1, pp. 40–53, 2015.

22. D. Andreasen, K. Van Leemput, R. H. Hansen, J. A. L. Andersen, and J. M. Edmund, “Patch-based generation of a pseudo CT from conventional MRI sequences for MRI- only radiotherapy of the brain Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain,” Med. Phys., vol. 42, no. 4, pp. 1596–1606, 2015.

23. C. M. Rank, C. Tremmel, N. Hünemohr, A. M. Nagel, O. Jäkel, and S. Greilich, “MRI-based treatment plan simulation and adaptation for ion radiotherapy using a classification-based approach.,” Radiat. Oncol., vol. 8, p. 51, Jan. 2013.

24. B. Demol, R. Viard, and N. Reynaert, “Monte Carlo calculation based on hydrogen composition of the tissue for MV photon radiotherapy,” J. Appl. Clin. Med. Phys., vol. 16, no. 5, pp. 117–130, 2015.

25. A. Chaikh, J.-Y. Giraud, E. Perrin, J.-P. Bresciani, and J. Balosso, “The choice of statistical methods for comparisons of dosimetric data in radiotherapy.,” Radiat. Oncol., vol. 9, no. 1, p. 205, Jan. 2014.

26. R. Ihaka and R. Gentleman, “R: A Language for Data Analysis and Graphics,” J. Comput. Graph. Stat., vol. 5, no. 3, pp. 299–314, Feb. 1996.

27. A. W. Bowman and A. Azzalini, Applied Smoothing Techniques for Data Analysis : The Kernel Approach with S-Plus Illustrations: The Kernel Approach with S-Plus Illustrations. Oxford: Oxford University Press, 1997.

28. A. Ahnesjö and M. M. Aspradakis, “Dose calculations for external photon beams in radiotherapy,” Phys. Med. Biol., vol. 44, no. 11, pp. R99–R155, Nov. 1999.

29. L. Wang, X. Zhong, L. Zhang, D. Tiwari, and H. Mao, “Ultra-short TE (UTE) Imaging of Skull and a Quantitative Comparison of Skull Images Obtained from MRI and CT,” Proc. Intl. Soc. Mag. Reson. Med., vol. 18, p. 796, 2010.

30. F. Wiesinger, L. I. Sacolick, A. Menini, S. S. Kaushik, S. Ahn, P. Veit-haibach, G. Delso, and D. D. Shanbhag, “Zero TE MR Bone Imaging in the Head,” Magn. Reson. Med., vol. 75, no. 1, pp. 101–114, 2016.

4. Comparaison des recalages déformables de RayStation avec ITK/Elastix

Comme évoqué plus haut, les vecteurs de déformation utilisés en combinaison avec l'atlas sont déterminés en utilisant RayStation (RaySearch Laboratories). De plus, les recalages rigides entre le CT et l'IRM de l'atlas sont faites avec Masterplan (Nucletron). Le fait de dépendre de deux systèmes commerciaux n'est pas très attractif. Pour cette raison nous avons démarré une étude de comparaison entre les images pCT obtenues en utilisant les deux algorithmes commerciaux et celles obtenus avec ITK/Elastix. Cette étude est faite par un stagiaire de deuxième année de la formation DQPRM. Les résultats sont montrés en figure 27.

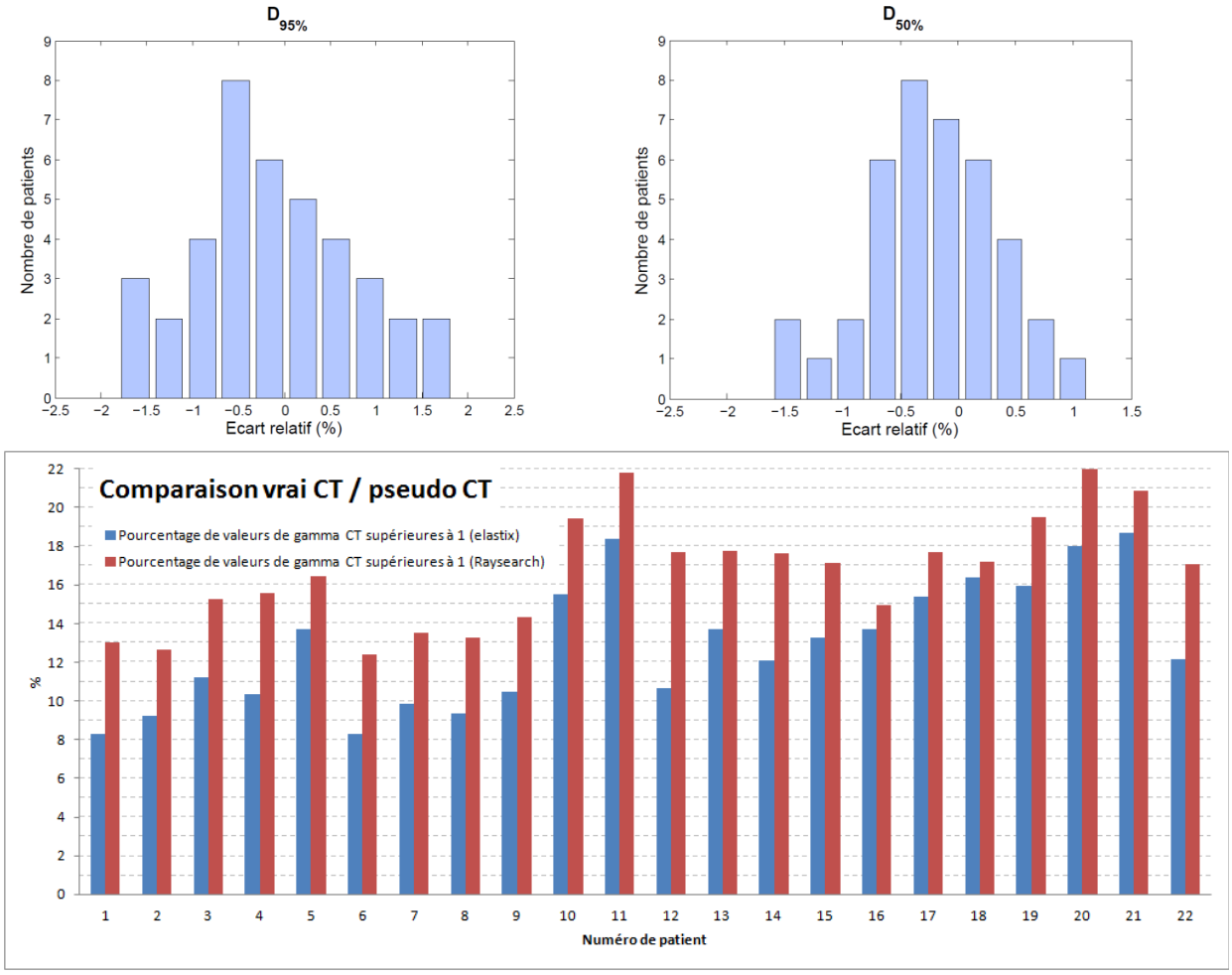


Figure 27 : Comparaison entre les deux algorithmes de régistration d'images déformable (ITK/Elastix, RayStation/Masterplan). En haut une comparaison dosimétrique : le D_{95} et le D_{50}, calculé avec Moderato pour les 22 patients. En bas, des tests gamma, afin de comparer les images pseudo-CT généré avec les deux méthodes, et aussi la comparaison entre les deux IRMs déformés.

Clairement, la qualité des images pCT, obtenues utilisant ITK/Elastix est très comparable avec celles décrites dans l'article 4 (utilisant RayStation et Masterplan). Les recalages déformables et rigides sont correctement décrits par ITK/Elastix, ce qui nous offre la possibilité d'avoir un système complètement automatisé, indépendant des outils commerciaux. En conséquence, les résultats décrits dans l'article 5 (plus bas) sont obtenus en remplaçant RayStation/Masterplan par ITK/Elastix.

5. Article 5 : L'utilisation des séquences ZTE dans la méthode atlas

5.1 Introduction

Dans le premier article (article 4), on a vu que la méthode hybride combinant l'atlas et les niveaux de gris de l'image IRM, et utilisant des séquences T_1 avait bien fonctionné pour la plupart des patients. Les cavités d'air sont mieux reproduites qu'avec la méthode d'atlas classique. Par contre, le problème des anomalies dans la géométrie du patient n'est pas encore résolu. En figure 28 on voit un patient qui a subi une intervention chirurgicale pendant laquelle on lui a enlevé un bout de crâne. Du tissu mou s'est introduit dans le trou.

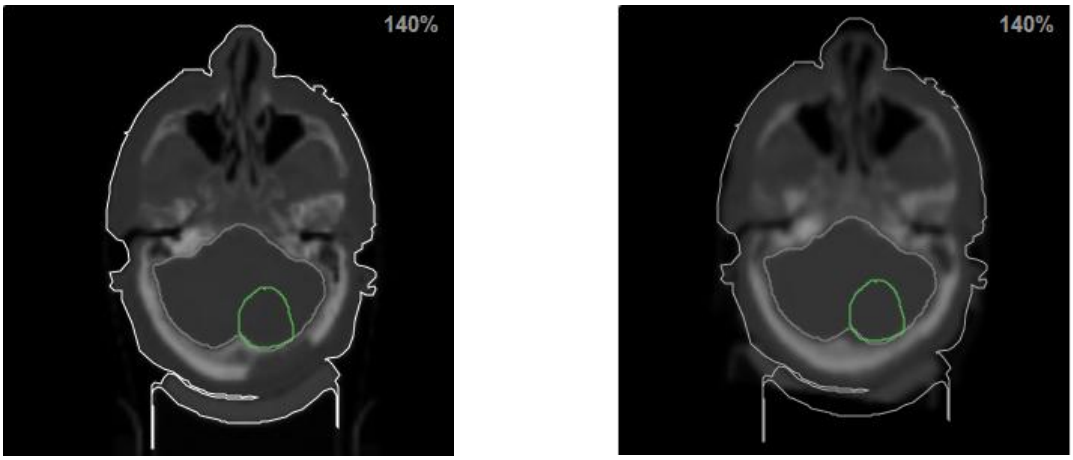


Figure 28 : Patient avec un trou dans le crâne visible dans le vrai CT (gauche) mais pas du tout reproduit dans le pCT, en utilisant la méthode hybride introduite dans article 4.

La méthode hybride fait déjà mieux que la méthode d'atlas classique, mais le problème est inhérent à l'utilisation des séquences T_1, où l'os n'est pas du tout visible. La méthode proposée ici va soustraire l'image IRM atlas déformée de l'IRM du nouveau patient. Si le recalage déformable a bien fonctionné cette différence est nulle (ou minimale). Dans les zones où la différence est plus importante (où le recalage n'a pas bien fonctionné à cause d'une anomalie dans la géométrie du nouveau patient), il faut utiliser une autre méthode. Le plus simple serait de convertir directement les niveaux de gris de l'IRM en UH. Mais ca n'est pas possible car il n'y a pas de relation directe entre l'IRM et le CT, surtout si on utilise une séquence conventionnelle comme le T_1 ou l'épaisseur de l'os du crâne est déjà complètement sous-estimée. C'est pour cette raison que nous avons décidé d'implémenter notre méthode utilisant des images ZTE. Même si on a vu plus haut qu'on ne peut pas encore utiliser ces séquences pour une conversion directe, l'épaisseur de l'os du crâne est quand-même correcte. Alors la méthode de détection de différences entre IRM atlas déformé et IRM du nouveau patient devrait beaucoup mieux marcher qu'avec la séquence T_1. Le but de l'article 5 était alors de focaliser complètement sur le patient avec le trou dans le crâne et de voir si notre méthode hybride utilisant des séquences ZTE serait capable de correctement reconstruire le trou dans le pCT.

5.2 Article 5

Zero TE MRI-only treatment planning for stereotactic radiation therapy of brain tumors after resection

Christine Boydev[1], Benjamin Demol[1,2,3], David Pasquier[1], Nick Reynaert[1]
[1] *Department of radiotherapy, Centre Oscar Lambret, Lille, France*
[2] *AQUILAB SAS, Loos Les Lille, France*
[3] *IEMN, UMR CNRS 8520, Villeneuve d'Ascq, France*

Article soumis à *Physics in Medicine and Biology*, Special Issue, le 05.09.16

Abstract. Using magnetic resonance imaging (MRI) as the sole imaging modality for patient modeling in radiation therapy is a challenging task due to the need to derive electron density information from MRI and construct a so-called pseudo-computed tomography (pCT) image. The purpose of this study was to propose an ordinary atlas-based method to derive synthetic CT numbers from an MR image acquired using a single ZTE sequence, combined with a novel *post hoc* classification-based correction method to improve the correctness of the pCT in regions of large anatomical discrepancies between the atlas and the patient using local intensity information. In the proposed two-stage correction process, a classification of the MR patient image was performed and pCT numbers were replaced by an inverse-distance weighted sum of the intensities of neighbors of the same class and with similar intensities. To evaluate results, the mean absolute error in bins of 20 HU was calculated with respect to the true planning CT scan of the patient. The proposed method improved pCT number estimation, in particular in the bone resection region where an ordinary atlas propagation scheme failed in generating soft tissue-equivalent CT numbers.

Keywords. MRI-only treatment planning, ZTE sequence, pseudo-CT, atlas-based method, head cancer

I. Introduction

Interest in magnetic resonance imaging (MRI) for use in radiation therapy (RT) planning and delivery has steadily increased over the last years due to its superior soft-tissue contrast compared to that of computed tomography (CT), and its potential for probing chemical composition, physiological behavior of tissues and changes in response to RT. These advantages, as well as the non-ionizing nature of MRI, the motivation to remove MRI-to-CT image fusion uncertainties and the recent development of commercial combined MRI-linear accelerators, account for the growing enthusiasm for MRI-only RT planning. However, there are challenges in using MRI as the sole imaging modality for patient modeling in RT, including lack of electron density information required for dose calculation and digitally reconstructed radiograph (DRR) generation for patient position verification. Indeed, image intensity in MRI, unlike in CT, does not reflect electron density and no clear

relationship between MRI signal and electron density has clearly been demonstrated using any of the existing standard imaging sequences.

Different methods to derive electron density information from MRI have been proposed. Four main approaches can be distinguished: atlas-based methods, patch-based, segmentation based and classification-based methods.

With atlas-based methods, an atlas comprised of an aligned pair of MR and CT images is used. The atlas MR image is registered to a patient MR image using deformable registration and the resulting transformation is applied to the atlas CT image to generate a predictive pseudo-CT image of the patient. This method was investigated by (Greer et al. 2011, Dowling et al. 2012) using an average atlas and T2-w MR images of male pelvis and by (Gudur et al. 2014) using a single random atlas and T1-weighted (T1-w) images of the brain. Also (Burgos et al. 2014) mapped each CT/T1-w MRI pair of a multiple brain atlas to the patient MRI and, in order to generate linear attenuation information, for each pixel in the patient, they fused CT intensities of all registered atlases at the investigated position on the basis of local image similarity measures between the patient MR image and each atlas MR image. The main advantages of these methods are that they are fully automatic and they generate realistic pseudo-CT images. The main drawbacks of these methods are that they ignore differences of tissue-equivalent electron density between atlas and patient and that registration algorithms may be unable to deform an atlas onto a patient with atypical or abnormal anatomy such as bone resections or implants.

With patch-based methods, a database of patches is constructed on the basis of pairs of aligned CT/MR images from different patients. By doing an intensity-based nearest neighbor search in the database and a similarity-weighted average of the CT numbers of the selected patches, a continuous-valued pCT number is assigned to each pixel of the test patient MR image. (Andreasen et al. 2015) used this technique with a conventional T1-w MRI sequence, and without using deformable registrations.

With segmentation-based methods, a direct conversion of MR intensities to HUs is performed. The simplest approach is to assign a uniform water-equivalent electron density to the whole patient (Beavis et al. 1998). To account for the heterogeneous nature of biological tissues, some authors used a more advanced 2-step approach consisting of manually segmenting tissue types relevant to dose calculations (typically, bone, lung and soft tissue) or specific organs, and assignment of bulk/uniform densities to each segment ((Lee et al. 2003, Chen et al. 2007, Lambert et al. 2011, Jonsson et al. 2010) using T1- or T2-weighted images of male pelvis, lung and head and neck. However, this approach leads to electron density errors derived from inter-observer contouring variability and lacks detailed representation of patient anatomy. A more sophisticated approach was proposed by (Korhonen et al. 2014), consisting of manually segmenting bones in male pelvic T1-T2*-weighted MR images and applying a direct non-uniform dual-model conversion of MRI intensities to HUs within and outside the bone segment.

To bypass the difficult task of manually segmenting bone, air cavities and soft tissue with traditional MR sequences, in particular in the brain where air and bone exist side by side and cannot be separated due to their similarly low signal, automatic algorithms are necessary. Authors in (Stanescu et al. 2008) performed atlas-based auto-segmentation of the scalp, bone and brain in T2-w head images using an atlas-based non-rigid registration technique and subsequently assigned bulk electron

densities to these head sub-structures. However, air regions were not segmented as they were considered irrelevant for dosimetric calculations.

Some authors tried to address the problem of automatically separating bone, air and soft tissue and linking MRI intensities to CT numbers in the head by using unsupervised classification -or regression-based methods (Keereman et al. 2010, Catana et al. 2010, Johansson et al. 2011, Berker et al. 2012, Hsu et al. 2013) as new research-oriented ultra short echo time (UTE) sequences were developed to respond to the critical need for visualizing tissues with short T2 such as cortical bone (T2∼0.5ms) (Robson et al. 2003, Robson and Bydder 2006, Tyler et al. 2007, Du and Bydder 2013). More specifically, dual-echo UTE (DUTE) sequences offer potential for elegantly differentiating bone from air. Indeed, with well-chosen echo times, the signal from bone is present in the first echo and not in the second echo, while signals from air and soft tissue are similar at both echo times. (Johansson et al. 2011, Berker et al. 2012, Hsu et al. 2013) combined DUTE and other standard MR imaging sequences. Authors in (Keereman et al. 2010) classified brain into 3 tissue classes (bone, soft tissue and air) using only the DUTE sequence (with 0.14ms/1.8ms echo times), and subsequently, assigned uniform linear attenuation coefficients to classes. To do so, they calculated the R2 map from the MR images acquired at the two echo times, ignoring pixels classified as air in the first-echo image using region-growing and thresholding techniques for mask generation, and after thresholding the R2 map to discriminate bone from soft tissue, they assigned uniform linear attenuation coefficients to classes. Derivation of attenuation-correction map for PET/MRI using the DUTE sequence with 0.07ms/2.46ms echo times was also performed for head imaging by (Catana et al. 2010).

Hybrid approaches combining the above-described methods have been investigated with the hope of compensating for the limitations of the methods if used alone. (Hofmann et al. 2008) combined local pattern recognition and atlas registration for T1-w brain images. (Gudur et al. 2014) developed an electron density mapping method using standard T1-w brain MRI and combining both intensity and geometry information into a probabilistic Bayesian framework.

This paper is an extension of a previous work where we proposed a new method to map CT numbers to MRI scans using a single-atlas propagation scheme followed by a post hoc correction of the CT numbers using local intensity information (Demol et al. 2016). The intensity of each pixel **p** in the deformed atlas CT image was replaced by an inverse-distance weighted sum of the intensities of some selected neighbors. The neighbors selected were those that had their MR intensity in the deformed atlas MR image close enough to the intensity of pixel **p** in the patient MR image. That boiled down to assigning a HU value to each pixel by calculating *on the fly* a local MRI-to-CT intensity conversion based on its neighborhood. We obtained promising results using standard T1-w head images for 22 patients who underwent CyberKnife stereotactic treatment. However, we found that our method failed in accurately estimating CT numbers in bone resection regions.

In the current paper, we therefore focused our efforts on trying to improve results in this very specific scenario where the patient had part of the skull removed by surgery prior to RT, causing surrounding soft tissue to naturally fill in the hole. For that purpose, we used MR images obtained with a zero echo time (ZTE) sequence resulting in higher differentiation between air, bone and soft tissue and hence, improved performance of the correction filter. In addition to our previous implementation, a classification of the patient MR image was performed and the correction filter was applied using local information of neighboring pixels of the same class for a more accurate correction

of CT numbers in the bone resection region. To the best of our knowledge, no publications to date have addressed the problem of electron density estimation from MRI in bone resection regions of the head, and specifically from ZTE MRI.

II. Materials and methods

A. Data description

The study population comprised 4 patients. One of these 4 patients underwent resection of part of the occipital bone, in the back of the head, prior to RT and was used as the test patient for this study. The three remaining patients were considered as atlas candidates. The normalized mutual information similarity metric was calculated between the MR image of each atlas candidate and the MR image of the test patient, and the most similar patient among the three candidates was chosen as the atlas. MR and CT images were acquired using the MR450w 1.5T GE scanner and the SOMATOM Definition AS Siemens scanner, respectively. For each patient, a planning CT image was acquired under the clinical routine protocol in treatment position, and additionally, for our research purposes, an MRI scan was acquired in the same position using the 3D Silenz ZTE sequence, with the following parameters: 12 µs echo time, 741µs repetition time, 1° flip angle, 240-260 mm FOV, 1.2 mm pixel size and 244 Hz/pixel bandwidth. For CT acquisition, the peak voltage, the X-ray tube current and the exposure time were 120 kVp, 133 mA and 1000 ms, respectively. Images were retrospectively analyzed after conversion from DICOM format to meta-image format (.mhd). For all our processing, MR and CT images had a size of 256x256x256 and 512x512x223 pixels, respectively, and a spacing of 0.94x0.94x0.94 mm^3 and 0.98x0.98x1 mm^3, respectively.

B. Pseudo-CT generation algorithm

The workflow of the image processing in the proposed framework is summarized in figure 1.

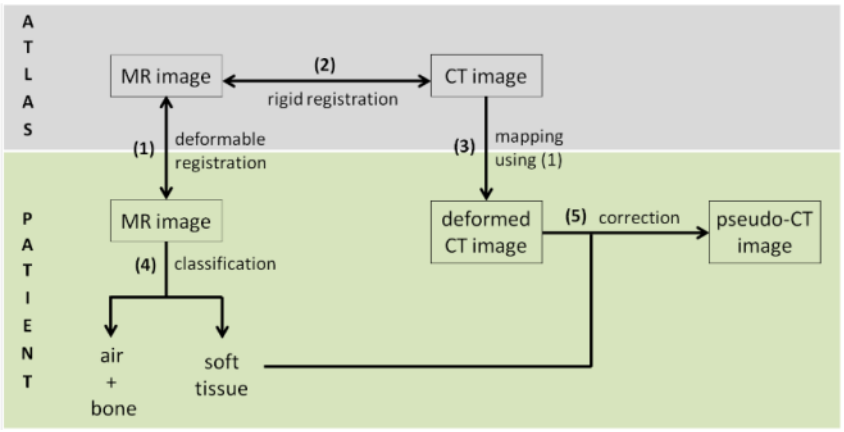


Figure 1: Pipeline of the data processing followed in this study.

B.a. Mapping of the atlas CT image onto the patient MRI anatomy

A single atlas was carefully chosen among the set of atlas candidates. To do so, the normalized mutual information similarity measure between the MR image of each atlas candidate and the MR image of the patient was computed and the atlas candidate with the highest similarity measure was

chosen as the atlas. Thereafter, the CT and MR images of the atlas were aligned using rigid registration as deformations within the head of the same patient are considered negligible (step (2) in figure 1). Furthermore, the atlas MR image was mapped onto the patient MR image using B-spline free-form deformable registration with Mattes mutual information similarity metric (step (1) in figure 1), and the resulting transformation was finally applied to the atlas CT image (step (3) in figure 1). The result was an estimated CT number map adapted to the patient MRI anatomy, in the physical space coordinate system of the patient MR image (i.e. same image grids). To improve accuracy, post-processing operations were carried out before producing the final pCT image (steps (4-5) of figure 1).

B.b. Post-processing correction

The post-processing correction step (steps (4-5) in figure 1) aimed at improving the correctness of the CT number map estimated through the ordinary atlas propagation scheme described in steps (1-3). A simple approach was to smooth the pCT image generated by steps (1-3) to remove noise. For example, replacing each pixel intensity of the generated pCT by an inverse-distance weighted sum of the intensities of the pixels in the neighborhood had the effect of reducing registration uncertainties. As a matter of fact, not all pixels in the neighborhood of a given pixel of interest (POI) were of the same tissue type as that POI, primarily at interfaces. Sometimes, applying a simple smoothing filter could be sufficient to reduce registration errors. However, it may be the case that not all pixels in the neighborhood should contribute to the smoothing, on pain of introducing unwanted additional correction errors.

Our approach aimed to select only pixels in the neighborhood that deserved to contribute to the smoothing, i.e. those that belonged to the same class as the POI. For that purpose, prior to correction, the patient MR image was classified using an expectation-maximization (EM) based clustering algorithm (Fedorov et al. 2012). Although the ZTE sequence allowed for a better separation between air, bone and soft tissue than standard sequences (e.g. T1-w), it did not completely solve the problem of separation as illustrated in figure 2, with no clear bijective relation between CT and MR intensities, and classification into three tissue classes yielded unacceptable errors. However, a classification into two tissue classes, one for air and bone together and one for soft tissue, yielded satisfactory results. Separating soft tissue pixels from the rest of the pixels was sufficient to meet the purpose of this study, i.e. correcting CT numbers at pixels actually belonging to soft tissue in the bone resection region. The following correction process was carried out.

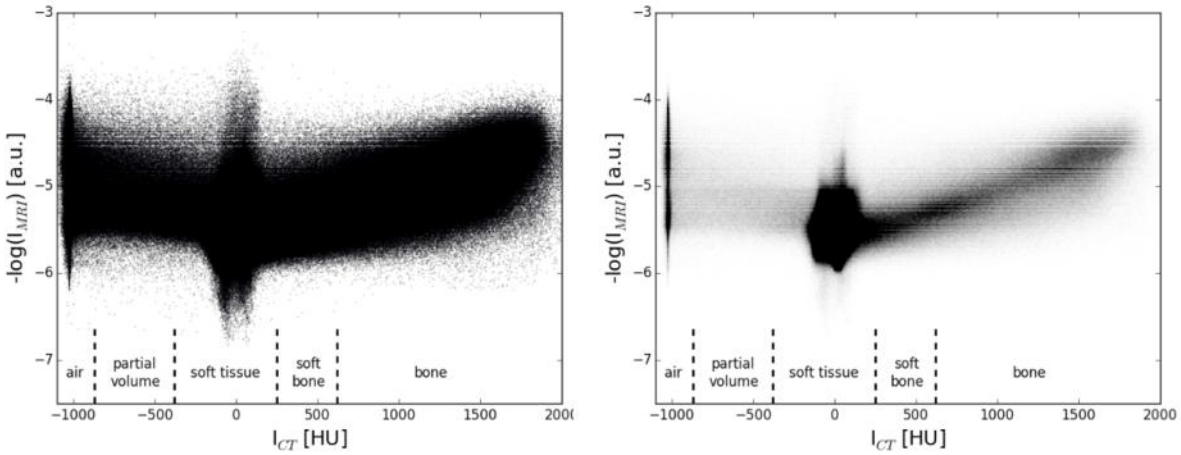


Figure 2: Relation between CT numbers and MRI intensities using (a-b) the T1-w image, the (c-d) ZTE image and the CT image of the test patient after alignment, for all pixels inside the head (same mask used). Scatter plots have two different marker sizes: (a,c) high size (area of 0.01 points squared) and (b,d) the lowest size afforded by matplolib (Hunter 2007), that is, an area of 0.000003 points squared.-log(I_{MRI}) values are shown in (a-d) in order to compare with (Wiesinger et al. 2016).

Let I_{MRI} , $I_{defAtlasMRI}$ and I_{pCT} denote the patient MR image, the deformed atlas image and the pCT image generated from steps (1-3) of figure 1, respectively, and after normalization of grayscale values between the MR images by histogram matching. Let $C_{\boldsymbol{p}} \in \{0,1\}$ be the estimated class of a pixel **p**, with 0 for air and bone, and 1 for soft tissue.

Loop 1: for each pixel **p** of the patient image grid in the patient image grid (figure 3 (a)),

- 1. loop 2: iterate over the neighborhood of **p** to find all pixels $\{\boldsymbol{q_i}\}_{i=0,1}$ such that

$$C_{\boldsymbol{q_i}=C_{\boldsymbol{p}}} \tag{1}$$

and

$$\left|\frac{I_{defAtlasMRI}(\boldsymbol{q_i}) - I_{MRI}(\boldsymbol{p})}{I_{MRI}(\boldsymbol{p})}\right| < A \tag{2}$$

where A is a given threshold (values ranging from 0:05 to 1 were tested) and the neighborhood is a box of 9x9x9 pixels, i.e. of radius 4 pixels (figure 3 (b)),

- 2. Replace I_{pCT}(**p**) by $\sum_i \frac{\alpha_i}{\sum_j \alpha_j} I_{pCT}(\boldsymbol{q_i})$ where $\alpha_i = \frac{1}{d_i}$ if $\boldsymbol{q_i} \neq \boldsymbol{p}$, otherwise $\alpha_i = \frac{2}{d_0}$ (arbitrarily) with d_i the distance between **p** and $\mathbf{q_i}$, and d_0 the minimum spacing between pixels (figure 3(c)).

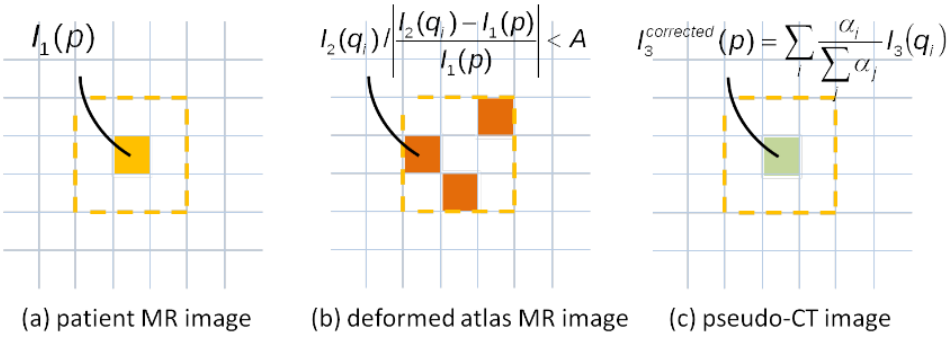


Figure 3: *Post hoc* correction of the pCT image: strategy for selecting the neighbors (shown in (b)) of a given pixel of interest (POI) of the patient MR image (shown in (a)) that were used for correcting its HU value (shown in (c)). Only neighbors of the same class as that of the POI were eligible for selection. All images were in the coordinate system of the patient MR image. I_1, I_2 and I_3 are the patient MR image, the deformed atlas MR image, and the pCT image, respectively. i and j are mute indices addressing pixels in the neighborhood, p is the position of the POI, q_i is the position of a pixel in the neighborhood of p, α_i and α_j are inverse-distance weighting coefficients and A is a given threshold.

Ensuring that pixel $\mathbf{q_i}$ in the neighborhood of **p** belonged to the same tissue type as that of **p** aimed at preventing from selecting a pixel $\mathbf{q_i}$ that showed a similar MR intensity as **p**, without actually belonging to the same tissue class. In this work, although only soft tissue was separated from other tissues through a two-class image segmentation, correction was applied to pixels of both classes. This post-processing correction step was presented in (Demol et al. 2016) without making a distinction between classes. No classification prior to correction was conducted and each neighbor of a POI that verified equation 2 was selected for correction, without considering whether its tissue type was identical to that of the POI, which could possibly cause correction errors. The main new contributions here were the following. The method integrated a classification step and was evaluated on a patient having undergone bone resection, using a new dataset of ZTE MR images.

C. Validation

To evaluate the correctness of the CT number estimate, the pCT image was compared to the true planning CT (tCT) scan of the patient by calculating the pixel-to-pixel mean absolute error (MAE) histogram using consecutive, non-overlapping bins of size 20 HU across the HU scale as in (Andreasen et al. 2015): $MAE(HU) = \frac{1}{N}\sum_{i=0}^{N}|CT(i) - pCT(i)|$ where N is the number of pixels having intensity values falling into interval [*HU*-10, *HU*+10] in the tCT scan. The MAE histogram was calculated in the head, as well as in a smaller, rectangle-shaped region of interest, ($ROI_{resection}$), containing the bone resection area to reveal the potential of our method, and illustrated in figure 4.

D. Software

For this study, all the data processing and visualization were performed on a Linux computer with distribution openSUSE LEAP 42.1, x86 64 architecture, Intel Core i7-5600U, 3.1 GHz processors, 2 cores/2 threads per core, 16 GB RAM and 4 MB cache. For the implementation of our methods, the following open-source software, based on C++, was used:

- the Insight Toolkit ITK (Johnson et al. 2015)
- `elastix`, based on the ITK library (Klein et al. 2010)
- the ITK-based Command Line Image Toolkit `clitk` for the pre-processing of the images, and the integrated image-visualization tool VV (Rit et al. 2011)

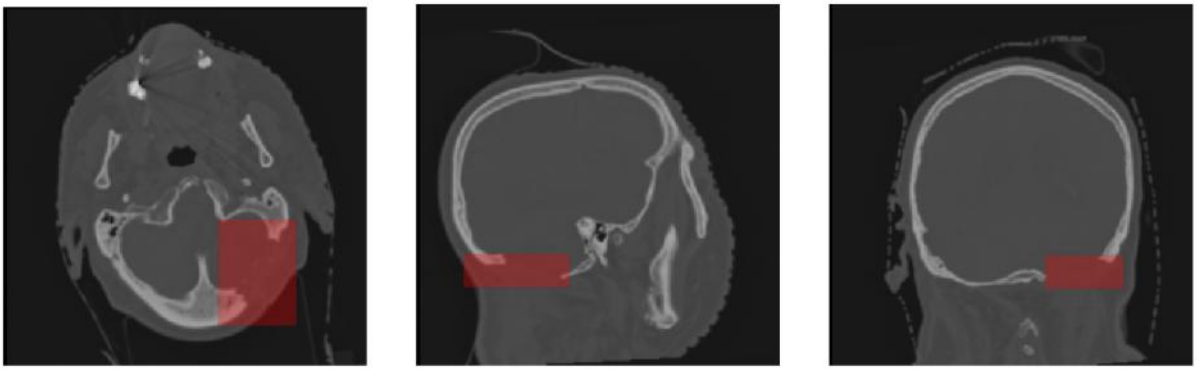


Figure 4: Rectangle-shaped region of interest (in red) used for MAE calculations and containing the bone resection region, overlaid on the planning CT scan of the test patient, in (a) axial, (b) sagittal and (c) coronal planes.

III. Results

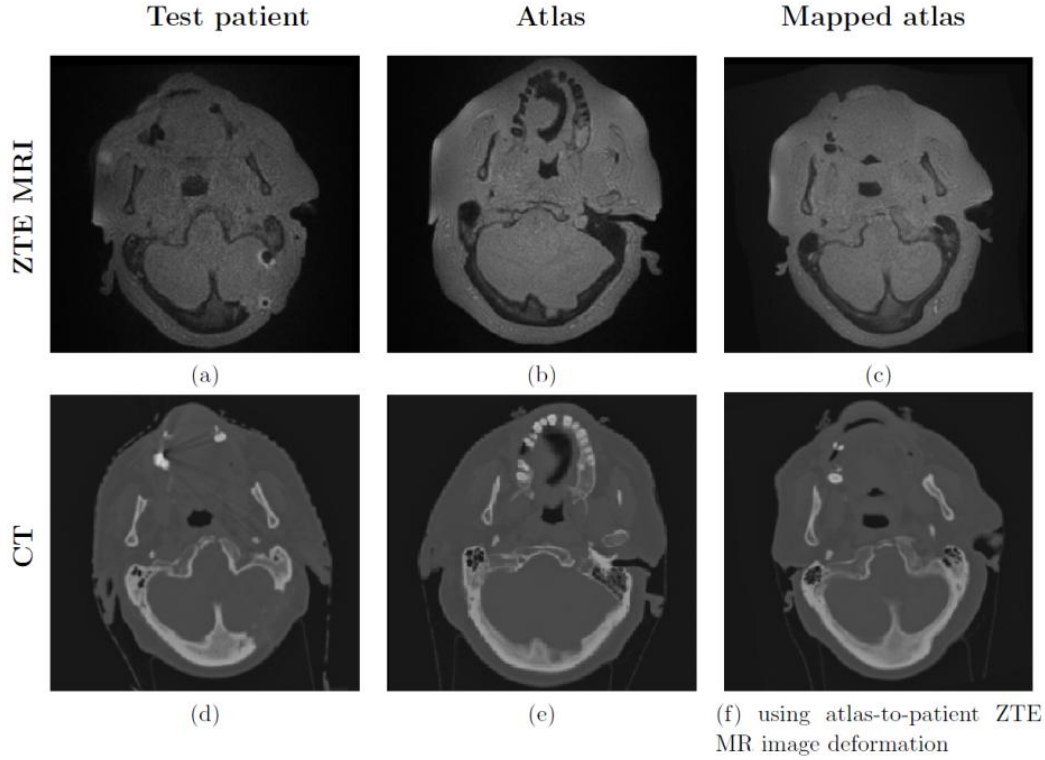


Figure 5: Axial (a-c) ZTE MRI and (d-f) CT slices of (first column) the test patient, (second column) the atlas and (third column) the atlas mapped using the transformation resulted from atlas-to-patient MR image deformable registration and before the proposed *post hoc* correction.

In figure 5, axial slices of the MR and CT images of the test patient, the atlas and the atlas mapped onto the patient are shown. Figure 5f corresponds to the pCT image generated by the atlas propagation scheme described in steps (1-3) of figure 1, and before applying the proposed *post hoc* classification-based correction. Thus, we observed that the atlas mapping resulted in the assignment of cortical bone CT numbers in the bone resection region of the test patient. Expectedly, the atlas

propagation scheme was unable to generate correct CT numbers in a region where the anatomy of the atlas and that of the patient did not match, as in $ROI_{resection}$.

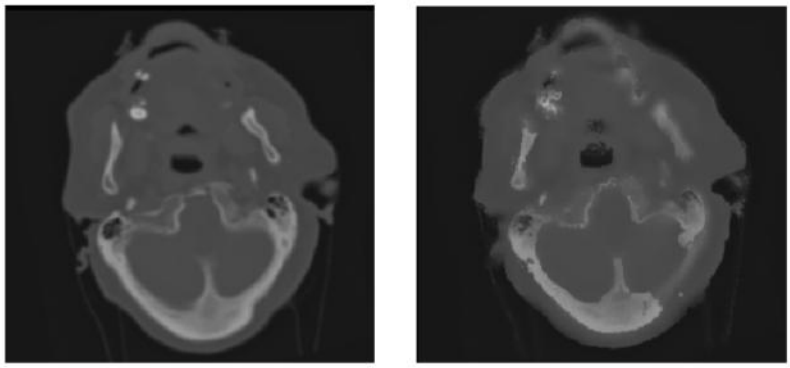


Figure 6: Slice of the pCT image of the test patient (a) after using the simple smoothing with a neighborhood radius of 1 pixel and (b) after applying the proposed *post hoc* classification-based correction with A = 0:3 and a neighborhood radius of 4 pixels, using the same window level as in figure 5.

Figure 6 highlights the improvement produced by the proposed *post hoc* correction scheme (with A=0.3 and a neighborhood radius of 4 pixels) in the pCT image, especially in $ROI_{resection}$, in comparison with a simple smoothing method (with a neighborhood radius of 1 pixel), showing the same axial slice as in figure 5. In addition, the result of the classification of the test patient MR image is displayed in figure 7. To evaluate the performance of the classifier, we segmented the tCT scan into 2 classes using thresholding. All pixels with CT numbers lower than -380 and higher than 250 were considered as air, partial volume and bone, and intermediary CT numbers were associated to soft tissue, similarly to (Rank et al. 2013). Table 1 reports the proportions of pixels calculated in different tissue types in the head in the tCT scan. Visually, a good agreement between the classified patient MR image and the segmented tCT image was reached, as depicted in figure 7. Quantitatively, we found a sensitivity of 89% (97%) and a specificity of 96% (82%) in the head (in $ROI_{resection}$). The sensitivity measures the percentage of soft tissue pixels that were correctly identified as such. The specificity measures the percentage of bone and air pixels that were correctly identified as such.

air	partial volume	soft tissue	soft bone	bone
< -870 HU	[-870; -380] HU	[-380; 250] HU	[250; 620] HU	> 620 HU
2%	1%	81%	4%	12%

Table 1: Proportion of pixels in different tissue types in the head based on thresholding of the tCT scan of the patient.

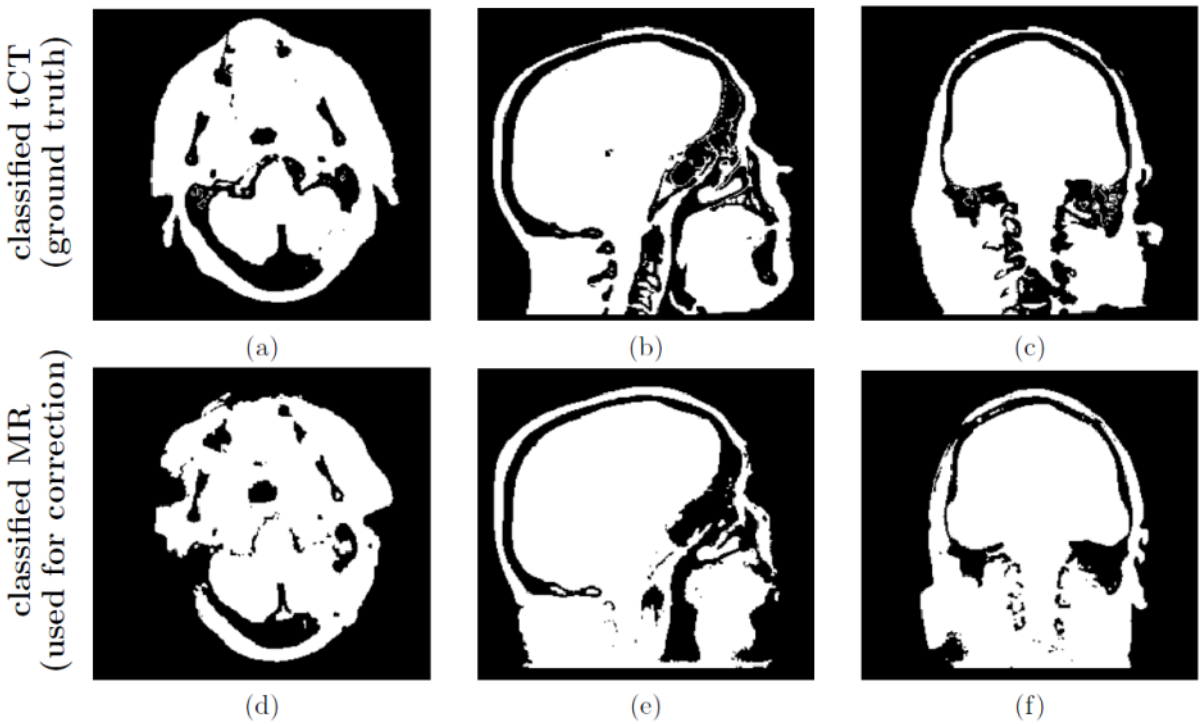


Figure 7: Visual comparison of the classified MR image partitioned into two tissue classes (one for air and bone together and one for soft tissue) against the tCT segmented by thresholding (ground truth), in (a,d) axial, (b,e) sagittal and (c,f) coronal planes.

Figure 8 provides graphs representing the pixel-to-pixel MAE as a function of the tCT value and calculated in consecutive, non-overlapping bins of 20 HU across the HU scale for pixels in the head, as well as in $ROI_{resection}$. Also, the cumulative number of pixels in all of the HU bins up to the specified bin was displayed on each graph to get a better sense of the proportion of pixels falling into the specified HU bin. Firstly, the simple correction approach consisting of replacing each pixel intensity of the generated pCT by an inverse distance weighted sum of the intensities of the pixels in the neighborhood regardless of what tissue type they belonged to was tested with different sizes of neighborhood ranging from 3 to 9 pixels (radii from 1 to 4 pixels) (figures 8a and 8b). Higher sizes of neighborhood lead to lower errors for partial volume pixels and bone pixels with CT numbers up to 1000 HU, but to higher errors for bone pixels with CT numbers higher than 1000 HU. No differences were found in the soft tissue range, in particular in $ROI_{resection}$. To improve the quality of CT number estimation for soft tissue pixels in $ROI_{resection}$ while limitating its deterioration for bone pixels with high CT numbers, we tested the proposed classification-based correction method for different values of A ranging from 0.1 to 0.7. The inspection of the MAE revealed an improvement in the HU estimation for CT numbers up to 1500 HU for lower values of A in the head (figures 8a), and up to 1000 HU in $ROI_{resection}$ (figures 8b) compared to results before correction.

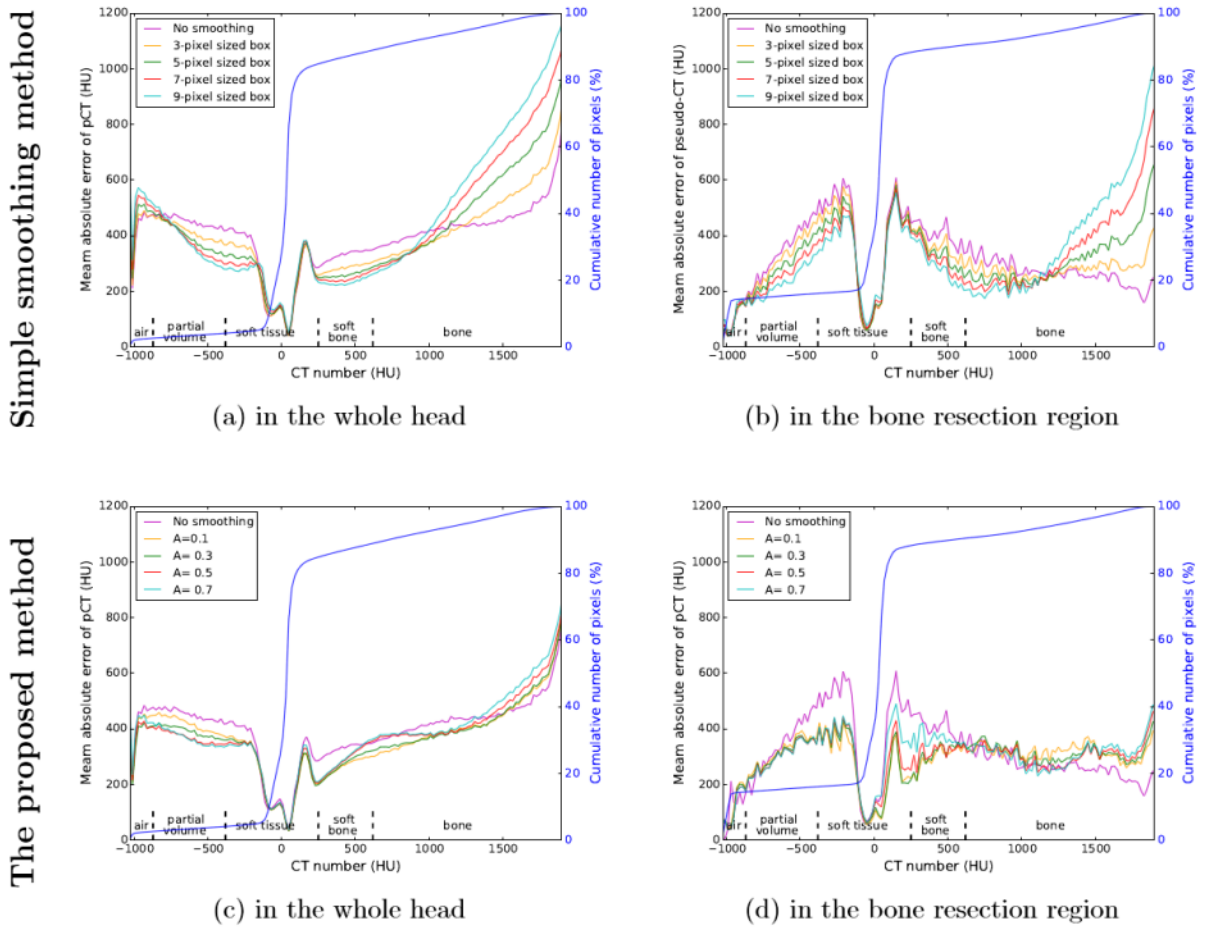


Figure 8: Effect of the pCT correction using (a-b) a simple smoothing with four different sizes of neighborhood and (c-d) our proposed method with a neighborhood radius of 4 pixels, considering pixels only (a,c) within the head and (b,d) in $ROI_{resection}$. The cumulative number of pixels was also shown to enhance the range of HUs with higher proportion of pixels (curve in blue, with a differently scaled y-axis on the right).

IV. Discussion

The analysis of the results produced with the simple smoothing for different neighborhood sizes (figures 8a and 8b) showed an improvement in CT number estimation after correction, except for CT numbers higher than 1000 HU and equivalent to air where MAEs were larger. These pixels most probably concerned areas of interfaces between bone (air) and soft tissue or air (bone) where very different intensity values contributed to the smoothing regardless of their tissue type. The best compromise seemed to be obtained with a neighborhood size of 3 pixels (i.e. radius of 1 pixel). However, additionally to errors in high CT number estimation, the simple smoothing method failed in estimating correct CT numbers at pixels located at the bone resection, as illustrated in figure 6a.

On the contrary, the proposed *post hoc* correction method improved CT number estimation in the cortical bone resection region as shown in figure 6b, and in figure 8d compared with figure 8b for soft tissue pixels. We recall that in the tCT scan, the bone that was removed by resection gave way to soft tissue. The best compromise for the proposed correction method seemed to be A=0.1 and A=0.3. Also, lower MAEs were found for high CT numbers with the proposed method (with A=0.3) compared with the simple smoothing correction method. Nevertheless, the proposed correction

method failed in improving correctness of CT number estimation for dense bones (HU>1500) in comparison with the absence of correction. This could be caused by classification errors, misleading the selection of neighboring pixels in the correction process (see condition 1). In fact, the classifier was found to have a specificity of 82% in $ROI_{resection}$, which meant that 18% of pixels classified as belonging to the second class, that is, bone or air, were actually wrongly classified. Another reason could be the lack of intensity separation between bone and other tissue types as depicted in figure 2, also misleading the selection of neighboring pixels in the correction process (see condition 2).

The proposed *post hoc* correction method showed promising results and merits further investigation. In this study, a simple EM-based classifier was used and a more sophisticated clustering method would most probably enhance the classification performance and hence, the correctness of the neighboring pixel selection in the proposed correction process. Furthermore, the quality of the ZTE images could be improved to accentuate separation between tissues, and improve the correctness of the proposed correction process.

V. Conclusion

In this work, we proposed an ordinary atlas-based method to derive synthetic CT numbers from an MR image acquired using a single ZTE sequence, combined with a novel *post hoc* classification-based correction method to improve the correctness of the pCT in regions of large anatomical discrepancies between the atlas and the patient using local intensity information. We demonstrated the potential of our correction method in a patient having undergone bone resection surgery prior to radiation therapy.

Acknowledgments

The authors thank centre Léonard de Vinci (Dechy, France) where the data sets used in this study were acquired. This work was part of a PhysiCancer project financed by the French Institut National du Cancer (INCA).

References

Andreasen D, Van Leemput K, Hansen R H, Andersen J A L and Edmund J M 2015 Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain. Med Phys 42(4), 1596-1605.

Beavis A, Gibbs P, Dealey R and Whitton V 1998 Radiotherapy treatment planning of brain tumours using MRI alone. Br J Radiol 71(845), 544-548.

Berker Y, Franke J, Salomon A, Palmowski M, Donker H C W, Temur Y, Mottaghy F M, Kuhl C, Izquierdo-Garcia D, Fayad Z A, Kiessling F and Schulz V 2012 MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/dixon MRI sequence. J Nucl Med 53(5), 796-804.

Burgos N, Cardoso M J, Thielemans K, Modat M, Pedemonte S, Dickson J, Barnes A, Ahmed R, Mahoney C J, Schott J M, Duncan J S, Atkinson D, Arridge S R, Hutton B F and Ourselin S 2014 Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies. IEEE Trans Med Imaging 33(12), 2332-2341.

Catana C, van der Kouwe A, Benner T, Michel C J, Hamm M, Fenchel M, Fischl B, Rosen B, Schmand M and Sorensen A G 2010 Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype. J Nucl Med 51(9), 1431-1438.

Chen L, Nguyen T B, Jones E, Chen Z, Luo W, Wang L, Price, Jr R A, Pollack A and Ma C M C 2007 Magnetic resonance-based treatment planning for prostate intensity-modulated radiotherapy: creation of digitally reconstructed radiographs. Int J Radiat Oncol Biol Phys 68(3), 903-911.

Demol B, Boydev C and Reynaert N 2016 Atlas-based method applied to intracranial treatments using Cyberknife device in MR-only treatment planning 55èmes Journées Scientifiques de la Société Française de Physique Médicale.

Dowling J A, Lambert J, Parker J, Salvado O, Fripp J, Capp A,Wratten C, Denham JWand Greer P B 2012 An atlas-based electron density mapping method for magnetic resonance imaging (MRI)- alone treatment planning and adaptive MRI-based prostate radiation therapy. Int J Radiat Oncol Biol Phys 83(1), e5-11.

Du J and Bydder G M 2013 Qualitative and quantitative ultrashort-TE MRI of cortical bone. NMR Biomed 26(5), 489-506.

Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller J V, Pieper S and Kikinis R 2012 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 30(9), 1323-1341.

Greer P B, Dowling J A, Lambert J A, Fripp J, Parker J, Denham J W, Wratten C, Capp A and Salvado O 2011 A magnetic resonance imaging-based workflow for planning radiation therapy for prostate cancer. Med J Aust 194(4), S24-S27.

Gudur M S R, Hara W, Le Q T, Wang L, Xing L and Li R 2014 A unifying probabilistic Bayesian approach to derive electron density from MRI for radiation therapy treatment planning. Phys Med Biol 59(21), 6595-6606.

Hofmann M, Steinke F, Scheel V, Charpiat G, Farquhar J, Aschoff P, Brady M, Sch□olkopf B and Pichler B J 2008 MRI-based attenuation correction for PET/MRI: a novel approach combining pattern recognition and atlas registration. J Nucl Med 49(11), 1875-1883.

Hsu S H, Cao Y, Huang K, Feng M and Balter J M 2013 Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy. Phys Med Biol 58(23), 8419-8435.

Hunter J D 2007 Matplotlib: A 2D graphics environment Computing In Science & Engineering 9(3), 90-95.

Johansson A, Karlsson M and Nyholm T 2011 CT substitute derived from MRI sequences with ultrashort echo time. Med Phys 38(5), 2708-2714.

Johnson H J, McCormick M M and Ibanez L 2015 The ITK Software Guide.

Jonsson J H, Karlsson M G, Karlsson M and Nyholm T 2010 Treatment planning using MRI data: an analysis of the dose calculation accuracy for different treatment regions. Radiat Oncol 5, 62.

Keereman V, Fierens Y, Broux T, De Deene Y, Lonneux M and Vandenberghe S 2010 MRIbased attenuation correction for PET/MRI using ultrashort echo time sequences. J Nucl Med 51(5), 812-818.

Klein S, Staring M, Murphy K, Viergever M A and Pluim J P W 2010 elastix: a toolbox for intensitybased medical image registration. IEEE Trans Med Imaging 29(1), 196-205.

Korhonen J, Kapanen M, Keyriläinen J, Seppälä T and Tenhunen M 2014 A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer. Med Phys 41(1), 011704.

Lambert J, Greer P B, Menk F, Patterson J, Parker J, Dahl K, Gupta S, Capp A, Wratten C, Tang C, Kumar M, Dowling J, Hauville S, Hughes C, Fisher K, Lau P, Denham J W and Salvado O 2011 MRI-guided prostate radiation therapy planning: Investigation of dosimetric accuracy of MRI-based dose planning. Radiother Oncol 98(3), 330-334.

Lee Y K, Bollet M, Charles-Edwards G, Flower M A, Leach M O, McNair H, Moore E, Rowbottom C and Webb S 2003 Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. Radiother Oncol 66(2), 203-216.

Rank C M, Tremmel C, Hünemohr N, Nagel A M, Jäkel O and Greilich S 2013 MRI-based treatment plan simulation and adaptation for ion radiotherapy using a classification-based approach. Radiat Oncol 8, 51.

Rit S, Pinho R, Delmon V, Pech M, Bouilhol G, Schaerer J, Navalpakkam B, Vandemeulebroucke J, Seroul P and Sarrut D 2011 VV, a 4D slicer. in `Medical Imaging and Computer-Assisted Intervention MICCAI, Fourth International Workshop on Pulmonary Image Analysis' pp. 171-175.

Robson M D and Bydder G M 2006 Clinical ultrashort echo time imaging of bone and other connective tissues. NMR Biomed 19(7), 765-780.

Robson M D, Gatehouse P D, Bydder M and Bydder G M 2003 Magnetic resonance: an introduction to ultrashort TE (UTE) imaging. J Comput Assist Tomogr 27(6), 825-846.

Stanescu T, Jans H S, Pervez N, Stavrev P and Fallone B G 2008 A study on the magnetic resonance imaging (MRI)-based radiation treatment planning of intracranial lesions. Phys Med Biol 53(13), 3579-3593.

Tyler D J, Robson M D, Henkelman R M, Young I R and Bydder G M 2007 Magnetic resonance imaging with ultrashort TE (UTE) PULSE sequences: technical considerations. J Magn Reson Imaging 25(2), 279-289.

Wiesinger F, Sacolick L I, Menini A, Kaushik S S, Ahn S, Veit-Haibach P, Delso G and Shanbhag D D 2016 Zero TE MR bone imaging in the head Magnetic Resonance in Medicine 75(1), 107-114.

6. Suite du projet PhysiCancer

6.1 La méthode multi-atlas

Notre méthode hybride était d'abord basée sur un atlas constitué d'un patient sélectionné « manuellement ». Le patient ne présentait pas d'anomalies dans sa géométrie ni de distorsions dans ses images IRM, et une densité d'os correspondant à la moyenne de la population des patients. Dans l'article 5, la paire « atlas-nouveau patient » est sélectionnée plus formellement en utilisant le critère de similarité concernant les images IRM. A l'origine, on ne pouvait pas faire mieux par manque de données de Douai.

Dans la littérature on trouve plusieurs méthodes multi-atlas [60][66][67]. Dans ce cas, l'atlas consiste en une dizaine de paires CT-IRM avec des géométries diverses. L'application la plus simple est de comparer l'IRM du nouveau patient avec les IRMs de l'atlas et de chercher le cas qui est le plus similaire pour le recalage déformable. Une autre option est de combiner tous les cas atlas. La valeur d'un voxel dans le pCT est une somme pondérée des valeurs obtenues en utilisant tous les cas de l'atlas. Les facteurs de pondération sont déterminés par la similarité.

Dans la suite de notre projet, il faut tester la valeur ajoutée de cette méthode par rapport à notre méthode (utilisation d'une simple paire CT-IRM comme atlas). On peut déjà imaginer que dans le crâne le multi-atlas va moins apporter que dans le pelvis ou le thorax où les différences entre les patients peuvent être beaucoup plus importantes. L'article de Sekine et al [67] a montré une valeur ajoutée pour le multi-atlas même dans le crâne.

Il faut insister sur le fait que, même avec une méthode multi-atlas, on ne va pas résoudre les anomalies étudiées plus haut (trou dans le crâne par exemple). En plus, pour le moment, on a toujours négligé l'importance des CT dans l'atlas. Le choix du cas, ou les facteurs de pondération utilisés sont toujours basés sur la similarité entre les images IRM. Mais avant d'inclure des cas atlas il faut aussi tenir compte d'autres facteurs, comme les distorsions dans les images IRM ou la qualité du CT. Là aussi, il faut éviter les artéfacts (implants dentaires par exemple) et s'assurer que le protocole du scanner CT soit toujours le même ou tout au moins que les UH ne soient pas trop influencées. Un autre problème est que la densité de l'os (par exemple dans le crâne) peut différer d'un patient à l'autre. Cette information n'est pas disponible dans les images IRM. Ce sont donc les valeurs du CT atlas qui sont directement utilisées pour définir celles du pCT.

6.2 Autres indications cliniques

Jusqu'à maintenant, nous n'avons étudié que le crâne. L'avantage du crâne est qu'il ne faut pas nécessairement positionner le patient en position de traitement sur l'IRM. Pour les autres indications, cela devient par contre indispensable. Il faut alors des données des patients effectivement simulées sur l'IRM avec une table plate et tous les systèmes de contention. Nous avons récemment reçu des données IRM et CT de dix patients prostate, simulés en position de traitement, du HUS de Helsinki (Finlande). Nous allons maintenant nous focaliser sur ces cas en utilisant notre méthode hybride. Pour ces patients il n'y pas d'images ZTE, mais on peut imaginer que dans le pelvis ça sera moins important. L'utilisation d'un multi-atlas sera par contre indispensable.

On va installer entretemps le nouvel IRM au COL. Il sera partiellement dédié à la radiothérapie avec la possibilité de simuler les patients en position de traitement.

Les patients foie seront très intéressants à traiter en IRM-seule. Le GTV est souvent très difficile à définir sur les images CT et le recalage entre IRM et CT (pour propager les contours) est extrêmement difficile aussi. Là aussi, la génération des images pCT demande l'utilisation d'une méthode multi-atlas. Les images ZTE ne seront, là encore, pas nécessaires. Une difficulté additionnelle sera l'impact des mouvements en IRM. Il faut utiliser des séquences rapides ou faire de l'IRM 4D.

Enfin, le challenge le plus important sera les traitements du poumon. Ils seront étudiés à la fin du projet. Pour ces traitements l'IRM 4D est indispensable.

6.3 Les images ZTE : en routine et fondamental

Jusqu'à maintenant, nous avons utilisé les séquences ZTE dans la méthode atlas. Une utilisation directe (conversion directe entre IRM et CT sans déformations) semble plus élégante. Pour le moment, la qualité en termes de distinction entre air et os de ces images n'est pas assez élevée. Afin d'optimiser ces séquences, nous collaborons avec le LTSI et peut-être prochainement avec le laboratoire Siemens à l'EPFL (Ecole Polytechnique Fédérale de Lausanne) en Suisse. Ce labo serait prêt à installer leurs séquences ZTE (séquence PETRA) à l'ICO de Nantes (deuxième partenaire dans le projet PhysiCancer), où ils utilisent un IRM Siemens.

Une fois notre nouvelle séquence optimisée (une séquence ZTE en deux temps par exemple, qu'on pourrait appeler double ZTE), on peut de nouveau essayer de convertir directement les valeurs IRM en valeurs CT. On pourrait même essayer de déterminer le contenu hydrogène dans l'image afin d'avoir une relation physique entre intensité IRM et sections efficaces d'interactions dans notre plateforme Monte Carlo, ce qui était le but original du projet.

Même dans le cas où une conversion directe est possible, il faut encore démontrer la valeur ajoutée par rapport à la méthode atlas. Une conversion directe a l'air moins robuste concernant les distorsions en IRM. Effectivement, ces distorsions seront propagées directement vers le pCT. En utilisant la méthode atlas, ce sont les images CT de l'atlas qui vont déterminer directement les valeurs dans le pCT. Il faut donc bien comparer la robustesse des deux méthodes de conversion. Cela fera l'objet d'études plus poussées aussi. Finalement, une fois la méthode de génération pCT optimale trouvée, il faudra encore démontrer l'apport de cette nouvelle méthode par rapport à la méthode de propagation de contours (combinaison de l'IRM et du CT comme on le fait actuellement en routine). Ceci devrait normalement être facile à démontrer. En effet, une erreur dans la génération d'un pCT va introduire des UH fausses dans les images, et par conséquence une déviation du calcul de dose de quelques pourcents. Par contre, le contour du GTV sera au bon endroit. Si on propage les contours, la dose sera peut-être plus précise, mais l'irradiation sera peut être faite à côté de la tumeur. Alors, intuitivement, la valeur ajoutée de la radiothérapie en IRM seule est claire, mais il faut quand-même le démontrer plus formellement.

6.4 Achat du nouvel équipement au Centre Oscar Lambret

Au Centre Oscar Lambret, l'idée est vraiment de cibler les traitements de plus en plus sur l'IRM. Pour les traitements en IRM-seule, l'achat d'un simulateur IRM est prévu en 2017.

En plus de cette machine, l'installation d'un IRM-Linac est prévue dans le futur proche. De cette façon, toute la chaîne de traitement sera basée sur l'IRM. Les images de référence, déterminées pendant la simulation IRM seront parfaitement compatibles avec les images de positionnement et de suivi pendant les irradiations. L'avantage de l'IRM-Linac est la possibilité de produire des images IRM en temps réel pendant tout le temps du traitement (et de faire du Gating ou du tracking). Le contraste est également beaucoup plus élevé qu'avec le CT et offre la possibilité de voir la tumeur en temps réel, au lieu de fiduciels. En plus, un tel système offre la possibilité d'adapter le traitement avant chaque séance prenant en compte les changements de géométrie du patient (amaigrissement, diminution du volume de la tumeur, …). On peut aussi « facilement » déterminer la dose effectivement délivrée en utilisant les images du patient pendant le traitement (de cette façon, les chapitres 2 et 3 de ce texte se rencontrent). Si l'IRM fonctionnelle est disponible sur cette machine, on pourra même faire le suivi biologique pendant le traitement et utiliser l'IRM comme biomarqueur (et de nouveau, si nécessaire, adapter le traitement).

Chapitre 4 : Dosimétrie fondamentale

1. Conversion Dose to medium/Dose to water

1.1 Introduction

Depuis qu'on a introduit les algorithmes Monte Carlo dans les systèmes de planification, une discussion est lancée dans la littérature scientifique concernant la façon d'exprimer la dose : en dose dans l'eau ou en dose dans le milieu (par exemple dose dans l'os) [68]. Dans les TPSs conventionnels, le patient est modélisé comme de l'eau avec une densité électronique variable. Dans chaque voxel, les UHs du scan sont converties linéairement en densité électronique, mais la vraie composition des tissus n'est pas prise en compte. Dans la gamme d'énergie de la radiothérapie moderne en photons (6 MV – 25 MV) toutes les interactions rayons-matière sont dominées par l'effet Compton qui ne dépend que de la densité électronique. C'est pourquoi cette approche a toujours donné des résultats acceptables. Par contre, dans un algorithme Monte Carlo, il y a la possibilité de définir explicitement la composition de chaque voxel, comme c'est décrit dans le chapitre 3 (article 3) et dans l'introduction. On peut utiliser des inserts tissus-équivalents pour étalonner le CT (en densité et en composition) ou, de manière plus avancée, appliquer un étalonnage stœchiométrique. Dans ce cas, l'algorithme Monte Carlo va calculer la dose dans le milieu. Si on veut combiner ou comparer des résultats obtenus par un algorithme Monte Carlo avec celles d'un TPS classique, il faut appliquer une conversion de l'une vers l'autre. La différence entre dose dans l'eau et dose dans le milieu peut être importante, principalement dans l'os et dans l'air.

La première question qui se pose, qui est plutôt philosophique, est dans quel milieu faut-il exprimer la dose ? La référence clef concernant cette discussion est le « point-counterpoint » entre Helen Liu et Paul Keall [68]. Les arguments pour et contre la conversion de la dose dans le milieu vers la dose dans l'eau (repris dudit article) sont résumés comme suit:

Pour la dose dans l'eau :

- La dose dans l'eau est plus compatible avec les résultats cliniques historiques
- La dose dans l'eau est plus compatible avec les données biologiques (TCP/NCTP)
- La dose dans l'eau est plus compatible avec l'étalonnage de la machine, car tous les protocoles de dosimétrie sont basés sur la dose dans l'eau

Pour la dose dans le milieu

- La dose dans le milieu est plus compatible avec les effets biologiques
- La conversion entre la dose dans le milieu et la dose dans l'eau introduit une incertitude trop importante, et on perd l'avantage d'utiliser un algorithme de calcul de dose précis comme le Monte Carlo

La plupart de ces arguments sont très faibles ou même faux :

Pour la dose dans l'eau :

- Historique clinique déterminé par D_{eau} : Sans le savoir les algorithmes basés sur le collapsed cone (Ahnesjö [69]) ont la capacité de fournir la dose dans le milieu comme c'est montré dans le paragraphe 1.3 de ce chapitre.
- D_{eau} plus compatible avec TCP/NTCP : oui, mais on peut convertir ces données. En plus, l'incertitude des valeurs de TCP/NTCP est très élevée.
- D_{eau} plus compatible avec l'étalonnage : De nouveau on peut facilement convertir ces données.

Pour la dose dans le milieu :

- D_{med} plus compatible avec les effets biologiques : l'article de Walters et al [70] montre clairement que, même dans l'os, D_{eau} est plus corrélée avec la dose calculée dans un micro-CT parce-que les cellules radio sensibles dans l'os sont en fait équivalent eau.
- La conversion entre la dose dans le milieu et la dose dans l'eau introduit une incertitude trop importante, et on perd l'avantage d'utiliser un algorithme précis comme le Monte Carlo : Cet argument est complètement faible, parce qu'il faut en tout cas convertir si on veut combiner ou comparer les données. Et en plus, comme c'est montré dans l'article ci-dessous, ces incertitudes ne sont pas tellement élevées.

En résumé, il n'y pas d'arguments importants pour décider comment rapporter la dose. Il faut simplement faire un choix (presque arbitraire) et être cohérent après. En tout cas, l'élément le plus important est la consistance. Si on veut cumuler les données dans des études multicentriques ou localement dans un service de radiothérapie, il faut appliquer des objectifs et des contraintes de dose qui sont compatibles avec la façon dont le TPS exprime la dose.

Les erreurs dans la littérature scientifique autour de ce sujet sont la conséquence d'une mauvaise interprétation de la dose dans l'eau. En principe il y a deux « doses dans l'eau » :

- La dose dans l'eau comme calculée par les TPS, en modélisant le patient comme de l'eau avec densité électronique variable
- La dose dans l'eau obtenue par conversion de la dose dans le milieu (calculée avec un système de MC), utilisant le rapport des pouvoirs d'arrêts : en principe ça donne la dose dans une petite sphère d'eau centré dans le voxel du medium : « dose dans l'eau dans l'os »

A cause du fait que ces deux quantités sont mélangées, toute la discussion est déraillée. Biologiquement c'est la dose dans l'eau dans l'os qu'on veut (comme démontré dans [70]), mais les TPS n'ont jamais calculé cette quantité. Plus de détails sont donnés dans l'Article 6.

Un autre argument qui peut jouer aussi est l'algorithme de calcul de dose utilisé pour l'optimisation du plan : quand la dose est exprimée en dose dans le milieu, l'optimisation est moins robuste, principalement quand il y a de l'air ou de l'os dans le PTV. Dans ce cas le système d'optimisation va artificiellement booster la fluence des photons dans ces régions (parce qu'il y a moins d'atténuation dans l'os et dans l'air que dans l'eau pour la même densité électronique. Ceci est dû, comme nous l'avons vu au chapitre 3, à la concentration d'hydrogène inférieure dans ces matériaux). Si le patient bouge pendant le traitement ou s'il n'est pas parfaitement positionné, la dose délivrée dans ces zones sera trop élevée.

1.2 Article 6

Dose to medium or dose to water: solely a matter of mass energy attenuation coefficients

F. Crop[1], H. Palmans[2], I Kawrakow[3], E. Sterpin[4,5] and N. Reynaert[1]

1. Centre Oscar Lambret, Lille, France

2. National Physics Laboratory, Teddington, UK

3. Viewray

4. Katholieke Universiteit Leuven, Department of Oncology, Laboratory of Experimental Radiotherapy, Leuven, Belgium

5. Université catholique de Louvain, Institut de Recherche expérimentale et Clinique, Center of Molecular Imaging, Radiotherapy and Oncology, Brussels, Belgium

Abstract

Purpose: The aim of the current paper is to demonstrate that conversion factors between dose to medium and dose to water for photon beams should be based on mass energy attenuation coefficients.

Methods: A theoretical explanation is introduced establishing the inadequacy of stopping power values when converting dose to medium to dose to water. Monte Carlo calculations are performed in a simple bone phantom, validating the theoretical model. A bone slab is modeled and calculations are performed in bone and in water having the same electron density as bone. Special attention was paid on the importance and the range of the interface effects. Calculations were performed for MV photons beams but MV electron beams are also considered.

Results: The Monte Carlo simulations clearly confirm the theoretical model. The dose to medium to dose to water conversion problem cannot be considered as a cavity problem as the composition of all voxels is modified simultaneously, leading to large electron fluence differences. For photon beams, the secondary electron fluence is modified by two effects. On one hand, fewer electrons are generated in bone because of the lower attenuation coefficients compared to water with the same electron density, which tends to increase the secondary electron fluence in water compared to bone. On the other hand, the range of these secondary electrons is larger in bone than in water with bone density which leads to an inverse effect. The first effect is defined by the ratio of mass attenuation coefficients; while the second by the ratio of stopping powers which is compensating the stopping power ratio present in the formal equation of the ratio of dose to water to dose to medium. Only at interfaces, stopping power ratios are playing a role.

Conclusion: As the range of the interface effects is less than commonly used voxel sizes for MV photon beams, one can in general use the ratio of attenuation coefficients to convert dose to medium to dose to water. As a direct consequence, uncertainties on stopping powers (due to uncertainties on the ionization potential of tissues) have negligible impact on the dose to medium to dose to water conversion.

Keywords: Dose to water, Monte Carlo, dosimetry

I. Introduction

For a couple of decades, the discussion regarding dose to medium/dose to water has led to the introduction of a number of arguments in favor of one or the other [1]. It is generally accepted that conventional treatment planning systems (TPSs) have provided dose to water with variable electron density. With the introduction of Monte Carlo algorithms in clinical routine (for QA purposes or for actual treatment planning), the question regarding a shift of paradigm towards dose to medium (as provided by a Monte Carlo dose algorithm), led to an important scientific debate [1]. This debate generally addresses two issues: 1) what is the best quantity to score (dose to medium or dose to water) with respect to the biological effect of radiation? ; 2) how can we convert accurately dose distributions to ensure consistency between previous and more recent dose calculation algorithms [2][3][4][5]?

Several arguments were developed in favor of one quantity or another and were summarized in a point – counterpoint discussion [1]:

In favor of dose to water:

- More consistent with clinical practice
- More consistent with TCP/NCTP data
- More consistent with standard dosimetry and calibration of the machines

In favor of dose to medium:

- Dose to medium is closely related to biology
- The conversion factors introduce a large uncertainty

The discussion has regained importance because of the introduction of Acuros XB (Varian) in clinical routine that has the capability to compute dose to medium or convert it to dose to water using stopping power ratios as suggested by Siebers et al [2], which is the generally accepted method. A recent paper of Andreo [4] revisited the topic. The most important message in this paper was the fact that the stopping powers of tissue have a large uncertainty because of the addition rule (Bragg rule [6]) for determining the ionization potential. This ionization potential is well known for water, but this is not the case for tissues. Because of that, conversion from dose to medium to dose to water should be avoided.

An important aspect to realize is that in this debate three quantities are being compared, which has led to wrong interpretations:

- Dose to medium as calculated by Monte Carlo algorithms
- Dose to water with variable electron density as calculated by “conventional TPSs”
- Dose to water converted from dose to medium using stopping power ratios: provides dose to water in the water material of a cell in an otherwise non-water medium

As we will demonstrate in the current paper, the two quantities, providing dose to water are not identical. We will go into the conversion method regarding dose to medium and dose to water with variable electron density. Our main aim is to obtain consistency between "conventional" and "modern" TPS dose calculations. The discussion on the best quantity to score with respect to biology is not considered in the current paper. In the remaining of this paper, dose to water will be considered as the dose provided by modeling the patient as water with variable electron density.

II. Materials and Methods

II.1 Theoretical model

The conversion of dose to medium to dose to water is formally defined by equation 1, using the definition of dose as the integral over the (secondary) electron fluence and the stopping powers of the corresponding media (water and tissue).

$$\frac{D_w}{D_m} = \frac{\int \phi_w(E) \left(S/\rho \right)_w dE}{\int \phi_m(E) \left(S/\rho \right)_m dE} \quad (1)$$

In the model used by Siebers et al [2], a voxel is considered as a small cavity. Thus the secondary electron fluence is considered as being equal and the dose ratio becomes a ratio of stopping powers. This is an acceptable approximation when dose to water is considered as the dose in a very small water volume inside a non-water voxel. But, as stated above, this is not how a conventional TPS calculates dose. Another conversion method should be considered for an actual comparison between conventional TPS results and the dose to medium provided by a Monte Carlo algorithm. Taking into account the voxel sizes currently used in radiotherapy treatment planning, and the energy of the secondary electrons, voxels cannot be considered as small cavities. As shown in the paper of Siebers et al, but also by Andreo [4] the average energy of the secondary electrons is around 300 keV. These electrons have a range of less than 1 mm in bone (so smaller than the voxel size, even when using a 1 mm resolution). So, certainly for bone, these voxels can almost be considered as "large cavities" where one should apply a ratio of mass energy absorption coefficients instead of stopping powers. In general, depending on the voxel size, the voxels can be considered as intermediate cavities (Burlin cavity theory [7]). Furthermore, it can even be stated that this conversion problem has nothing to do with cavity theory. The cavity approach would be correct if only the composition of a single voxel is modified, without changing the surrounding voxels. And this is obviously not what we are doing when performing the conversion. The "dose-to-water" as computed by multiplying dose to medium with the stopping power ratio has in fact a very specific meaning that has no established record in past clinical practice. It is the dose absorbed by a small cavity of water

placed into a medium of different composition. The potential biological relevance of this quantity has been addressed by other authors [5]. However, we are interested in practice in the conversion between a Monte Carlo dose distribution that is calculated by taking into account the tissue composition of all voxels, and a "conventional TPS" calculation, modeling all voxels as water with variable electron density. In other words, we should be converting the composition of all voxels at once. In that case the impact on the secondary electron fluence in equation 1 is becoming very important. To introduce our theoretical model we will focus on the case of bone (the only relevant material regarding the conversion, as will be argued further). In a certain region (consisting of a large number of voxels), the electron fluence is influenced by two competitive processes:

- Because of the different mass attenuation coefficients, more electrons will be generated in water (having the electron density of bone) than in bone. This difference can be calculated using the ratio of mass attenuation coefficients.
- Because of the differences in mass stopping powers, the range of these secondary electrons is higher in bone than in water (again having the electron density of bone). This difference can be approximated by using the ratio of mass stopping powers.

Combining these arguments leads to equation 2:

$$\frac{\phi_w}{\phi_m} \approx \frac{\left(\mu_{en}/\rho\right)_w}{\left(\mu_{en}/\rho\right)_m} \frac{\left(S/\rho\right)_m}{\left(S/\rho\right)_w} \qquad (2)$$

When we introduce equation (2) in equation (1), the stopping powers are eliminated and the ratio is given by a ratio of mass energy absorption coefficients (this ratio can be considered equivalent to that of the mass attenuation coefficients for the studied energy range). This model is only valid in voxels where a transient electronic equilibrium exists. At interfaces, things are more complicated. For example, when considering the interface between water and bone, the electron fluence at the entrance of this bone region will be identical in bone and water. In that case the dose ratio equals the stopping power ratio. Uncertainties on stopping powers because of an uncertain ionization potential will only influence the conversion at these interfaces. But, as the secondary electron energy is relatively low (on average 300 keV for a 6 MV beam), these interface effects will only be important at a limited distance from the interface. One can expect that photon and electron backscatter influences the dose ratios as well, even in the water region, just before entering the bone slab.

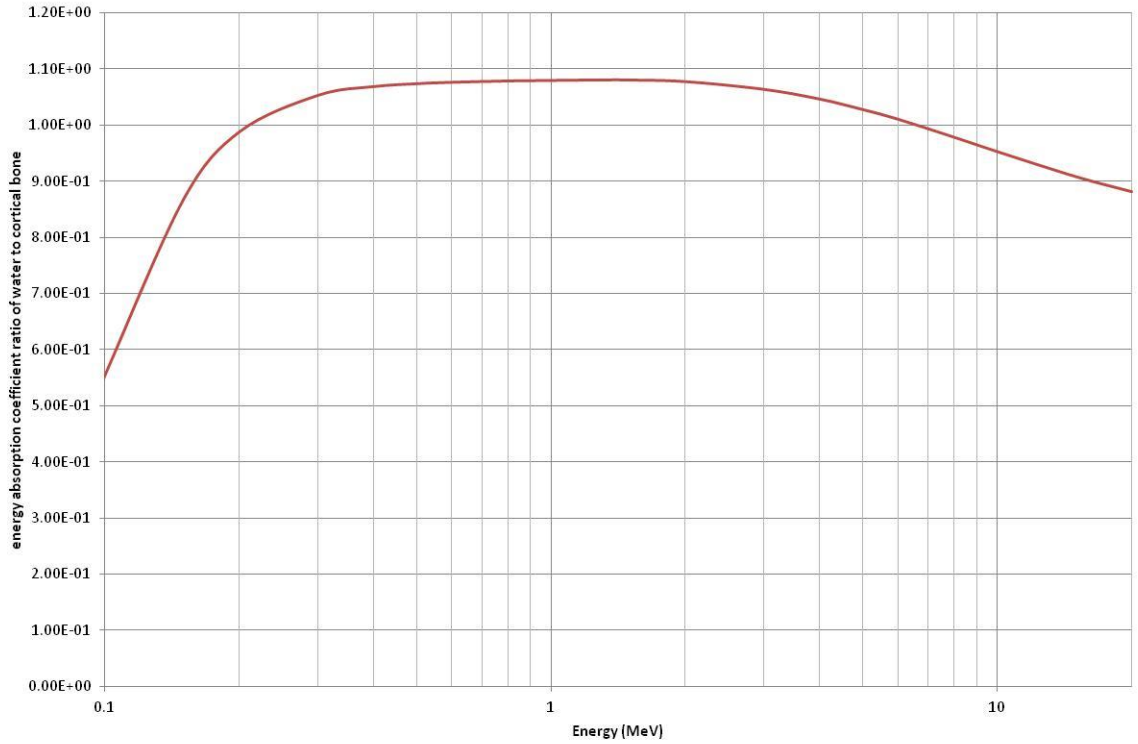


Figure 1: ratio of attenuation coefficients. The ratio for a 6MV averaged over the cyberknife spectrum is 1.06, for a 20MV spectrum this ratio equals 1.01. (Data from NIST database)

II.2 Monte Carlo simulations

All simulations were performed using EGS++ [8]. The geometry is depicted in figure 2.

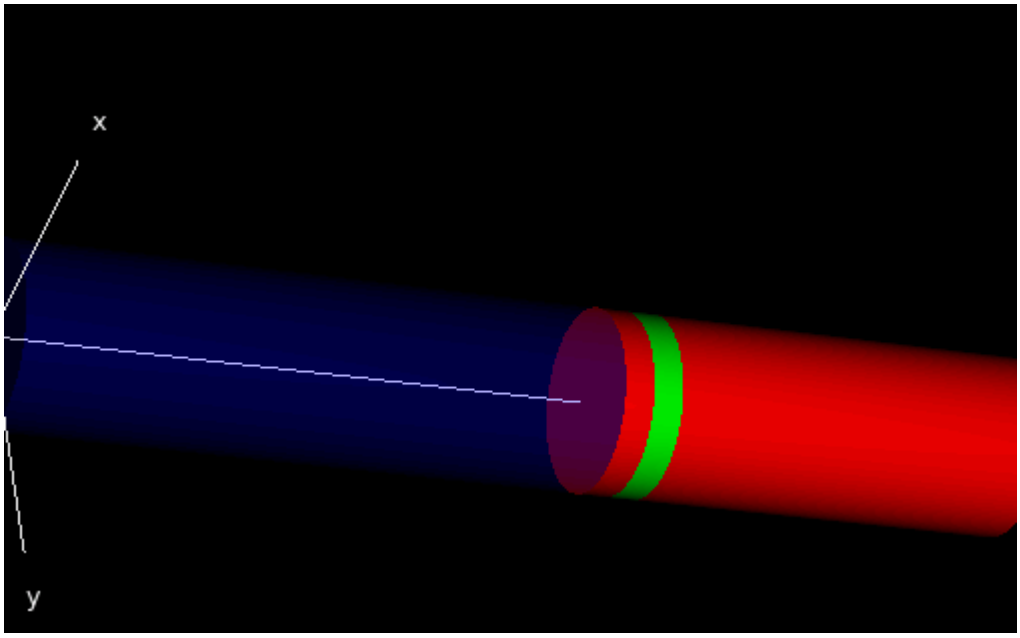


Figure 2: phantom geometry, in blue: air, red: water and green: cortical bone.

A bone slab of 2 cm thickness between a 2 cm water slab and a 26 cm water slab was modeled. A Cyberknife™ 6-cm collimator MC model was used, although the results are general for whatever 6 MV beam. To also consider a high energy beam, a 20MV spectrum (clinic with flattening filter) was taken from the EGSnrc database. PDD's were calculated in a small cylinder of 0.5 cm diameter through the geometry taking into account the different material compositions, providing dose to medium. The resolution was increased near the entrance and the exit of the bone slab to determine the interface effect. Different materials were created by varying a) material density, b) ionization potential and c) material composition (water/bone) as summarized in Table 1.

The electron density was considered for the generation of all materials. Table 1 shows the different material compositions. First of all, the respective density correction files, taking into account both ionization potential and electron density, were generated from the ESTAR [9] database. These density correction files and ionization potential were subsequently used in the generation of the PEGS4 [10] data. Bone composition follows the ICRU 44 [11] definition, with a standard mass density of 1.92 g/cm^3. In order to obtain equivalent electron density, the mass density of water is defined as 1.78 g/cm3. These density correction files, together with the indicated ionization potential were applied during the creation of the respective PEGS4 files.

	Material composition	Ionization potential (eV)	mass density (g/cm^3)
H2O_IPH2O_DENSH2O	water	75	1
H2O_IPB_DENSH2O	water	112	1
BONE_IPB_DENSB	bone	112	1.92
BONE_IPH2O_DENSB	bone	75	1.92
H2O_IPB_DENSB	water	112	1.78
BONE_IPH2O_DENSH2O	bone	75	1.078
H2O_IPH2O_DENSB	water	75	1.78
BONE_IPB_DENSH2O	bone	112	1.078

Table 1: Description of the different simulations performed, combining material composition, ionization potential and mass density.

II. RESULTS

III.1 Monte Carlo simulations

The obtained PDDs for four different simulations (bone modeled as water with electron density of bone and ionization potential of bone, water with bone electron density and ionization potential of water, bone and bone with ionization potential of water) are shown in

figure 3. In figure 4 the ratio of dose to water to dose to bone is shown. Inside the bone slab the ratio between dose to water and dose to bone is on the order of 1.05-1.06 which corresponds to the ratio of mass attenuation coefficients. At the interfaces the ratio is a result of the effects due to both the different ionization potential and material composition (see section III.2).

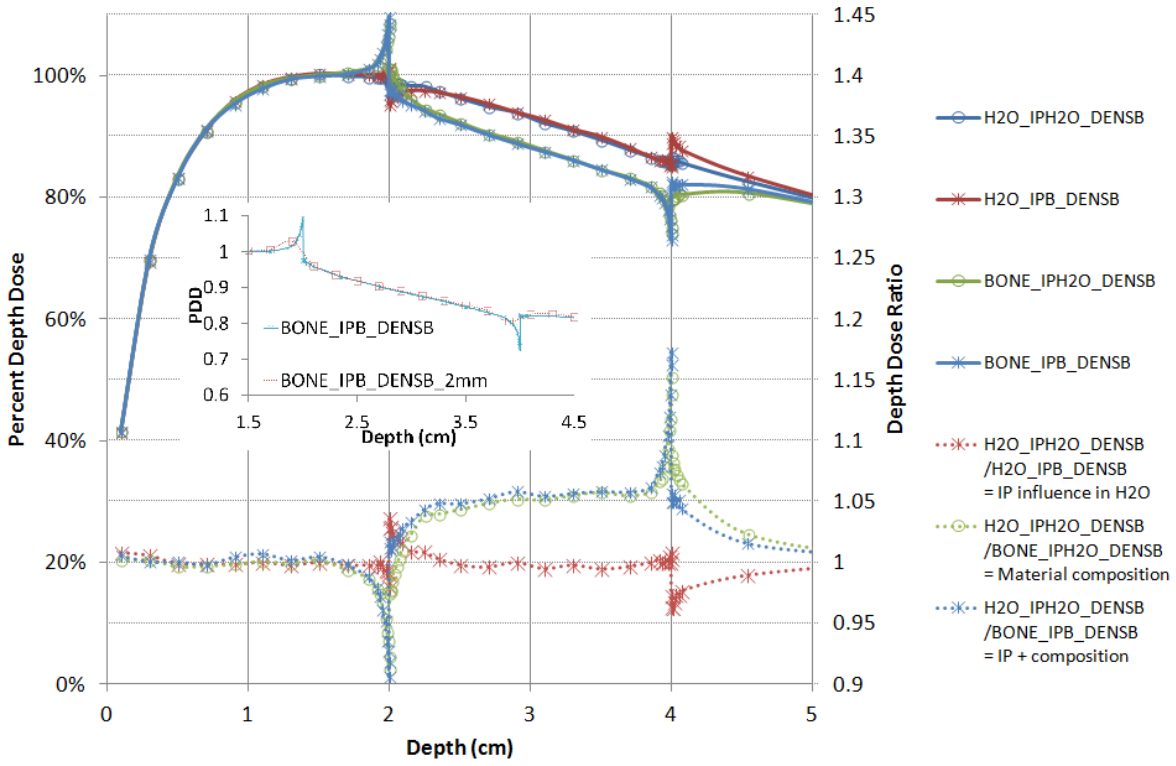


Figure 3: 6 MV depth dose curves for different material compositions (left axis). The dashed lines (right axis) show the different ratio's between the depth dose curves. When using 2 mm voxels, we see the interface effect being smoothed out (inset)

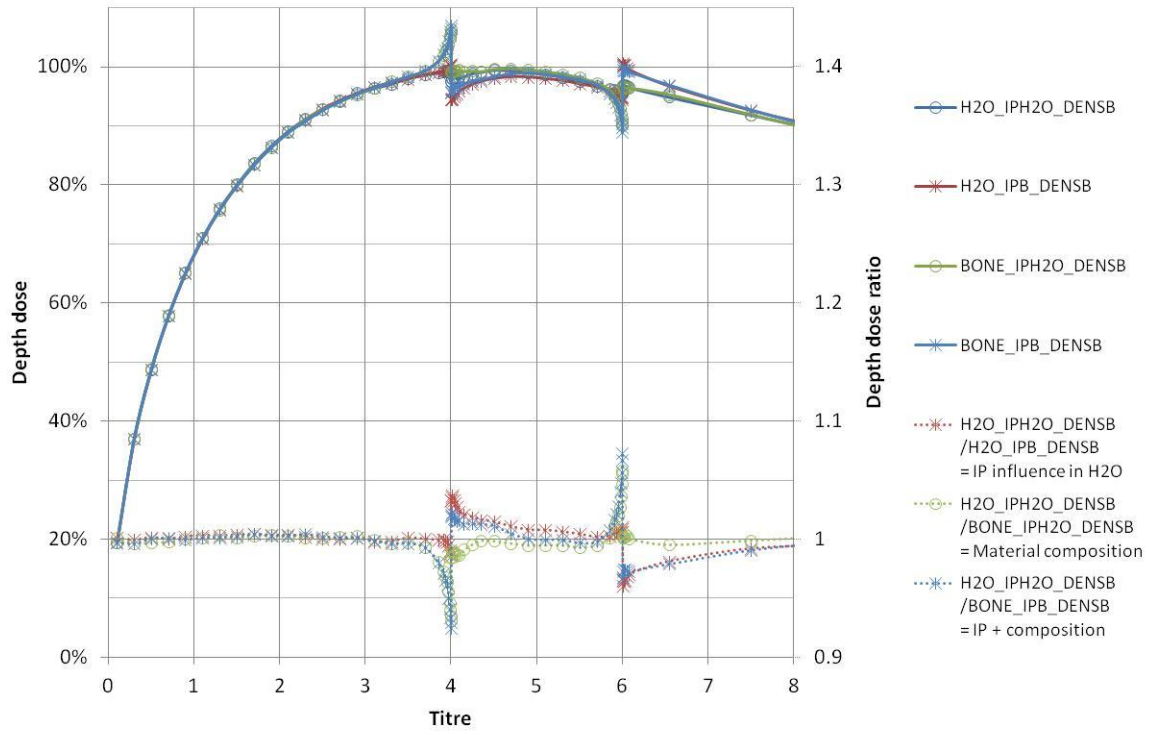


Figure 4: 20 MV depth dose curves for different material compositions (left axis). The dashed lines (right axis) show the different ratio's between the depth dose curves.

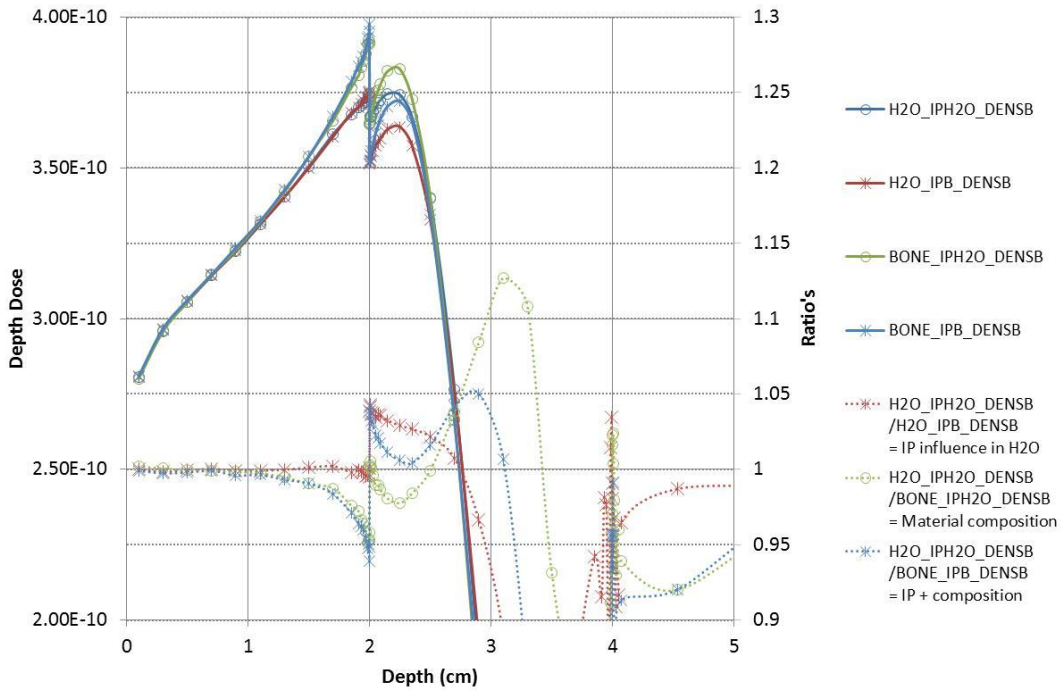


Figure 5: Depth dose curves and associated ratio's for 9 MeV electron beam. Water material H2O_IPH2O_DENSB was modeled using the actual mass density of 1.92 g/cm3 for electron beams

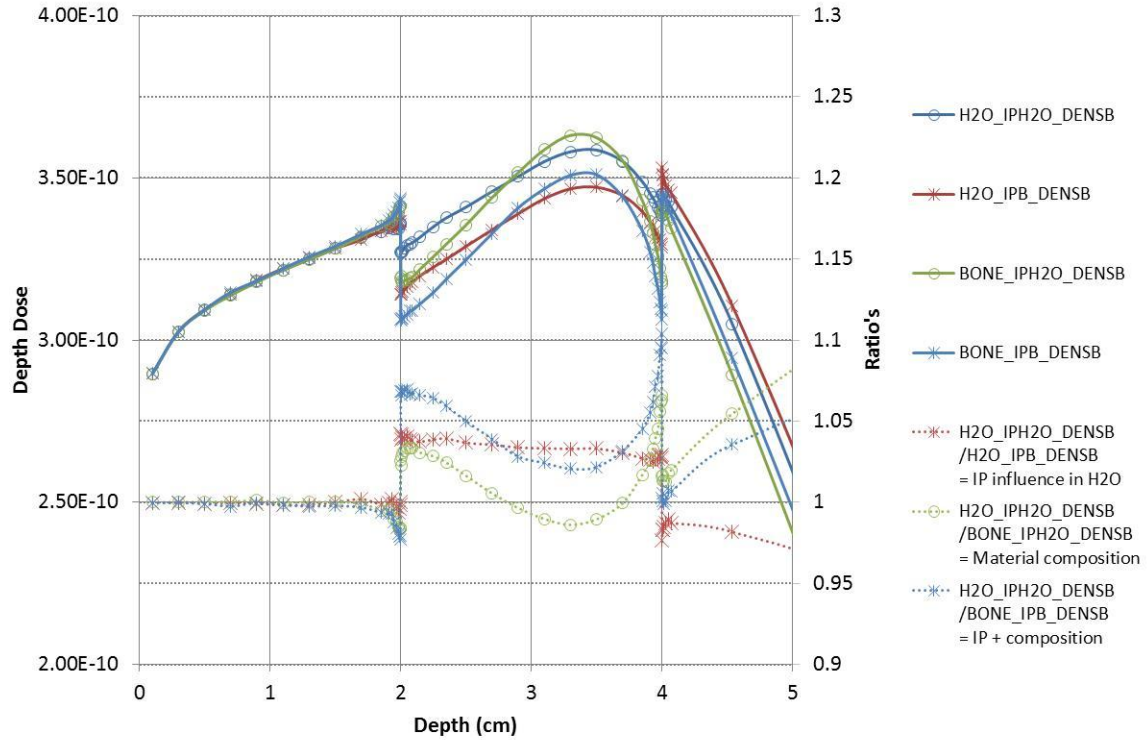


Figure 6: Depth dose curves and associated ratio's for 18 MeV electron beam. Water material H2O_IPH2O_DENSB was modeled using the actual mass density of 1.92 g/cm3 for electron beams

III.2 Interface effects

Different origins of interface effects can be distinguished as:

a) Ionization potential influence

Figure 2 illustrates the effect on the PDD's by changing the ionization potential of the material. Whatever the material, the same ratios at the interface due to the ionization potential change are obtained. The presence of a material with a higher ionization potential, but the same composition, results thus in a higher dose in the water region: both leading and trailing edge of the standard water regions.

The dose in the higher ionization potential material is lower on both trailing and leading edge. This accounts for about 4% dose difference on the downstream edge of the interface. The absolute behavior at the interface is however also influenced by the composition (see c).

b) Density influence

Artificially raising the water density leads to higher attenuation but no significant interface effect.

c) Material composition influence

There is a strong influence of the material composition on the interface effect. Only the composition of the material is changed when determining the ratio H2O_IPH2O_DENSB over BONE_IPH2O_DENSB (green circles in figure 3). This leads to the before mentioned 5-6% difference in dose to medium, but also to important interface effects. The material composition thus leads to an additional 12% difference in the leading edges of the materials.

III. DISCUSSION

The Monte Carlo simulations clearly confirm the simple theoretical model introduced in the current paper. The dose ratio within the bone slab is actually determined by the ratio of mass attenuation coefficients. The stopping power ratio is effectively cancelled by taking into account the modification of the secondary electron fluence. Consequently, modifying the ionization potential in that region has no influence at all on the dose ratios.

There is also an interface effect. In the conventional conversion method, we would never have modified the dose in the water regions at the interface as the composition is not modified. This seems to be caused by back-scatter of low-energy secondary electrons, or lack of electron equilibrium in the exit region of the bone region. At the interfaces, the impact of uncertainties on the ionization potential also becomes larger, as the dose ratio is influenced by the stopping powers in that region.

The interface effect is governed by the material composition. Electron backscatter is based on a) the Coulomb force with the nucleus, and thus increasing with the mean Z of the respective compound but also b) multiple scattering [12]. This interface effect dependence on electron energy is twofold: on the one hand, high energy (secondary) electrons will penetrate further back, but on the other hand, high energy (secondary) electrons will tend to backscatter less. This can be seen in figures 5 and 6. This interface effect is thus a combination of a) material composition and b) electron energy. Because of the limited range of the secondary electrons of photon beams (having an average energy below 300 keV) these interface effects are only significant in a couple of mm around the interface, but due to the relative low energy, the relative dose level in the small interface is higher than for pure electron beams. In current treatment planning calculations, one hardly ever uses voxel sizes below 1 mm. The interface effect can also be seen in the results of Ma et al [3], but as they used 2 mm voxels, the effect was smoothed out (see inset figure 3). In clinical practice these interface effects may be ignored and as a general dose conversion, the ratio of mass attenuation coefficients can be used in all voxels. When calculating using a higher resolution, or when requiring a higher precision, the dose conversion, which is a balancing between the stopping power ratio and the ratio of mass attenuation coefficients should be determined on the fly for every electron individually. The solution that is currently accepted as the general truth, using stopping power ratios fails to establish a meaningful and accurate link between ‘conventional’ and Monte Carlo dose engines. Or in other words, dose to water obtained by

a ‘conventional’ TPS has nothing to do with dose to water obtained by converting dose to medium using a ratio of stopping powers.

Another point is the conversion of dose to water to dose to medium for electron beams. In figures 5 and 6 (18MeV figure) we see that in the bone layer between 2 and 4 cm, the conversion is not consistent, considering ionization potential or material composition. As indicated by Ma et al [3], this is due to the changing fluence and stopping power ratio with depth and thus energy. However, the interface effect remains, albeit slightly different compared to photons. This illustrates that the interface effect for photon beams is caused by the secondary electrons (backscatter).

We focused purely on bone in the current paper as this is the only medium where the actual dose conversion problem is relevant. Only in air and bone is the mass energy absorption ratios between dose to medium and dose to water relatively large (around 5-6 % for a 6 MV beam). But air should be handled differently. When a PTV contains air cavities (e.g. the head and neck region), one should evaluate DVHs with the air removed as we are not interested in the dose in air. The same is true for trachea, intestines, rectum … In lung, air is saturated with hydrogen, and even low density lung tissue should be modeled using the cross sections of soft tissue.

Current paper focuses purely on the dose conversion method. It does not provide an answer on the question whether we should use dose to medium or dose to water. Most of the arguments enumerated in the introduction section are either wrong, or at least very weak:

- Dose to water is more consistent with clinical routine: is this actually the case? The collapsed cone algorithm developed by Ahnesjö et al [13] is certainly able to provide dose to medium if the attenuation coefficients of the different media are explicitly used. This is for example the case in Masterplan^{TM} (Nucletron). So, without realizing, a part of clinical routine is effectively based on dose to medium.
- Dose to water is more consistent with TCP/NTCP data: yes, but what is the quality of these data? Is the precision of TCP data higher than the 5-6 % conversion factor between dose to water and dose to medium? And why would it not be possible to convert the TCP to dose relation from dose to water to dose to medium, instead of converting all MC treatment plans for the next decades?
- Dose to water is more consistent with the calibration of the machine: yes, but the conversion is easily made.
- Dose to medium is closely related to biology: the paper of Walters et al [5], clearly demonstrated that, even in bone, dose to water converted using stopping power ratios (dose to water in a cavity) is more closely related to biology, because the radiosensitive cells in bone are water-like. However, this dose to water computation method has no established clinical record as clinical routine data is mostly based on dose to water with scaled electronic density (conventional algorithms) and, to some extent and without realizing it, dose to medium (Ahnesjö et al [13] algorithm in

Masterplan™ (Nucletron)). If the quantity dose to water in a cavity is what would be selected, than we need to convert both dose to medium (obtained using Monte Carlo) using stopping power ratios and dose to water (obtained using a conventional TPS) using a combination of ratios of stopping power and mass energy absorption coefficients. In that case, as demonstrated in the paper of Andreo et al [4], the uncertainty on stopping powers become very important.

- Converting dose to medium to dose to water introduces a large uncertainty: this is only partially true as illustrated by the results in current paper. As the conversion is based on mass attenuation coefficients instead of stopping powers, the uncertainty is limited. Anyway, if we decide to not convert dose to medium then we will have to convert all doses provided by TPSs that calculate dose to water. So in any case, we need to convert when combining results obtained with different TPSs.

A recent paper has demonstrated that dose to medium optimization presents intrinsic incompatibilities with the PTV concept [14]. However, this is not entirely specific to the dose to medium/dose to water debate as this issue might be solved using a robust optimizer.

Consistency is the key. There is no real argument strong enough to select between dose to water and dose to medium. One should simply ensure that the protocols used in clinical studies and even in clinical routine in general, are consistent with the way dose is calculated. Furthermore, dose is an intermediate parameter introduced to quantify biological effects caused by irradiation, no matter if it is expressed as dose to medium or dose to water. The fact that the ratio of mass energy absorption factors is energy dependent, which is explicitly taken into account when calculating dose to medium, while this is ignored when calculating in water with scaled electron density; is an argument in favor of expressing dose as dose to medium. However, mass attenuation coefficients ratios are close to unity for soft tissues, making the variation of these ratios with beam energy a third order effect with negligible clinical relevance.

IV. CONCLUSION

In the current paper, we have proved that the conversion from dose to medium to dose to water for MV photon beams should be based on mass attenuation coefficients and not on stopping powers. The conversion has nothing to do with cavity theory as one is not converting the composition of one single voxel. Because the complete geometry is modified simultaneously, the secondary electron fluence is largely modified, partly because of a difference in secondary electron generation (governed by mass attenuation coefficients) and partly because of a difference in electron range in different media (governed by stopping powers). This leads to a cancelling of stopping powers in the dose ratio and to the fact that uncertainties on stopping powers (or on the ionization potential of tissues) do not influence the dose conversion. Only very close to interfaces, the situation gets more complicated and

the conversion is a combination of backscatter effects due to material composition, the stopping power ratios and the ratio of attenuation coefficients. But this is only valid at the level of small distances smaller than the currently used voxel sizes in treatment planning systems. Furthermore, we demonstrated that there are no real arguments, regarding the choice between dose to medium and dose to water, as long as protocols are consistent with the way dose is calculated.

ACKNOWLEDGEMENTS

We would like to acknowledge Stefaan Vynckier, Jan Seuntjens and Hugo Bouchard for the fruitful discussions on the current topic.

REFERENCES

1. Liu HH, Keall P, Hendee WR: **D[sub m] rather than D[sub w] should be used in Monte Carlo treatment planning**. *Med Phys* 2002, **29**:922.

2. Siebers J V, Keall PJ, Nahum a E, Mohan R: **Converting absorbed dose to medium to absorbed dose to water for Monte Carlo based photon beam dose calculations.** *Phys Med Biol* 2000, **45**:983–995.

3. Ma C-M, Li J: **Dose specification for radiation therapy: dose to water or dose to medium?** *Phys Med Biol* 2011, **56**:3073–3089.

4. Andreo P: **Dose to "water-like" media or dose to tissue in MV photons radiotherapy treatment planning: still a matter of debate.** *Phys Med Biol* 2015, **60**:309–37.

5. Walters BRB, Kramer R, Kawrakow I: **Dose to medium versus dose to water as an estimator of dose to sensitive skeletal tissue.** *Phys Med Biol* 2010, **55**:4535–4546.

6. Bragg WH, Kleeman R: **XXXIX. On the α particles of radium, and their loss of range in passing through various atoms and molecules**. *Philos Mag Ser 6* 1905, **10**:318–340.

7. Burlin TE: **A general theory of cavity ionisation.** *Br J Radiol* 1966, **39**:727–734.

8. Lippuner J, Elbakri I a, Cui C, Ingleby HR: **Epp: A C++ EGSnrc user code for x-ray imaging and scattering simulations.** *Med Phys* 2011, **38**:1705–1708.

9. Berger MJ, Coursey JS, Zucker MA, Chang J: **ESTAR : Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions**. *NIST, http://www.nist.gov/pml/data/star/index.cfm* 2011.

10. Nelson WR, Hirayama H, Rogers DW: **The EGS4 Code System**. *System* 1985(SLAC-265):398.

11. Griffiths HJ: **Tissue Substitutes in Radiation Dosimetry and Measurement.** *Radiology* 1989(October):1989.

12. Scott W: **THICK-TARGET BREMSSTRAHLUNG WHICH INCLUDES MULTIPLE by We Wuyne Scott**. *NASA Tech Note TN D-4063* (August 1967).

13. Ahnesjö a: **Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media.** *Med Phys* 2013, **16**:577–592.

14. Sterpin E: Potential pitfalls of the PTV concept in dose-to-medium planning optimization. *Phys. Med.* 2016, **32(9)**:1103-10

FIGURE CAPTIONS

Figure 1: ratio of attenuation coefficients. The ratio for a 6MV averaged over the cyberknife spectrum is 6%, for a 20MV spectrum this ratio equals 1 %. (Data from NIST database)

Figure 2: phantom geometry, in blue: air, red: water and green: cortical bone.

Figure 3: 6 MV depth dose curves for different material compositions (left axis). The dashed lines (right axis) show the different ratio's between the depth dose curves. When using 2 mm voxels, we see the interface effect being smoothed out (inset)

Figure 4: 20 MV depth dose curves for different material compositions (left axis). The dashed lines (right axis) show the different ratio's between the depth dose curves.

Figure 5: Depth dose curves and associated ratio's for 9 MeV electron beam. Water material H2O_IPH2O_DENSB was modeled using the actual mass density of 1.92 g/cm3 for electron beams

Figure 6: Depth dose curves and associated ratio's for 18 MeV electron beam. Water material H2O_IPH2O_DENSB was modeled using the actual mass density of 1.92 g/cm3 for electron beams

1.3 Le collapsed cone de Ahnesjö peut donner la D_{med}

La superposition convolution en collapsed cone applique les équations suivantes (copié de la référence Ahnesjö et al [69]) :

$$T_E(\mathbf{r}) = (r/r_0)^2 \frac{\mu(E,\mathbf{r})}{\rho(\mathbf{r})} \Psi_E(\mathbf{r}_0) \exp\left[-\int_{\mathbf{r}_0}^{\mathbf{r}} \mu(E,\mathbf{l}) dl \right], \quad (1)$$

et

$$D(\mathbf{r}) = [1/\rho(\mathbf{r})] \int\int\int\int T_E(\mathbf{s}) \rho(\mathbf{s}) h(E,\mathbf{s},\mathbf{r}) d^3\mathbf{s}\, dE, \quad (2)$$

La dose D(r) est calculée avec une convolution du Terma (énergie totale perdue par les photons) avec les kernels (énergie déposée par les électrons et photons secondaires autour d'une « source » de Terma). Le Terma (défini à l'équation (1)) est déterminé par la fluence énergétique des photons, l'inverse carré des distances et le facteur d'atténuation exponentiel. Mais le facteur clef dans cette équation est le facteur d'atténuation μ(E,r) : si on reprend ces équations pour calculer la dose dans l'eau, on obtient :

$$D_{CS,water}(r) = [1/\rho(r)] \int\iiint T_{E,water}(s) \rho(s) h_{water}(E,s,r) d^3 s dE$$

Avec le Terma dans l'eau donné par :

$$T_{E,water}(r) = (r/r_0)^2 \frac{\mu_{water}(E,r)}{\rho(r)} \Psi_E(r_0) \exp[-\int \mu(E,l) dl]$$

Si on compare la dernière équation avec les équations (1) et (2) de l'article, on obtient :

$$D_{CS,water} = D_{CS} \frac{\mu_{water}}{\mu_{medium}} \Leftrightarrow D_{CS} = D_{CS,water} \frac{\mu_{medium}}{\mu_{water}}$$

Et comme nous l'avons démontré dans l'article 6, le rapport des facteurs d'atténuation correspond à la conversion entre la dose dans l'eau et la dose dans le milieu. Alors $D_{CS} = D_{med}$. Si l'algorithme d'Ahnesjö est implémenté, en utilisant des facteurs d'atténuation différents pour chaque milieu, la dose obtenue par le TPS correspond à D_{med}. C'est effectivement le cas pour le TPS Masterplan™ de Nucletron. C'est pourquoi, sans le savoir, une partie des résultats cliniques obtenus dans le passé (avec des TPS « classiques ») est quand-même basée sur la dose dans le milieu.

2. Prescription de dose dans les poumons

2.1 Introduction

Depuis quelques années, on a compris que la qualité de l'algorithme de calcul de dose est extrêmement importante dans les poumons. A cause de la faible densité des poumons, le parcours des électrons secondaires est relativement élevé et il n'y a souvent pas d'équilibre électronique dans la tumeur. Le PTV est constitué du GTV, qui a normalement la densité des tissus mous, et d'une marge qui contient du tissu poumon de faible densité. La dose calculée par un algorithme qui ne tient pas compte du transport des électrons secondaires (un algorithme du type A) sera largement surestimée (comme c'est montré à la figure 19 du chapitre 2) dans cette marge. C'est pourquoi, historiquement, on a toujours surestimé la dose dans les poumons. L'erreur faite dépend largement de la taille de la tumeur, de la localisation dans les poumons (centrale ou périphérique) [80] et de la densité des poumons qui est souvent très faible en cas de maladie. La question se pose alors de la description de dose. Un plan de traitement optimisé avec un algorithme approximatif et, une prescription de 3x20 Gy sur 95% du PTV verra sa dose diminuer lorsqu'il sera recalculé en Monte Carlo. Dans l'exemple précédent, on n'obtiendra peut-être que 40 Gy. Même si les résultats cliniques sont positifs, cela pose des problèmes. Si la différence entre les algorithmes de type A et B était systématiquement la même, il n'y aurait pas de problème. Dans ce cas, on prescrirait systématiquement 40 Gy calculé en Monte Carlo (qui correspondrait avec la dose qu'on a systématiquement donnée dans le passé). Cependant, la différence entre les deux algorithmes dépend, comme c'est expliqué plus haut, de la taille de la tumeur, de la densité poumons, et de plusieurs autres paramètres… Il faut donc en tout cas prescrire en utilisant les résultats des calculs Monte Carlo pour traiter les patients de façon cohérente. La solution qui est maintenant la référence est d'utiliser un algorithme du type B, comme le Monte Carlo ou le collapsed cone, et de prescrire des doses différentes en fonction du volume de la tumeur [80]. De cette façon la dose est mieux standardisée et correspond quand-même avec la routine clinique du passé.

Il y a, par contre, un deuxième problème qui fait le sujet de l'article 7. Ce problème est lié à la densité hétérogène du PTV dans les poumons : sur quel volume faut-il prescrire si on utilise un algorithme de type B pour calculer la dose ? Si on prescrit sur 95% du PTV, la dose au niveau de la tumeur même (l'effet biologique) dépend de la taille de la tumeur, et des mêmes paramètres déjà évoqués plus haut. Alors, même si on calcule correctement la dose et qu'on essaie de prescrire correctement, on continue à avoir une grande variabilité de la dose à la tumeur (GTV) entre les patients. L'article 7 prouve qu'il faut absolument prescrire sur le GTV au lieu du PTV pour traiter les patients de façon cohérente.

2.2 Article 7

Lacornerie *et al. Radiation Oncology* 2014, **9**:223
http://www.ro-journal.com/content/9/1/223

RADIATION ONCOLOGY

METHODOLOGY Open Access

GTV-based prescription in SBRT for lung lesions using advanced dose calculation algorithms

Thomas Lacornerie[1*], Albert Lisbona[2], Xavier Mirabel[3], Eric Lartigau[3] and Nick Reynaert[1]

Abstract

Background: The aim of current study was to investigate the way dose is prescribed to lung lesions during SBRT using advanced dose calculation algorithms that take into account electron transport (type B algorithms). As type A algorithms do not take into account secondary electron transport, they overestimate the dose to lung lesions. Type B algorithms are more accurate but still no consensus is reached regarding dose prescription. The positive clinical results obtained using type A algorithms should be used as a starting point.

Methods: In current work a dose-calculation experiment is performed, presenting different prescription methods. Three cases with three different sizes of peripheral lung lesions were planned using three different treatment platforms. For each individual case 60 Gy to the PTV was prescribed using a type A algorithm and the dose distribution was recalculated using a type B algorithm in order to evaluate the impact of the secondary electron transport. Secondly, for each case a type B algorithm was used to prescribe 48 Gy to the PTV, and the resulting doses to the GTV were analyzed. Finally, prescriptions based on specific GTV dose volumes were evaluated.

Results: When using a type A algorithm to prescribe the same dose to the PTV, the differences regarding median GTV doses among platforms and cases were always less than 10% of the prescription dose. The prescription to the PTV based on type B algorithms, leads to a more important variability of the median GTV dose among cases and among platforms, (respectively 24%, and 28%). However, when 54 Gy was prescribed as median GTV dose, using a type B algorithm, the variability observed was minimal.

Conclusion: Normalizing the prescription dose to the median GTV dose for lung lesions avoids variability among different cases and treatment platforms of SBRT when type B algorithms are used to calculate the dose. The combination of using a type A algorithm to optimize a homogeneous dose in the PTV and using a type B algorithm to prescribe the median GTV dose provides a very robust method for treating lung lesions.

Background

The differences between radiotherapy dose calculation algorithms that take into account the electron transport phenomenon inside non-homogeneous environments (category 3&4 or type B) and those that do not (category 1&2 or type A) are well known [1]. The clinical implications of Monte-Carlo dose calculation algorithms [2,3] have been described, as well. However, the differences between type A and type B algorithms for small fields – with a penumbra region that occupies most of the field area – depend on many factors, such as lung density, beam sizes, and the position and size of the target. It is still not clearly understood how to adapt the protocols based on type A algorithms. Xiao et al. [4] concluded their study about heterogeneity corrections and RTOG trial 0236 in 2009 stating *"the information provided by the study will be used for future protocols,"*. But almost 3 years later Li et al. stated, about heterogeneities and the RTOG trial 0813 [5], that *"further studies are expected to establish protocol criteria for MC dose calculations"*. The RTOG trial 0915 requires the use of a type B algorithm with a PTV edge prescription and an isodose between 60 and 90% of the dose maximum. That variability leads to significantly different GTV doses as we will demonstrate. Many articles have pointed out the discrepancies between type A and type B algorithms in this journal [6-8] and in others for many years [9] but they are almost always focused on results for the PTV. We believe that the PTV edge prescription is not a good

* Correspondence: t-lacornerie@o-lambret.fr
[1]Service de Physique Médicale, Centre Oscar Lambret, Lille, France
Full list of author information is available at the end of the article

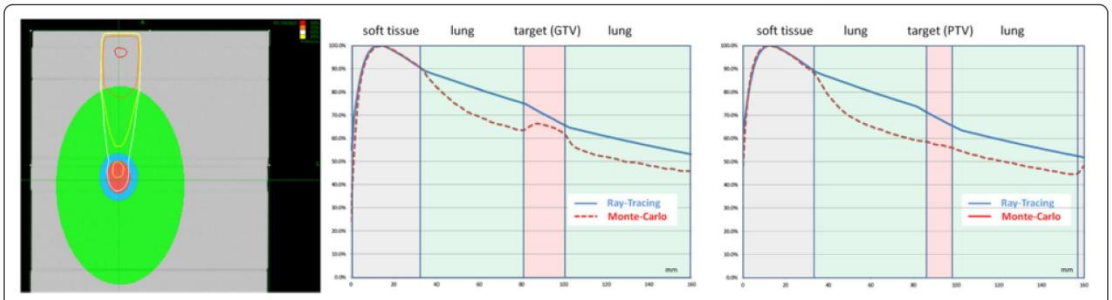


Figure 1 PDD for 30 mm collimator calculated with Ray-Tracing and Monte-Carlo algorithms in phantom with different densities (left) on the beam axis (central) and 1 cm off axis (right), GTV (red, density = 1), PTV (blue), lung (green, density = 0.3).

approach to be consistent between cases and platforms as will be shown below.

Discrepancies between Monte-Carlo (a type B algorithm) and Ray-Tracing algorithms (a type A "equivalent path length" algorithm) were monitored at the beginning of our experience with Cyberknife (Accuray Inc, Sunnyvale, CA) five years ago, although the protocol applied to peripheral lesions for dose prescription was the standard protocol defined by other teams for: 60 Gy in three fractions encompassing the PTV [10,11] calculated with type A algorithms.

Because the Cyberknife has a 6-MV output and no flattening filter, the range of secondary electrons in lung can be up to 4 cm. In Figure 1, differences between Ray-Tracing and Monte-Carlo for a 30–mm beam in a phantom representing a 14-mm peripheral lung lesion is illustrated. Similar effects between any type A and type B algorithm can be observed.

Ray-Tracing overestimates the dose in the lung because it disregards the lack of lateral electronic equilibrium and changes of scattered dose. The rebuild-up in the target observed with Monte-Carlo is ignored by Ray-Tracing. Behind the rebuild-up zone at the centre of the target, Ray-Tracing overestimates the dose because of the absence of scattering in lung tissue. On the edge of the target, in the PTV margin, the differences are even more pronounced because the lack of electronic equilibrium is dominant in this area, and Ray-tracing applies the change of density observed on the beam axis whereas the interface between lung and target is not at the same depth. The same kind of differences between every type A and type B algorithm can be observed: they are maximal in lung at the edge of the PTV (Figure 2). The smaller the lesion, and the beam, the higher the discrepancy between type A and type B algorithms in lung [12]. Because of that, modifying the prescription is more complicated for SBRT than for wide-field RT treating bigger volumes. Although the dose of 60 Gy is overestimated when calculated by a type A algorithm, local control has been very high [13], so this actual dose level should be maintained. Comparison between Pencil Beam and Collapsed Cone (Oncentra Masterplan, V4.1, Nucletron Elekta, Veenendaal, Netherlands) for the Varian Clinac® (Varian Medical Systems Inc., Palo Alto, CA) and comparison of Pencil-Beam and Monte-Carlo dose calculations for Novalis® (Brainlab AG, Feldkirchen, Germany) lead to the same conclusions. Current study will be applied to these treatment platforms.

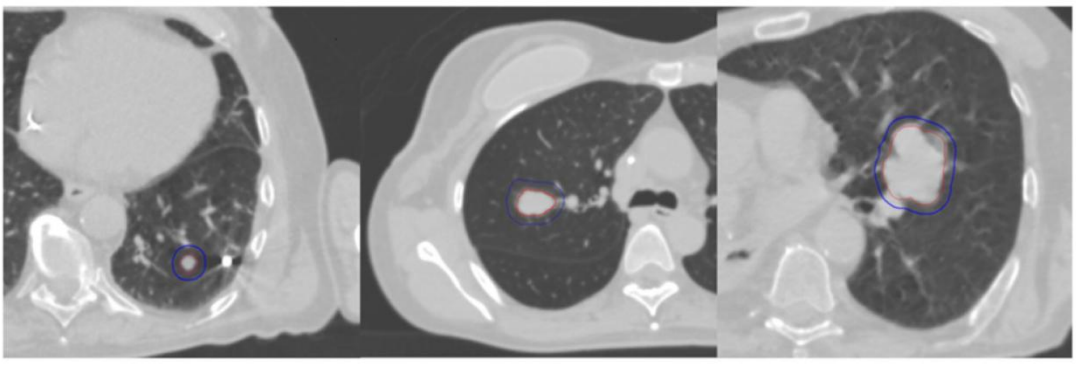


Figure 2 GTV and PTV for the 3 different lesions analyzed.

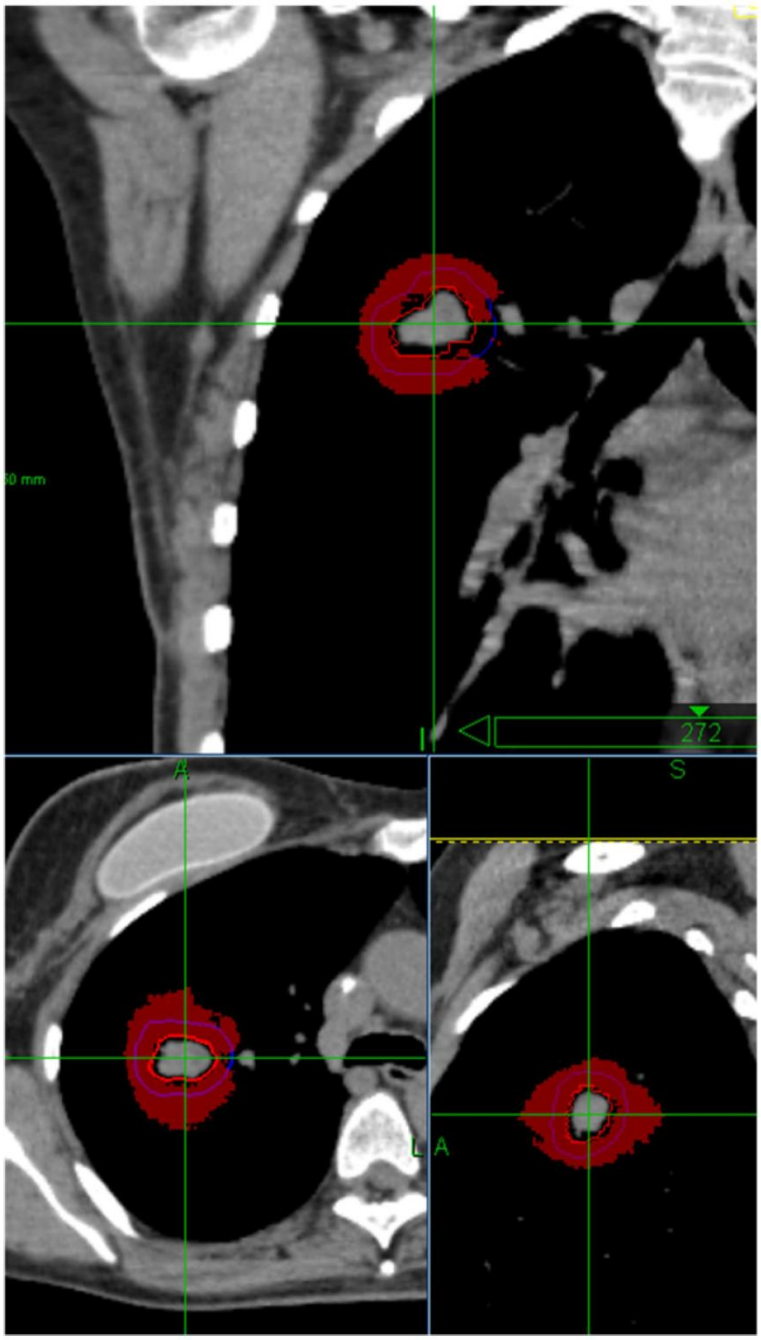


Figure 3 Differences (in red) larger than 15 Gy for case 2 on Cyberknife between Ray-Tracing and Monte-Carlo.

A systematic review for NSCLC stage I SBRT was recently published [14] showing a wide variability in reporting. In the 45 identified studies there were 22 different treatment schemes. One dose level only, probably PTV based, was reported for each study, ranging from 15 Gy to 72.5 Gy in 1 to 12 fractions. In another paper van Baardwijk et al. try to determine the best therapeutic ratio [15] despite the differences between type A and B algorithms considering only dose at the edge of the PTV. The questions in these articles will not be answered correctly when considering only the PTV.

Our goal is to establish a consistent dose prescription and reporting in a heterogeneous environment that is applicable for different SBRT treatment modalities.

Methods

A dose calculation experiment is proposed to discuss the different prescription methods. This exercise is not meant to be an exhaustive demonstration but should primarily provide a number of arguments to put into perspective the current prescription methods.

Three different cases with three different sizes of peripheral lung lesions (<1 cm, ~2 cm, ~4 cm) were examined (Figure 3). Each case was calculated with a type A and a type B algorithm on three different platforms: CyberKnife with Ray-Tracing and Monte-Carlo (Multiplan v.9.0, 1% statistical uncertainty and resolution of 1×1×1 mm), Novalis 6 MV with Pencil-Beam and Monte-Carlo (Iplan v.4.1.2, 0.5% statistical uncertainty and resolution of 2×2×1 mm), and Clinac 6 MV (Varian) with Pencil-Beam and Collapse Cone (Masterplan v4.0, resolution of 2×2×1 mm). The PTV was defined as GTV + 5 mm, for the treatment with gating or with real-time tracking, details of ballistics are given in Table 1. The targets were not enlarged for prophylactic treatment (RTOG 0236 & 0915).

Scenario A: 60 Gy was prescribed to 95% of the PTV using type A algorithms (RTOG 0236). For the Clinac plans, the margin from PTV to TV was 5 mm, as in conventional 3D-CRT, consequently the field sizes were larger than 3×3 cm^2 for all cases. The GTV D98% (near min) and the D50% (median) were compared. Each plan was recalculated with a type B algorithm while keeping the monitor units per beam constant. The D95% of the PTV was calculated and again the D98% and D50% for the GTV were evaluated.

Scenario B: Because type A algorithms overestimate the delivered dose, 48 Gy instead of 60 was prescribed with type B algorithms, renormalizing the plans used in Scenario A, to 95% of the PTV, as proposed in the STARS protocol by the MD Anderson Cancer center. The different doses to the GTV were observed for all combinations of cases and platforms.

Scenario C: Finally, to avoid discrepancies among median GTV doses, a dose of 54 Gy was prescribed to this parameter using a type B algorithm. These plans were first optimized with type A algorithms to ensure a homogeneous photon fluence in the PTV (−5%, +15% of prescription dose), recalculated with type B algorithms and normalized in a way that the GTV D50% equals the prescription dose (i.e. 54Gy).

Because the three lesions were in the centre of lung, the doses to organs at risk – esophagus, ribs, and heart – were very low for all cases; only the normal tissue dose to lung was recorded. A large number of criteria concerning lung toxicity have been published [16]. Most of them were calculated for the current study, yet not all of them are published here for the sake of brevity. The lung volume receiving <10 Gy was presented to show that in all scenarios, the criteria proposed by AAPM report of TG 101 were fulfilled [17].

Results

Scenario A

When prescribing 60 Gy to 95% of the PTV using a type A algorithm, the maximum discrepancies in the median dose to the GTV were moderate among studied platforms: 3.6 Gy (6% of the prescription dose (Dp)), and cases: 4.5 Gy (7.5% of Dp) (Table 2). The maximum dose variance among the nine plans was 6 Gy (10% of Dp), which is common in SBRT in homogenous tissues. For the GTV near-minimum dose (D98%), used as proposed by the ICRU report 83, the maximum differences were comparable: 4 Gy (6.7% Dp) among techniques and 2.4 Gy (4% Dp) among cases. For the two platforms with conventional stereotactic margins, Novalis and CyberKnife, higher median doses were obtained for larger lesions. For the Clinac

Table 1 Description of cases and irradiation techniques

	Case 1	Case 2	Case 3
GTV	1.0 cm^3	3.6 cm^3	32.8 cm^3
PTV	6.4 cm^3	14.5 cm^3	75.0 cm^3
Novalis (Non-IMRT)	10 coplanar and non opposed beams	9 coplanar and non opposed beams	10 coplanar and non opposed beams
Cyberknife (circular collimators and flattening filter free)	62 non coplanar beams	58 non coplanar beams	52 non coplanar beams
Clinac (Non-IMRT) prescription like classical 3DRT with type A, extra-margin from PTV to field edge	5 coplanar and non opposed beams	7 coplanar and non opposed beams	7 coplanar and non opposed beams

Table 2 Prescription and calculation with type A algorithms (scenario A)

PTV D95% = 60 Gy type A									
	GTV D98% type A (Gy)			GTV D50% type A (Gy)			Normal lung receiving less than 10 Gy (cm^3)		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Novalis	61.5	62.0	61.8	61.6	63.2	64.1	3808	3468	2227
Cyberknife	61.6	64.0	62.3	62.8	64.9	67.3	3799	3488	1992
Clinac	60.7	60.0	60.3	62.3	61.3	64.0	3681	3308	1804

using a 5-mm margin from PTV to TV, i.e. using a conventional prescription as for classical 3DRT, the median dose was almost the same for the three cases.

Prescription with type A algorithms and re-calculation with type B algorithms

When a type B algorithm was used, for each platform, the PTV D95% was smaller than the prescription dose according to the type A algorithm (Figure 4). The maximum difference was 19.6 Gy (32.7% Dp) (Table 3). Moreover a large range of doses (40.4–54.4, 23.3% Dp) was observed. For the GTV D98%, the maximum difference between cases was 12.4 Gy (20.7% Dp), the maximum difference between platforms was 3.6 Gy (6% Dp). For the GTV D50%, the maximum difference between cases was 13.2 Gy (22%) and the maximum difference between platforms was 3.2 (5.3%). The differences regarding the volume of lung that received less than 10 Gy for cases 1 and 2 (3% maximum) were small. Type A algorithms overestimated the high dose but the low doses (<10 Gy) are quite similar.

Scenario B

With a prescription of 3x16 Gy to 95% of the PTV using type B algorithms, the maximum differences for median doses of the GTV were: 13.4 Gy (28% of the prescription dose) among the platforms and 11.6 Gy (24.2%) among cases (Table 4). For GTV D98% (near-min) the maximum differences were comparable: 7 Gy (14.6%) (Figure 4) among platforms and 10.1 (21%) among cases (Figure 5).

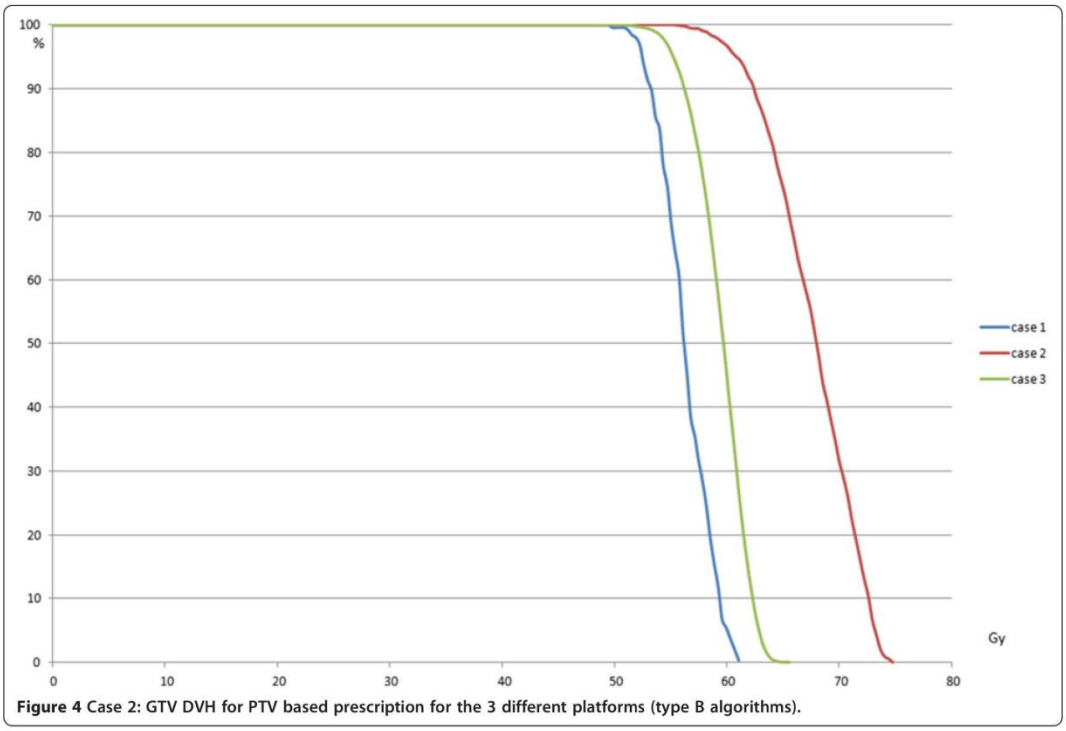


Figure 4 Case 2: GTV DVH for PTV based prescription for the 3 different platforms (type B algorithms).

Table 3 Calculation with type B algorithms with type A algorithms prescription (scenario A, same monitor units)

PTV D95% = 60 Gy type A												
	PTV D95% type B (Gy)			GTV D98% type B (Gy)			GTV D50% type B (Gy)			Normal lung receiving less than 10 Gy (cm^3)		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Novalis	43.7	40.5	52.2	47	47.7	59.4	49.8	53.9	63	3858	3456	2189
Cyberknife	42.7	40.4	49.5	46.3	49.7	55.9	50	57.1	61.5	3824	3537	2090
Clinac	44.6	50.3	54.4	48.7	51.3	58.1	51.4	57	62.1	3757	3259	1817

The reader is kindly referred to the supporting material for more details and additional DVHs (Additional file 1).

Scenario C

As illustrated above (scenario A), for plans prescribed at 95% of the PTV using type A algorithms, a large range of GTV50% values was obtained when recalculated using type B algorithms, Therefore we suggested a prescription dose of 54 Gy to 50% of the GTV when using a type B algorithm.

The maximum differences for GTV D98% were 1.8 Gy (3.4% of the prescription dose) among the platforms and 3.1 Gy (5.8%) among cases (Table 5) Of course with this scenario there are differences for the PTV D95%, the maximum is 9.4 Gy (17% of the prescription dose) among the platforms and 5.3 Gy (9.7% of the prescription dose) among cases.

Discussion

The accuracy of Monte Carlo dose calculation for the studied platforms, Novalis and CyberKnife, has been demonstrated [18,19] using anthropomorphic phantoms [20]. Therefore, it is advisable that advanced algorithms are used to obtain more accurate dose distributions to targets and organs at risk. The question is how to apply the well established prescription protocols that lead to positive clinical results.

On one hand, there are a lot of studies that report excellent local control [11,21,22] with different prescription schemes in which only the prescription dose to the PTV is provided without mentioning the algorithm used. On the other hand, differences between type A and type B algorithms have been observed in many studies [4,5,23,24] without clear consequences regarding the prescription. Hurkmans et al. have proposed for the ROSEL study [25] different dose conformity requirements according to the size of the PTV and the type of the algorithm used. The prescription dose was the same for all lesions based on the PTV, but the heterogeneity in the PTV and the dose constraints for the lung varied according to the size of the PTV. Van der Voort van Zyp et al. [26] have also performed a retrospective analysis of Monte-Carlo calculations on cases treated in their institution. They proposed three dose levels based on the size of the tumor calculated with Monte-Carlo and prescribed to the 95% of the PTV (48 Gy for tumors <3 cm, 51 Gy for tumors of 3–5 cm, and 54 Gy for tumors >5 cm), which can be considered as the most consistent proposition to date.

For PTV based prescription using type A algorithms a small variability for GTV D98% and GTV D50% among cases, and among platforms was obtained. Recalculation of the plans using a type B algorithm shows that, there is a dose escalation according to the size of the GTV with few discrepancies between the platforms (3.2 Gy, 5.2% of the prescription dose). When a type A algorithm is used with a PTV based prescription the dose delivered to the GTV is higher for larger lesions. That may explain the good local control of all the studies performed using type A algorithms. This result, also described in detail by van der Voort van Zyp et al. [26], led to their proposition of the three dose levels.

As the PTV includes a low-density region, lacking electronic equilibrium, the prescription, using type B algorithms, should not be based on this volume. For example, the prescription of 48 Gy to the PTV D95% in case 2 leads to a higher GTV median dose than for case 3 which is a larger lesion. Even if we prescribed like van der Voort van Zyp et al. (18) [26] suggest, i.e., 48 Gy to 95% of the PTV for case 2 and 51 Gy for case 3, the GTV median dose remains inferior to that of case 2 (61.5 Gy vs. 63 Gy for Novalis, 63.3 Gy vs. 67.8 Gy for CyberKnife). If the same

Table 4 Calculation with type B algorithms with prescription on PTV (scenario B)

PTV D95% = 48 Gy type B									
	GTV D98% type B (Gy)			GTV D50% type B (Gy)			Normal lung receiving less than 10 Gy (cm^3)		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Novalis	51.6	56.5	54.6	54.7	63.9	57.9	3824	3352	2219
Cyberknife	52.0	59.0	54.2	56.2	67.8	59.6	3777	3446	2218
Clinac	52.4	49.0	51.3	55.3	54.4	54.8	3723	3294	1927

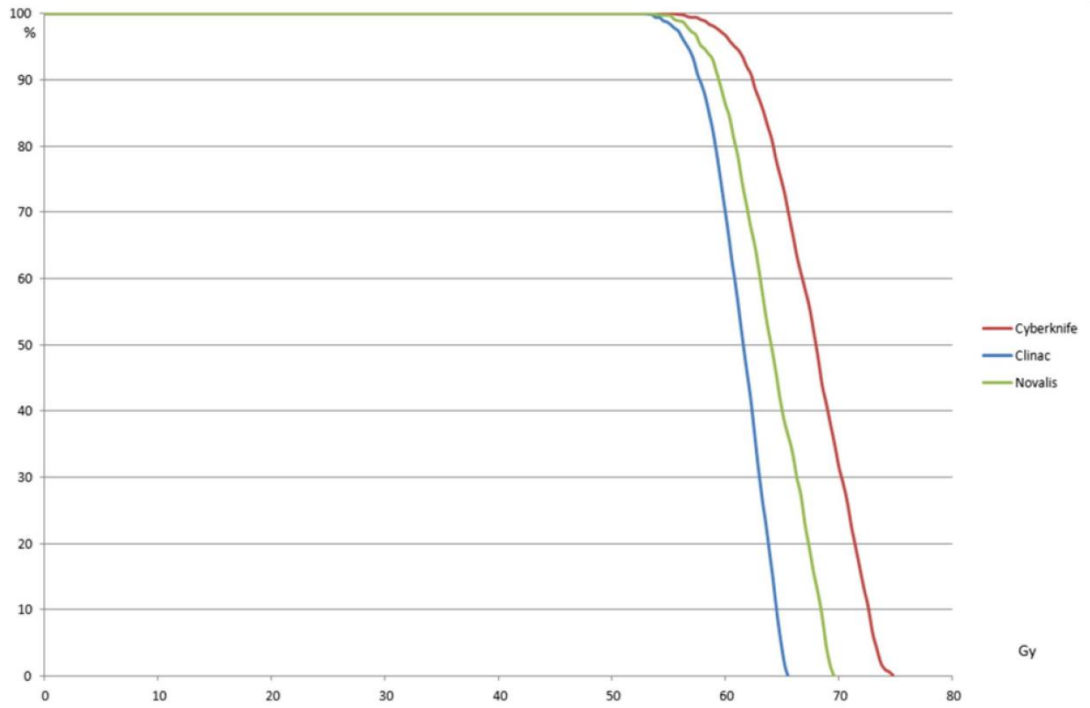


Figure 5 Cyberknife: GTV DVH for PTV based prescription for the 3 cases calculated with Monte-Carlo.

level of dose should be delivered to larger lesions, a PTV-based prescription should not be used.

The PTV is a fictitious volume created to ensure that the absorbed dose to the target equals the prescription dose, taking into account positioning uncertainties. Because of the low density surrounding the GTV, the minimum dose to the PTV is calculated in a low density region whereas the GTV has a density close to 1.

The PTV D95% depends on the lung density around the lesion and does not predict the dose to the GTV (28% variability in the median prescription dose and 14.6% for Dnear-min for the different treatment devices in our exercise). The difference between PTV D95% and GTV dose depends on the size of the lesion, lung density, location, treatment platform and even the operator. Our exercise is not intended to be exhaustive but only to show that one PTV criterion is not enough to characterize the dose distribution. The PTV margin in lung can be considered as the "flash" region margin used for tangential breast field, where part of the PTV is in air, while it is not used for prescription. Furthermore the statistical uncertainty of Monte-Carlo is up to 5% higher (or more for extreme case of low density lung) in the PTV periphery than on the GTV median dose (Figure 6) because there are fewer histories in low density region. With a prescription based on the PTV, the discrepancies between treatment systems are much more substantial with type B than with type A algorithms; 28% vs. 6%.

The best way to avoid these discrepancies among cases and treatment platforms is to prescribe to the GTV median dose. The GTV median dose is the most representative dose, in that the majority of the lesion receives ±5% of

Table 5 Calculation with type B algorithms with prescription on GTV D50% (scenario C)

GTV D50% = 54 Gy									
	GTV D98% type B (Gy)			PTV D95% type B			Normal lung receiving less than 10 Gy (cm^3)		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Novalis	51.0	47.8	50.9	47.4	40.6	44.7	3829	3455	2284
Cyberknife	50.0	47.0	49.1	46.1	38.2	43.5	3794	3563	2309
Clinac	51.2	48.6	50.5	46.9	47.7	47.3	3735	3299	1945

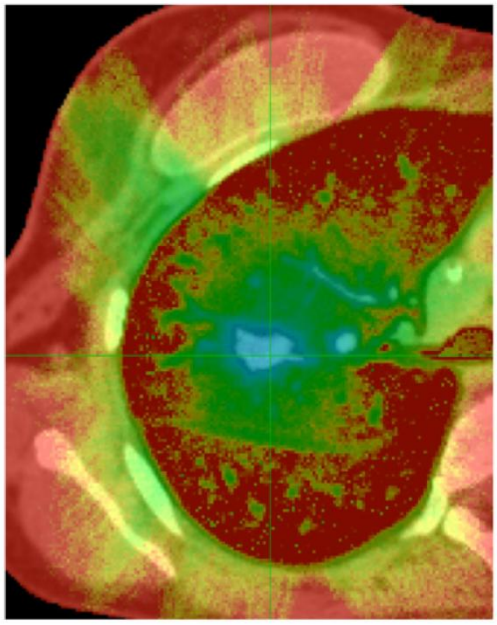


Figure 6 Map of % of statistical uncertainties for Monte-Carlo-Calculation for case 2 and Cyberknife.

the median dose even in the example with the highest degree of heterogeneity: case 2 treated with the CyberKnife. We believe that the GTV D50% is more relevant than the GTV near-minimum (D98%) because the GTV encompasses sometimes small low density regions, lacking electronic equilibrium too, so the uncertainty is larger than for the GTV D50%. However we think we should report this value.

During plan optimization the PTV needs to be targeted entirely though, because the GTV can be anywhere inside the PTV. This task is best achieved using a type A algorithm, available in all TPSs. In clinical practice, like in the exercise described in current article, a plan is created using a type A algorithm, covering 95% of the PTV by the prescription isodose. A steep gradient around the PTV is not really needed, because the effect on the DVH of the lung is very small and because the CTV extensions are not known, so a relatively homogenous fluence in the target is maintained. Then a type B algorithm is used to recalculate and rescale the result and to prescribe the dose to the GTV median dose (GTV D50%). As the median does not really depend on the minimum, the GTV median dose is almost constant no matter where the GTV is within the PTV envelop and, consequently, is very representative of the dose delivered and thus the biological impact to the tumor. For example if for case 2 treated with the Cyberknife, where the dose heterogeneity is largest (difference between PTV D95% and GTV50% and GTV98%), the GTV is shifted 5 mm towards the dose minimum in the PTV (overriding density of PTV-GTV and GTV shifted), (see Figure 7), the median dose to the GTV is 53.2 Gy instead of 54 Gy (a 1.5% variation): because of the rebuild-up effect the isodose moves with the GTV. For all other studied situations the variation would be lower because the dose heterogeneity is lower within the PTV. The combination of optimizing to the PTV using a type A algorithm, and prescribing to the median GTV dose using a type B algorithm is thus a very robust method.

The reoptimization with type B algorithms, studied in this article, is not described because it increases uncertainties of delivered dose. To compensate the lack of electronic equilibrium the system will increase the fluence in

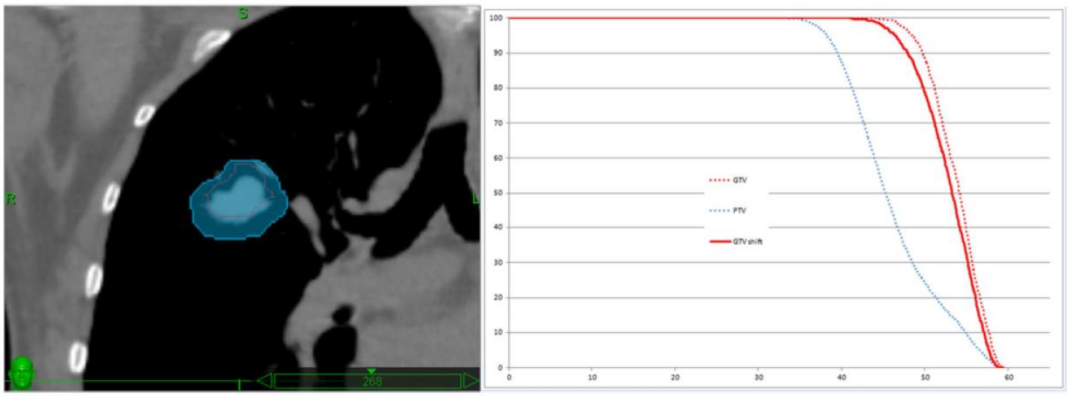


Figure 7 Effect of a shift of the GTV in the envelop of the PTV for case 2 and Cyberknife: The red contour indicates the position of the GTV shifted within the PTV (blue).

this region, this results in a non robust plan, as the GTV will be overexposed when it moves into the regions with increased photon fluence. As in all situations of lack of electronic equilibrium, like superficial neck nodes in build-up region, optimization increases the uncertainty on the delivered dose. New algorithms taking into account 4DCT and uncertainty of patient position should be used to have a robust optimization. As this robust optimization algorithms are not yet available in our TPSs, the method proposed in current work (optimizing on the PTV using a type A algorithm followed by recalculation using a type B algorithm) provides the best available alternative for the moment.

In this experiment, the breathing motion was considered controlled. If the PTV is defined from an ITV which includes all the positions of the GTV during the breathing cycle, the beam sizes used were larger so the lack of lateral electronic equilibrium was smaller but, still, the PTV D95% did not predict the minimum dose to the GTV, because the PTV margin consisted mostly of a low-density region. Therefore, also for the ITV method, we think it is better to prescribe to the GTV median. Without any gating or real-time tracking, the dose calculation is only a snapshot, but even then prescribing to the GTV D50% is more precise because it takes into account the rebuild-up. Actually the larger the margin the larger the difference between PTV minimum and GTV dose. That means that a large uncertainty in GTV position leads to a higher delivered dose. A GTV based prescription avoids this drawback.

Conclusion

The right dose needed for lung lesions treated by SBRT is not yet known. Advanced algorithms are needed and the GTV dose should be reported because the PTV prescription does not predict the dose to the GTV. Every team should report GTV D50% and D98% to enable a comparison of all results.

We suggest using type A algorithms to target the PTV during optimization and to recalculate dose with type B algorithms rescaling the prescription to the GTV median dose. Until robust optimization algorithms will be introduced in all TPSs, this appears to be the most robust way to avoid discrepancies from device to device and case to case and this with or without motion management (tracking, gating).

Additional file

Additional file 1: Slideshow.

Competing interests

The Centre Oscar Lambret has a collaboration agreement with Accuray Incorporated.

Authors' contributions

TL and AL designed the study and did the calculations of cases for different platforms. TL interpreted and analyzed the data, and drafted the manuscript. XM and EL did the contouring, initial prescriptions and participated to the discussion. NR analyzed data and helped a lot for the manuscript and the discussion. All authors read and approved the final manuscript.

Author details

[1]Service de Physique Médicale, Centre Oscar Lambret, Lille, France. [2]Service de Physique Médicale, Institut de Cancérologie de l'Ouest, Nantes, France. [3]Département Universitaire de Radiothérapie, Centre Oscar Lambret, Lille, France.

Received: 8 January 2014 Accepted: 29 September 2014
Published online: 16 October 2014

References

1. Papanikolaou N, Battista JJ, Boyer AL, Kappas C, Klein E, Mackie TR, Van Dy J: *Report of Task Group No. 65 of the Radiation Therapy Committee of the American Association of Physicists in Medicine: Tissue Inhomogeneity Corrections for Megavoltage Photon Beams.* Madison, WI: Medical Physics Publishing; 2004:1–130.
2. Reynaert N, van der Marck SC, Schaart DR, Van der Zee W, Van Vliet-Vroegindeweij C, Tomsej M, Jansen J, Heijmen B, Coghe M, De Wagter C: **Monte Carlo treatment planning for photon and electron beams.** *Radiat Phys Chem Oxf Engl* 2007, **76**(4):643–686.
3. Chetty IJ, Curran B, Cygler JE, DeMarco JJ, Ezzell G, Faddegon BA, Kawrakow I, Keall PJ, Liu H, Ma C-MC, Rogers DWO, Seuntjens J, Sheikh-Bagheri D, Siebers JV: **"Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning".** *Med Phys* 2007, **34**:4818. no. 12.
4. Xiao Y, Papiez L, Paulus R, Timmerman R, Straube WL, Bosch WR, Michalski J, Galvin JM: **Dosimetric evaluation of heterogeneity corrections for RTOG 0236: stereotactic body radiotherapy of inoperable stage I-II non-small-cell lung cancer.** *Int J Radiat Oncol Biol Phys* 2009, **73**(4):1235–1242.
5. Li J, Galvin J, Harrison A, Timmerman R, Yu Y, Xiao Y: **Dosimetric verification using Monte Carlo calculations for tissue heterogeneity-corrected conformal treatment plans following RTOG 0813 dosimetric criteria for lung cancer stereotactic body radiotherapy.** *Int J Radiat Oncol Biol Phys* 2012, **84**(2):508–513.
6. Takahashi W, Yamashita H, Saotome N, Iwai Y, Sakumi A, Haga A, Nakagawa K: **"Evaluation of heterogeneity dose distributions for Stereotactic Radiotherapy (SRT): comparison of commercially available Monte Carlo dose calculation with other algorithms".** *Radiat Oncol* 2012, **7**:20. no. 1.
7. Sikora M, Muzik J, Söhn M, Weinmann M, Alber M: **Monte Carlo vs. pencil beam based optimization of stereotactic lung IMRT.** *Radiat Oncol* 2009, **4**:64.
8. Schuring D, Hurkmans CW: **Developing and evaluating stereotactic lung RT trials: what we should know about the influence of inhomogeneity corrections on dose.** *Radiat Oncol* 2008, **3**:21.
9. Zhuang T, Djemil T, Qi P, Magnelli A, Stephans K, Videtic G, Xia P: **"Dose calculation differences between Monte Carlo and pencil beam depend on the tumor locations and volumes for lung stereotactic body radiation therapy".** *J Appl Clin Med Phys* 2013, **14**(no. 2):4011.
10. Nuyttens JJ, Prévost J-B, Praag J, Hoogeman M, Van Klaveren RJ, Levendag PC, Pattynama PMT: **Lung tumor tracking during stereotactic radiotherapy treatment with the CyberKnife: Marker placement and early results.** *Acta Oncol* 2006, **45**(7):961–965.
11. Collins BT, Erickson K, Reichner CA, Collins SP, Gagnon GJ, Dieterich S, McRae DA, Zhang Y, Yousefi S, Levy E, Chang T, Jamis-Dow C, Banovac F, Anderson ED: **"Radical stereotactic radiosurgery with real-time tumor motion tracking in the treatment of small peripheral lung tumors".** *Radiat Oncol* 2007, **2**:39.
12. Das IJ, Ding GX, Ahnesjö A: **"Small fields: Nonequilibrium radiation dosimetry".** *Med Phys* 2008, **35**:206. no. 1.
13. Brown WT, Wu X, Fayad F, Fowler JF, García S, Monterroso MI, de la Zerda A, Schwade JG: **Application of Robotic Stereotactic Radiotherapy to Peripheral Stage I Non-small Cell Lung Cancer with Curative Intent.** *Clin Oncol* 2009, **21**(8):623–631.
14. Soldà F, Lodge M, Ashley S, Whitington A, Goldstraw P, Brada M: **Stereotactic radiotherapy (SABR) for the treatment of primary non-small**

cell lung cancer; Systematic review and comparison with a surgical cohort. *Radiother Oncol* 2013, **109**(1):1–7.
15. van Baardwijk A, Tomé WA, van Elmpt W, Bentzen SM, Reymen B, Wanders R, Houben R, Ollers M, Lambin P, De Ruysscher D: **Is high-dose stereotactic body radiotherapy (SBRT) for stage I non-small cell lung cancer (NSCLC) overkill? A systematic review.** *Radiother Oncol* 2012, **105**(2):145–149.
16. Grimm J, LaCouture T, Croce R, Yeo I, Zhu Y, Xue J: **"Dose tolerance limits and dose volume histogram evaluation for stereotactic body radiotherapy".** *J Appl Clin Med Phys* 2011, **12:**3368. no. 2.
17. Benedict SH, Yenice KM, Followill D, Galvin JM, Hinson W, Kavanagh B, Keall P, Lovelock M, Meeks S, Papiez L, Purdie T, Sadagopan R, Schell MC, Salter B, Schlesinger DJ, Shiu AS, Solberg T, Song DY, Stieber V, Timmerman R, Tomé WA, Verellen D, Wang L, Yin F-F: **"Stereotactic body radiation therapy: The report of AAPM Task Group 101".** *Med Phys* 2010, **37:**4078. no. 8.
18. Petoukhova AL, van Wingerden K, Wiggenraad RGJ, van de Vaart PJM, van Egmond J, Franken EM, van Santvoort JPC: **"Verification measurements and clinical evaluation of the iPlan RT Monte Carlo dose algorithm for 6 MV photon energy".** *Phys Med Biol* 2010, **55:**4601. no. 16.
19. Ma C-M, Li JS, Deng J, Fan J: **Implementation of Monte Carlo Dose calculation for CyberKnife treatment planning.** *J Phys Conf Ser* 2008, **102:**012016.
20. Kry SF, Alvarez P, Molineu A, Amador C, Galvin J, Followill DS: **Algorithms used in heterogeneous dose calculations show systematic differences as measured with the Radiological Physics Center's anthropomorphic thorax phantom used for RTOG credentialing.** *Int J Radiat Oncol Biol Phys* 2013, **85**(1):e95–100.
21. van der Voort NC, van Zyp J-B, Prévost MS, Hoogeman J, van der Praag B, Holt PC, van Levendag RJ, Klaveren PP, Nuyttens JJ: **Stereotactic radiotherapy with real-time tumor tracking for non-small cell lung cancer: clinical outcome.** *Radiother Oncol* 2009, **91**(3):296–300.
22. Bral S, Gevaert T, Linthout N, Versmessen H, Collen C, Engels B, Verdries D, Everaert H, Christian N, De Ridder M, Storme G: **Prospective, risk-adapted strategy of stereotactic body radiotherapy for early-stage non-small-cell lung cancer: results of a Phase II trial.** *Int J Radiat Oncol Biol Phys* 2011, **80**(5):1343–1349.
23. Traberg Hansen A, Petersen JB, Høyer M, Christensen JJ: **"Comparison of two dose calculation methods applied to extracranial stereotactic radiotherapy treatment planning".** *Radiother Oncol* 2005, **77:**96–98. no. 1.
24. Sharma SC, Ott JT, Williams JB, Dickow D: **"Clinical implications of adopting Monte Carlo treatment planning for CyberKnife".** *J Appl Clin Med Phys* 2010, **11:**3142. no. 1.
25. Hurkmans CW, Cuijpers JP, Lagerwaard FJ, Widder J, van der Heide UA, Schuring D, Senan S: **"Recommendations for implementing stereotactic radiotherapy in peripheral stage IA non-small cell lung cancer: report from the Quality Assurance Working Party of the randomised phase III ROSEL study".** *Radiat Oncol* 2009, **4:**1.
26. van der Voort NC, van Zyp MS, Hoogeman S, van de Water S, van der Levendag PC, Holt B, Heijmen BJM, Nuyttens JJ: **Clinical introduction of Monte Carlo treatment planning: a different prescription dose for non-small cell lung cancer according to tumor location and size.** *Radiother Oncol* 2010, **96**(1):55–60.

doi:10.1186/s13014-014-0223-5
Cite this article as: Lacornerie *et al.*: GTV-based prescription in SBRT for lung lesions using advanced dose calculation algorithms. *Radiation Oncology* 2014 **9**:223.

2.3 L'optimisation du plan dans les poumons

Maintenant qu'on sait comment calculer la dose, quels sont les niveaux de dose à prescrire et sur quels volumes, il reste encore la question de l'optimisation du plan de traitement. Depuis de nombreuses années, on cherche des méthodes pour introduire les algorithmes de calcul de dose Monte Carlo dans la boucle d'optimisation. Malheureusement, un calcul de dose Monte Carlo prend beaucoup de temps, et pour l'optimisation, on doit faire beaucoup d'itérations. L'utilisation d'un algorithme approximatif va introduire deux erreurs : des erreurs de calcul de dose, mais aussi des erreurs d'optimisation [81]. La solution optimale obtenue avec un algorithme de type A peut devenir sub-optimale une fois recalculée en Monte Carlo. La solution logique est alors d'optimiser directement en Monte Carlo. On peut, par exemple, calculer la distribution de dose obtenue pour un certain nombre de « beamlets » et ensuite optimiser leur poids avant de finalement les segmenter (optimisation inverse) [81]. Une autre solution consiste à générer des segments concrets qui sont optimisés en poids et géométrie et on essaie d'introduire le Monte Carlo dans la boucle d'optimisation (« segment-based optimisation ») [82]. Plusieurs solutions ont été proposées dans le passé.

A l'inverse, on peut imaginer plusieurs exemples, où l'utilisation d'un algorithme de type B pendant l'optimisation est contre-productif. Les traitements du poumon sont peut-être le meilleur exemple. Le principe est expliqué en figure 29. La différence de densité entre le GTV même et la marge autour, va introduire des différences en fluence de photons si on utilise un algorithme de type B. Le système d'optimisation va essayer de compenser la faible densité dans la marge autour du GTV en augmentant la fluence. Si on optimise avec un algorithme du type A, la fluence des photons est homogène. Si on fait ensuite le calcul de dose pour une géométrie statique où le GTV est parfaitement centré dans le PTV, la dose sera correcte dans les deux cas. Par contre, dès le moment où l'on prend en compte les mouvements respiratoires du GTV dans le PTV, on va obtenir un sur-dosage si on a utilisé un algorithme de type B pour l'optimisation. Ceci est dû au mouvement de la tumeur dans les zones où la fluence est plus élevée. Si on a utilisé un algorithme de type A, la dose dans le GTV n'est pas modifiée, peu importe sa position dans le PTV (le but de l'utilisation d'un PTV est de garantir une dose constante au GTV quels que soient les incertitudes de positionnement). La solution robuste consiste à optimiser la dose au PTV avec un algorithme de type A, puis de recalculer la distribution de dose en utilisant un algorithme de type B pour la prescription faite sur le GTV (et non sur le PTV). En effet, dans beaucoup de cas, le GTV n'est pas homogène en densité (même le GTV contient souvent du tissu poumon de faible densité) et il vaut mieux prescrire sur 50 % du GTV (et non sur 95 %).

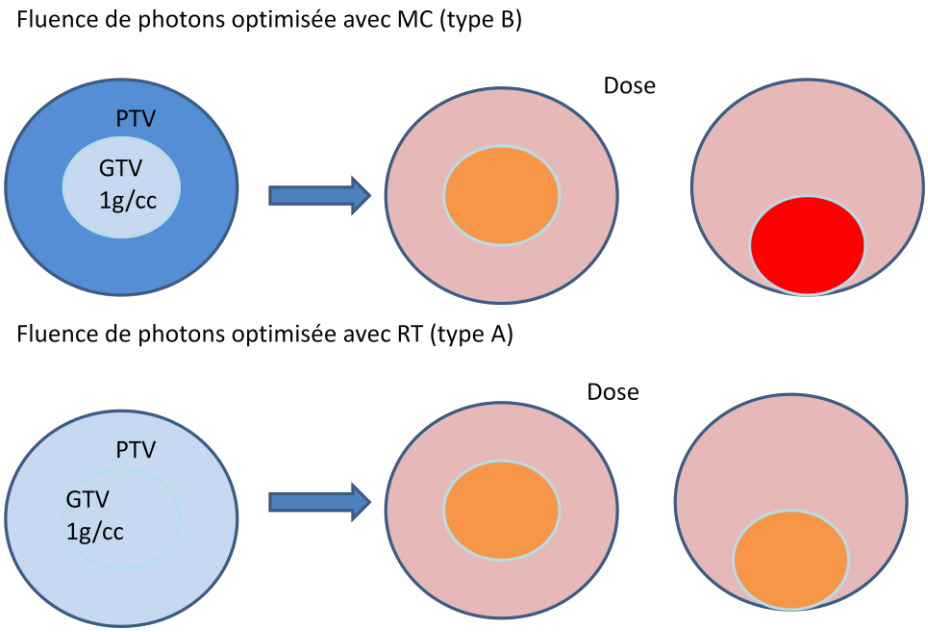


Figure 29 : L'impact du type de l'algorithme de calcul de dose pendant l'optimisation du plan de traitement pour les tumeurs poumons. A cause de la densité hétérogène il ne faut pas utiliser un algorithme du type B, sinon, la fluence des photons n'est pas homogène dans le PTV, et le GTV sera surdosé dans ces zones.

Chapitre 5 : Discussion générale + Futures Perspectives

1. Rapid learning

La radiothérapie devient plus individualisée car de plus en plus d'informations sont disponibles lors de la préparation du traitement (imagerie multi-modalités) et de sa réalisation [83]. Il est de plus en plus difficile de définir un groupe de patients homogène pour les études cliniques. Seuls 3 % des adultes traités en radiothérapie sont inclus dans des études cliniques [84], il est donc nécessaire d'utiliser les résultats des traitements qui ne sont pas inclus dans des études formelles (« rapid learning »). Il est même indispensable de combiner les données de plusieurs centres de radiothérapie [85] (« distributed learning »). Le consortium coordonné par Maastro est un exemple d'échange de données à l'échelle internationale [86].

A cause du surplus d'information, les systèmes informatisés deviennent indispensables à la prise de décision concernant le traitement du patient (CDSS, Clinical decision-support system). Pour chaque patient, plusieurs modalités d'imagerie sont à notre disposition, des tests sanguins, de l'information génomique, plusieurs modalités de traitement… La décision thérapeutique est de plus en plus prise en accord avec le patient en tenant compte de l'état général du patient, de son âge, de sa qualité de vie et des coûts du traitement. La figure 30 tirée de la référence [84] de Lambin et al représente cette philosophie.

Les premiers axes de recherche se situent au niveau de la table à droite de ce schéma (« Learn »). Il faut plus spécifiquement focaliser sur la qualité des données. A long terme, le Centre Oscar Lambret a l'ambition d'appliquer cette table en entier et d'introduire un système CDSS, mais il est avant tout nécessaire de développer les outils afin de standardiser le recueil des données médicales des patients.

L'erreur, qui est souvent faite pendant la génération d'un modèle CDSS pour la radiothérapie, est l'utilisation systématique des données de faible qualité [84]. Par exemple, si on veut comparer l'efficacité et/ou toxicité des traitements protons et photons, la conclusion dépend énormément de la qualité du traitement : si la tumeur n'est pas correctement délinéée ou si le patient n'est pas correctement positionné, les deux modalités de traitement ne seront pas efficaces et la conclusion de l'étude sera faussée.

Il faut alors augmenter la qualité de chaque étape de la chaîne de traitement en utilisant les images IRM, et déterminer précisément la dose effectivement délivrée pour les traitements avancés comme la RCMI et la Stéréotaxie.

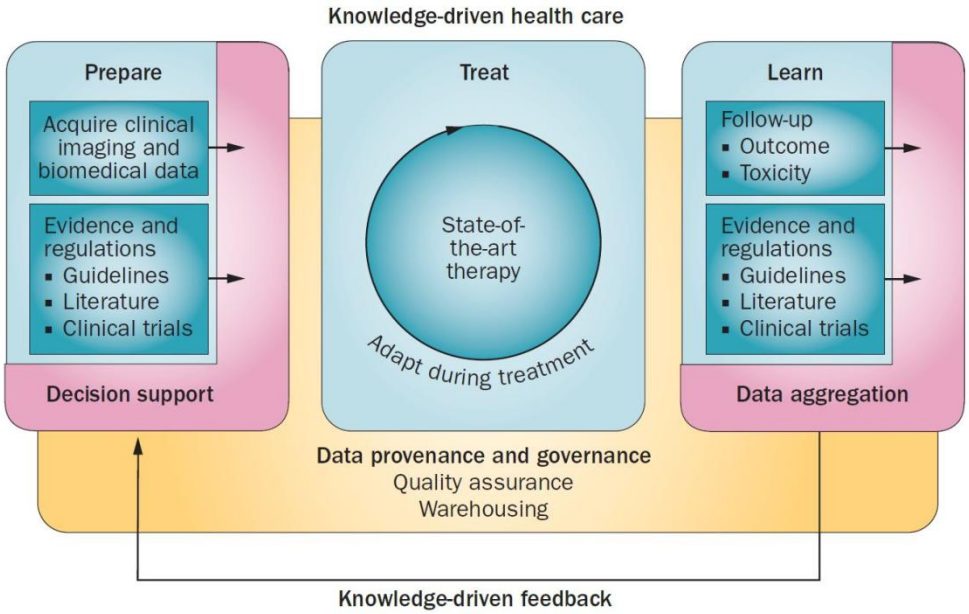


Figure 30 : Représentation d'un système CDSS (Lambin *et al*. [84]).

2. La dose réellement délivrée

Nous l'avons vu au chapitre 2, la distribution de la dose délivrée peut différer d'une façon significative de la dose planifiée. L'utilisation de la dose planifiée peut introduire un biais dans les bases de données et fausser les conclusions médicales. Même si une comparaison systématique entre dose administrée et planifiée offre déjà beaucoup d'informations utiles pour améliorer les traitements futurs, le but est de lier ces différences aux résultats cliniques en termes de toxicité, d'effets secondaires et de récidives locales. Pour le recueil de ces données, il faut utiliser des techniques informatisées (« data mining », Roelofs et al [86]).

3. La radiothérapie en IRM seule

Au chapitre 3, nous avons vu que l'introduction de l'IRM à chaque étape de la chaîne de traitement pourrait largement augmenter la qualité des traitements et en même temps améliorer la qualité des données utilisées pour le Rapid Learning:

- la délinéation des organes (meilleur contraste tissus mous en IRM)
- calcul de dose directement sur les images IRM
- positionnement pendant chaque séance de traitement
- IRM 4D pendant chaque séance : reconstruction de la dose administrée
- suivi de l'évolution de la tumeur entre les séances de traitement
- suivi (efficacité/toxicité) après le traitement

- IRM-Omics

Délinéation des organes et calcul de dose

Il est déjà reconnu que la délinéation des organes est plus facile sur l'IRM que sur le CT, grâce à un contraste dans les tissus mous plus élevé. La suite du projet PhysiCancer (INCa) permettra d'illustrer, en utilisant une plate-forme Monte-Carlo, que les calculs de dose sur l'IRM peuvent encore améliorer la précision, par rapport aux systèmes de planification de traitement actuellement disponibles, et ce, pour les modalités de radiothérapie avancées. L'impact des inexactitudes et des distorsions géométriques doit être éliminé en combinant et optimisant des séquences IRM en collaboration avec le LTSI de Rennes (par exemple DUTE plus spin écho, ou une Double-ZTE, …). Un simulateur IRM sera installé au COL. L'information sur le contenu d'hydrogène des tissus sera utilisée pour améliorer la précision du calcul de dose dans l'os et les poumons. Une machine de radiothérapie avec IRM intégrée existe déjà (ViewRay) et le consortium Elekta/Philips prévoit l'installation d'accélérateurs combinés avec une IRM en 2019. Le Centre Oscar Lambret envisage l'acquisition d'une de ces machines, leur utilisation nécessitant des algorithmes capables de calculer la dose directement sur des images IRM.

Comme nous l'avons montré au chapitre 3, le logiciel Moderato est déjà utilisé pour le calcul de dose sur les images IRM grâce à l'utilisation d'un pseudo-CT. Du fait de l'augmentation prévisible des traitements en IRM-seule, le système Monte Carlo sera adapté pour être en mesure d'importer et de convertir des données d'IRM (calculs classiques basés sur la densité de masse et la composition des tissus). Une conversion théorique entre la densité de protons et la densité de masse pour les régions osseuses devrait permettre une meilleure description de l'os. De plus, la teneur en hydrogène du tissu pulmonaire doit permettre d'effectuer des calculs plus précis du transport des particules et du dépôt d'énergie pour les traitements dans le thorax. Les séquences IRM doivent être optimisées, pour l'information tissulaire, la reproductibilité, le bruit et le positionnement du patient en utilisant des bobines dédiées (LTSI).

Un problème plus pratique à résoudre est la détection des fiduciels implantés autour/dans la tumeur pour le positionnement du patient ou le suivi des mouvements en temps réel pendant le traitement. Dans le projet « Gentle Radiotherapy » en Suède, un fournisseur de fiduciels essaye de générer des fiduciels compatibles avec l'IRM. Néanmoins, d'après le responsable du projet (Tufve Nyholm, communication personnelle), il faut en parallèle optimiser ou générer des séquences IRM plus spécifiques pour la détection des fiduciels.

Un autre détail pratique consiste à la génération des images DRRs (« digitally reconstructed radiographs ») qui sont indispensables pour le positionnement du patient. Au CyberKnife par exemple, le positionnement est réalisé avec un système d'imagerie radiographique stéréoscopique. La génération de ces images est complètement basée sur les densités dans les tissus. On peut imaginer que de petites erreurs de densité dans le pseudo-CT vont avoir un impact relativement important. Là également, la conversion entre IRM et pseudo-CT doit être optimisée par rapport aux méthodes actuellement disponibles.

L'IRM pour positionner le patient et déterminer la dose administrée

Avant chaque séance de traitement, le patient doit être positionné sur la machine de traitement comme il l'était pendant la simulation afin de délivrer le traitement. L'utilisation d'un système IRM 3D intégré à la machine de traitement serait parfaitement compatible avec la simulation IRM et un recalage entre ces deux images augmenterait la précision de positionnement. Ce système IRM embarqué peut aussi servir à déterminer des images 4D pendant le traitement. Ces images, comme cela a été décrit plus haut, peuvent être utilisées en combinaison avec les fichiers logs des paramètres machine, pour reconstruire la dose effectivement administrée. Cette méthode sera plus précise que les méthodes décrites au chapitre 2 où il faut reconstruire la géométrie du patient en utilisant les positions des fiduciels et le scan 4D de simulation.

L'IRM pour le suivi pendant et après le traitement

L'IRM fonctionnelle peut être utilisée entre les séances de traitement pour évaluer l'évolution de la tumeur. La combinaison de l'information anatomique et fonctionnelle (sans avoir besoin d'une imagerie irradiante comme le TDM et/ou le TEP), peut guider l'adaptation du traitement. La tumeur évolue pendant le traitement. Si elle diminue, les organes à risque vont entrer dans le champ d'irradiation (par exemple les parotides pour les traitements ORL, ou le rectum/la vessie pour les traitements gynécologiques). Adapter le traitement est un autre moyen de limiter les différences entre la dose planifiée et celle délivrée, et d'augmenter la qualité du traitement et des données sauvegardées dans une base de données qui sera utilisé comme système prédictif (CDSS).

Après le traitement, l'IRM fonctionnelle sera utilisée pour déterminer les effets cliniques comme l'efficacité du traitement et les toxicités.

Les Radiomics IRM (IRM-Omics)

L'évolution vers la médecine personnalisée est largement guidée par les technologies génomiques (GenOmics) et/ou protéomiques (proteomics) [85] qui sont utilisées pour déterminer une signature spécifique de la tumeur. Ces techniques demandent des biopsies (invasives) et une critique pourrait être qu'on ne teste que quelques cellules d'une partie spécifique de la tumeur.

Radiomics [87][88][89] est une évolution récente des techniques dans laquelle on va chercher de l'information quantitative dans le scanner (TDM) du patient, comme la géométrie du contour de la tumeur et l'homogénéité des valeurs de gris du scanner TDM dans ce contour. Ces paramètres sont apparemment liés à l'effet du traitement et sont utilisables dans un système prédictif (élément clef d'un CDSS). Un nombre de paramètres (« features ») sont d'abord testés en stabilité (par rapport à la reproductibilité d'images et de délinéation de la tumeur) et en pouvoir prédictif, en utilisant un premier groupe de patients pour lesquels les effets cliniques post-traitement sont déjà connus (survie, efficacité). Un système informatique détermine la signature optimale de la tumeur (combinaison de plusieurs « features »). Après, le pouvoir prédictif de cette signature est testé sur un deuxième groupe de patients. Il est démontré que la signature Radiomics est corrélée avec les signatures GenOmics. De cette façon, les Radiomics peuvent devenir un point majeur dans un

système prédictif comme un CDSS. A nouveau, ce modèle demande des données cliniques d'une qualité assez élevée, sinon le système sera entraîné fautivement.

L'inconvénient du Radiomics est la dose associée au scanner et le manque d'information fonctionnelle. Il y a effectivement la possibilité d'inclure des TDMs de perfusion, mais cela demande encore plus de dose au patient. L'IRM pourrait offrir une alternative. Le problème pour le moment est l'instabilité ou le manque de reproductibilité de l'IRM. Deux séries d'images, utilisant la même séquence sur le même patient ne donnent pas toujours des images identiques. Une sélection et une optimisation des séquences IRM utilisées est nécessaire. La collaboration avec le LTSI de Rennes, le NPL d'Angleterre et le HUS de Finlande sera très importante pour résoudre ce problème.

D'autre part, dans cette utilisation des radiomics IRM, il est également important de s'assurer de la reproductibilité des images IRM d'un centre à l'autre et d'un imageur à l'autre. Le contrôle qualité des images IRM est une nouvelle fois primordial pour cette utilisation. C'est une étape nécessaire pour trouver une signature numérique en utilisant les images IRM.

4. Les effets cliniques

Les différences entre la dose réellement administrée et la dose planifiée doivent être liées aux effets cliniques, comme la toxicité, les effets secondaires et les récidives locales. Cela demande une harmonisation précise du recueil des données cliniques. Initialement, le recueil des données sera manuel et ne sera pas possible pour tous les traitements. Une première option sera juste de noter s'il y a des toxicités ou des récidives locales pendant la consultation en utilisant le Système d'Information Hospitalier (SIH, DXCare au Centre Oscar Lambret) puis de détailler ces toxicités dans un formulaire avec la base de données cliniques (ClinSight) offline. Pour ces patients, il faudra saisir les données de l'imagerie IRM de diffusion, les niveaux de dose pour des volumes spécifiques dans les organes pour lesquels on a observé une toxicité… Le formulaire à remplir doit être bien standardisé, facile à utiliser (pour ne pas perdre trop de temps), compact mais complet. Après nous étudierons son automatisation en utilisant une base de données centralisée (data warehouse, data mining, BigData [85]). Ce logiciel doit être combiné avec des outils qui peuvent chercher automatiquement les données des patients (dans le PACS, dans DXCare, …), comme Roelofs et al [85] l'ont décrit par exemple , et remplir le document dans ClinSight. Les données doivent encore être validées systématiquement, mais on peut facilement gagner 50 % du temps par rapport à un recueil manuel et les résultats sont plus robustes. Cette base de données permettra de faire les analyses statistiques de manière aisée et de faire les liens entre dose administrée et effets cliniques.

Cette information pourra être utilisée ensuite pour mieux optimiser les traitements, pour adapter les contraintes utilisées pendant l'optimisation du traitement et faire évoluer les prescriptions.

6. Références

[1] E. E. Klein, J. Hanley, J. Bayouth, F. F. Yin, W. Simon, S. Dresser, C. Serago, F. Aguirre, L. Ma, B. Arjomandy, C. Liu, C. Sandin, T. Holmes, and A. A. of P. in M. Task Group, "Task Group 142 report: quality assurance of medical accelerators," *Med Phys*, vol. 36, no. 9, pp. 4197–4212, 2009.

[2] P. Andreo, T. Kanai, F. Laitano, V. Smyth, N. Zealand, and S. Vynckier, "Ansorbed Dose Determination in External Beam Radiotherapy," *Iaea Trs-398*, vol. 2004, no. April, 2004.

[3] J. B. Smilowitz, I. J. Das, V. Feygelman, B. A. Fraass, S. F. Kry, I. R. Marshall, D. N. Mihailidis, Z. Ouhib, T. Ritter, M. G. Snyder, and L. Fairobent, "AAPM Medical Physics Practice Guideline 5.a.: Commissioning and QA of Treatment Planning Dose Calculations - Megavoltage Photon and Electron Beams," *J. Appl. Clin. Med. Phys.*, vol. 16, no. 5, pp. 14–34, 2015.

[4] N. Papanikolaou, J. J. Battista, A. L. Boyer, C. Kappas, E. Klein, T. R. Mackie, M. Sharpe, and J. Van Dyk, "Report of Task Group No. 65 of the Radiation Therapy Committee of the American Association of Physicists in Medicine: Tissue inhomogeneity corrections for megavoltage photon beams," *Med. Phys*, pp. 1–130, 2004.

[5] V. Grégoire and T. R. Mackie, "State of the art on dose prescription, reporting and recording in Intensity-Modulated Radiation Therapy (ICRU report No. 83)," *Cancer/Radiotherapie*, vol. 15, no. 6–7. pp. 555–559, 2011.

[6] R. L. Stern, R. Heaton, M. W. Fraser, S. M. Goddu, T. H. Kirby, K. L. Lam, A. Molineu, and T. C. Zhu, "Verification of monitor unit calculations for non-IMRT clinical radiotherapy: report of AAPM Task Group 114.," *Med. Phys.*, vol. 38, no. 1, pp. 504–530, 2011.

[7] N. Reynaert, B. De Smedt, M. Coghe, L. Paelinck, B. Van Duyse, W. De Gersem, C. De Wagter, W. De Neve, and H. Thierens, "MCDE: a new Monte Carlo dose engine for IMRT.," *Phys. Med. Biol.*, vol. 49, no. 14, pp. N235–N241, 2004.

[8] I. J. Chetty, B. Curran, J. E. Cygler, J. J. DeMarco, G. Ezzell, B. a Faddegon, I. Kawrakow, P. J. Keall, H. Liu, C.-M. C. Ma, D. W. O. Rogers, J. Seuntjens, D. Sheikh-Bagheri, and J. V Siebers, "Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning," *Med. Phys.*, vol. 34, no. 12, pp. 4818–4853, 2007.

[9] N. Reynaert, S. Vandermarck, D. Schaart, W. Vanderzee, C. Vanvlietvroegindeweij, M. Tomsej, J. Jansen, B. Heijmen, M. Coghe, and C. Dewagter, "Monte Carlo treatment planning for photon and electron beams," *Radiat. Phys. Chem.*, vol. 76, no. 4, pp. 643–686, 2007.

[10] N. Reynaert, M. Coghe, B. De Smedt, L. Paelinck, B. Vanderstraeten, W. De Gersem, B.

Van Duyse, C. De Wagter, W. De Neve, and H. Thierens, “The importance of accurate linear accelerator head modelling for IMRT Monte Carlo calculations.,” *Phys. Med. Biol.*, vol. 50, no. 5, pp. 831–846, 2005.

[11] a Wagner, F. Crop, T. Lacornerie, F. Vandevelde, and N. Reynaert, “Use of a liquid ionization chamber for stereotactic radiotherapy dosimetry.,” *Phys. Med. Biol.*, vol. 58, no. 8, pp. 2445–59, 2013.

[12] D. W. O. Rogers, B. Walters, and I. Kawrakow, “BEAMnrc Users Manual,” *Source*, vol. 509, pp. 1–260, 2011.

[13] B. Walters, I. Kawrakow, and D. W. O. Rogers, “DOSXYZnrc Users Manual,” *NRCC Rep. PIRS-0794*, no. April, pp. 1–125, 2013.

[14] J. Lippuner, I. a Elbakri, C. Cui, and H. R. Ingleby, “Epp: A C++ EGSnrc user code for x-ray imaging and scattering simulations.,” *Med. Phys.*, vol. 38, no. 3, pp. 1705–1708, 2011.

[15] S. Jan, D. Benoit, E. Becheva, T. Carlier, F. Cassol, P. Descourt, T. Frisson, L. Grevillot, L. Guigues, L. Maigne, C. Morel, Y. Perrot, N. Rehfeld, D. Sarrut, D. R. Schaart, S. Stute, U. Pietrzyk, D. Visvikis, N. Zahra, and I. Buvat, “GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy.,” *Phys. Med. Biol.*, vol. 56, no. 4, pp. 881–901, 2011.

[16] Q. Chen, Y. Chen, M. Chen, E. Chao, E. Sterpin, and W. Lu, “A slit method to determine the focal spot size and shape of TomoTherapy system.,” *Med. Phys.*, vol. 38, no. 6, pp. 2841–2849, 2011.

[17] B. Vanderstraeten, P. W. Chin, M. Fix, A. Leal, G. Mora, N. Reynaert, J. Seco, M. Soukup, E. Spezi, W. De Neve, and H. Thierens, “Conversion of CT numbers into tissue parameters for Monte Carlo dose calculations: a multi-centre study.,” *Phys. Med. Biol.*, vol. 52, pp. 539–562, 2007.

[18] B. Demol, R. Viard, and N. Reynaert, “Monte carlo calculation based on hydrogen composition of the tissue for MV photon radiotherapy,” *J. Appl. Clin. Med. Phys.*, vol. 16, no. 5, pp. 117–130, 2015.

[19] B. De Smedt, B. Vanderstraeten, N. Reynaert, W. De Neve, and H. Thierens, “Investigation of geometrical and scoring grid resolution for Monte Carlo dose calculations for IMRT.,” *Phys. Med. Biol.*, vol. 50, no. 17, pp. 4005–4019, 2005.

[20] B. De Smedt, M. Fippel, N. Reynaert, and H. Thierens, “Denoising of Monte Carlo dose calculations: smoothing capabilities versus introduction of systematic bias.,” *Med. Phys.*, vol. 33, no. 6, pp. 1678–1687, 2006.

[21] P. Mansfield and A. A. Maudsley, “Medical imaging by NMR,” *Br. J. Radiol.*, vol. 50, no. 591, pp. 188–194, 1977.

[22] S. J. Gwyther, “Modern techniques in radiological imaging related to oncology,” *Ann. Oncol.*, vol. 5, no. suppl 4, pp. S3–S7, 1994.

[23] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance*

Imaging: Physical Principles and Sequence Design, vol. 1st. 1999.

[24] B. A. Jung and M. Weigel, "Spin echo magnetic resonance imaging.," *J. Magn. Reson. Imaging*, vol. 37, no. 4, pp. 805–17, 2013.

[25] M. Debois, R. Oyen, F. Maes, G. Verswijvel, G. Gatti, H. Bosmans, M. Feron, E. Bellon, G. Kutcher, H. Van Poppel, and L. Vanuytsel, "The contribution of magnetic resonance imaging to the three-dimensional treatment planning of localized prostate cancer," *Int. J. Radiat. Oncol.*, vol. 45, no. 4, pp. 857–865, 1999.

[26] C. Rasch, I. Barillot, P. Remeijer, A. Touw, M. Van Herk, and J. V. Lebesque, "Definition of the prostate in CT and MRI: A multi-observer study," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 43, no. 1, pp. 57–66, 1999.

[27] a Fiorentino, R. Caivano, P. Pedicini, and V. Fusco, "Clinical target volume definition for glioblastoma radiotherapy planning: magnetic resonance imaging and computed tomography.," *Clin. Transl. Oncol.*, vol. 15, no. 9, pp. 754–8, 2013.

[28] A. Thiagarajan, N. Caria, H. Schöder, N. G. Iyer, S. Wolden, R. J. Wong, E. Sherman, M. G. Fury, and N. Lee, "Target volume delineation in oropharyngeal cancer: Impact of PET, MRI, and physical examination," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 83, no. 1, pp. 220–227, 2012.

[29] X. Geets, J. F. Daisne, M. Tomsej, T. Duprez, M. Lonneux, and V. Grégoire, "Impact of the type of imaging modality on target volumes delineation and dose distribution in pharyngo-laryngeal squamous cell carcinoma: comparison between pre- and per-treatment studies," *Radiother. Oncol.*, vol. 78, no. 3, pp. 291–297, 2006.

[30] G. M. Cattaneo, M. Reni, G. Rizzo, P. Castellone, G. L. Ceresoli, C. Cozzarini, A. J. M. Ferreri, P. Passoni, and R. Calandrino, "Target delineation in post-operative radiotherapy of brain gliomas: Interobserver variability and impact of image registration of MR(pre-operative) images on treatment planning CT scans," *Radiother. Oncol.*, vol. 75, no. 2, pp. 217–223, 2005.

[31] a. Wagner, F. Crop, T. Lacornerie, F. Vandevelde, and N. Reynaert, "Use of a liquid ionization chamber for small beam dosimetry," *Phys. Medica*, vol. 28, p. S8, 2012.

[32] D. W. Litzenberg, S. W. Hadley, N. Tyagi, J. M. Balter, R. K. Ten Haken, and I. J. Chetty, "Synchronized dynamic dose reconstruction.," *Med. Phys.*, vol. 34, no. 1, pp. 91–102, 2007.

[33] M. a. Admiraal, D. Schuring, and C. W. Hurkmans, "Dose calculations accounting for breathing motion in stereotactic lung radiotherapy based on 4D-CT and the internal target volume," *Radiother. Oncol.*, vol. 86, pp. 55–60, 2008.

[34] J. V Siebers and H. Zhong, "An energy transfer method for 4D Monte Carlo dose calculation.," *Med. Phys.*, vol. 35, no. 2008, pp. 4096–4105, 2008.

[35] E. Heath, F. Tessier, and I. Kawrakow, "Investigation of voxel warping and energy mapping approaches for fast 4D Monte Carlo dose calculations in deformed geometries using VMC++.," *Phys. Med. Biol.*, vol. 56, pp. 5187–5202, 2011.

[36] J. M. Balter, R. K. Ten Haken, T. S. Lawrence, K. L. Lam, and J. M. Robertson, "Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 36, no. 1, pp. 167–174, 1996.

[37] P.-Y. Y. Bondiau, K. Bénézery, V. Beckendorf, D. Peiffert, J.-P. P. Gérard, X. Mirabel, A. Noël, V. Marchesi, T. Lacornerie, F. Dubus, T. Sarrazin, J. Herault, S. Marcié, G. Angellier, and E. Lartigau, "[CyberKnife robotic stereotactic radiotherapy: technical aspects and medical indications]," *Cancer Radiother.*, vol. 11, no. 6–7, p. 338—344, 2007.

[38] E. Lartigau, X. Mirabel, B. Prevost, T. Lacornerie, F. Dubus, and T. Sarrazin, "Extracranial stereotactic radiotherapy: Preliminary results with the cyberknife??," *Onkologie*, vol. 32, no. 4. pp. 209–215, 2009.

[39] A. S. Beddar, K. Kainz, T. M. Briere, Y. Tsunashima, T. Pan, K. Prado, R. Mohan, M. Gillin, and S. Krishnan, "Correlation between internal fiducial tumor motion and external marker motion for liver tumors imaged with 4D-CT," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 67, no. 2, pp. 630–638, 2007.

[40] L. A. Dawson, K. K. Brock, S. Kazanjian, D. Fitch, C. J. McGinn, T. S. Lawrence, R. K. Ten Haken, and J. Balter, "The reproducibility of organ position using active breathing control (ABC) during liver radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 51, no. 5, pp. 1410–1421, 2001.

[41] H. Shirato, Y. Seppenwoolde, K. Kitamura, R. Onimura, and S. Shimizu, "Intrafractional Tumor Motion: Lung and Liver," *Seminars in Radiation Oncology*, vol. 14, no. 1. pp. 10–18, 2004.

[42] E. D. Brandner, A. Wu, H. Chen, D. Heron, S. Kalnicki, K. Komanduri, K. Gerszten, S. Burton, I. Ahmed, and Z. Shou, "Abdominal organ motion measured using 4D CT," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 65, no. 2, pp. 554–560, 2006.

[43] D. P. Gierga, J. Brewer, G. C. Sharp, M. Betke, C. G. Willett, and G. T. Y. Chen, "The correlation between internal and external markers for abdominal tumors: Implications for respiratory gating," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 61, no. 5, pp. 1551–1558, 2005.

[44] E. W. Pepin, H. Wu, Y. Zhang, and B. Lord, "Correlation and prediction uncertainties in the CyberKnife Synchrony respiratory tracking system," *Med. Phys.*, vol. 38, no. 7, p. 4036, 2011.

[45] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, "The ITK Software Guide," *ITK Softw. Guid.*, vol. Second, no. May, p. 804, 2005.

[46] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 196–205, 2010.

[47] ICRU, *Prescribing, Recording and Reporting Photon-Beam Intensity Modulated Radiation Therapy (IMRT) (ICRU Report 83)*, vol. 10, no. 2. 2010.

[48] T. Nyholm and J. Jonsson, “Counterpoint: Opportunities and Challenges of a Magnetic Resonance Imaging-Only Radiotherapy Work Flow,” *Semin. Radiat. Oncol.*, vol. 24, no. 3, pp. 175–180, 2014.

[49] J. Korhonen, M. Kapanen, J. Keyriläinen, T. Seppälä, and M. Tenhunen, “A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer.,” *Med. Phys.*, vol. 41, p. 11704, 2014.

[50] S. Devic, “MRI simulation for radiotherapy treatment planning,” *Med. Phys.*, vol. 39, p. 6701, 2012.

[51] G. P. Liney, S. C. Owen, A. K. E. Beaumont, V. R. Lazar, D. J. Manton, and A. W. Beavis, “Commissioning of a new wide-bore MRI scanner for radiotherapy planning of head and neck cancer,” *Br. J. Radiol.*, vol. 86, no. 1027, 2013.

[52] A. Johansson, M. Karlsson, J. Yu, T. Asklund, and T. Nyholm, “Voxel-wise uncertainty in CT substitute derived from MRI.,” *Med. Phys.*, vol. 39, no. 6, pp. 3283–90, 2012.

[53] H. Saint-Jalmes, P. A. Eliat, J. Bezy-Wendling, A. Bordelois, and G. Gambarota, “ViP MRI: virtual phantom magnetic resonance imaging,” *MAGMA*, vol. 27, no. 5, pp. 419–424, 2014.

[54] B. Leporq, H. Saint-Jalmes, C. Rabrait, F. Pilleul, O. Guillaud, J. Dumortier, J. Y. Scoazec, and O. Beuf, “Optimization of intra-voxel incoherent motion imaging at 3.0 Tesla for fast liver examination,” *J. Magn. Reson. Imaging*, vol. 41, no. 5, pp. 1209–1217, 2015.

[55] D. Pasquier, T. Lacornerie, M. Vermandel, J. Rousseau, E. Lartigau, and N. Betrouni, “Automatic Segmentation of Pelvic Structures From Magnetic Resonance Images for Prostate Cancer Radiotherapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 68, no. 2, pp. 592–600, 2007.

[56] V. Keereman, Y. Fierens, T. Broux, Y. De Deene, M. Lonneux, and S. Vandenberghe, “MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences,” *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.*, vol. 51, no. 5, pp. 812–818, 2010.

[57] A. Johansson, M. Karlsson, and T. Nyholm, “CT substitute derived from MRI sequences with ultrashort echo time,” *Med Phys*, vol. 38, no. 5, pp. 2708–2714, 2011.

[58] D. M. Grodzki, P. M. Jakob, and B. Heismann, “Ultrashort echo time imaging using pointwise encoding time reduction with radial acquisition (PETRA),” *Magn. Reson. Med.*, vol. 67, no. 2, pp. 510–518, 2012.

[59] K. Eilertsen, A. Vestad, O. Geier, and A. Skretting, “A simulation of MRI based dose calculations on the basis of radiotherapy planning CT images,” *Acta Oncol.*, vol. 47, pp. 1294–1302, 2008.

[60] J. A. Dowling, J. Lambert, J. Parker, O. Salvado, J. Fripp, A. Capp, C. Wratten, J. W. Denham, and P. B. Greer, “An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 83, no. 1, 2012.

[61] U. Schneider, E. Pedroni, and A. Lomax, “The calibration of CT Hounsfield units for radiotherapy treatment planning.,” *Phys. Med. Biol.*, vol. 41, no. 1, pp. 111–24, 1996.

[62] G. Delso, F. Wiesinger, L. I. Sacolick, S. S. Kaushik, D. D. Shanbhag, M. Hullner, and P. Veit-Haibach, “Clinical Evaluation of Zero-Echo-Time MR Imaging for the Segmentation of the Skull,” *J. Nucl. Med.*, vol. 56, no. 3, pp. 417–422, 2015.

[63] F. Wiesinger, L. I. Sacolick, A. Menini, S. S. Kaushik, S. Ahn, P. Veit-Haibach, G. Delso, and D. D. Shanbhag, “Zero TE MR bone imaging in the head,” *Magnetic Resonance in Medicine*, 2015.

[64] M. S. R. Gudur, W. Hara, Q.-T. Le, L. Wang, L. Xing, and R. Li, “A unifying probabilistic Bayesian approach to derive electron density from MRI for radiation therapy treatment planning,” *Phys. Med. Biol.*, vol. 59, pp. 6595–6606, 2014.

[65] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, “Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy,” *Neuroimage*, vol. 46, no. 3, pp. 726–738, 2009.

[66] H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, “Multi-atlas segmentation with joint label fusion,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 611–623, 2013.

[67] T. Sekine, A. Buck, G. Delso, E. ter Voert, M. Huellner, P. Veit-Haibach, and G. Warnock, *Evaluation of atlas-based attenuation correction for integrated PET/MR in human brain - application of a head atlas and comparison to true CT-based attenuation correction*, vol. 4849. 2015.

[68] H. H. Liu, P. Keall, and W. R. Hendee, “D[sub m] rather than D[sub w] should be used in Monte Carlo treatment planning,” *Med. Phys.*, vol. 29, no. 5, p. 922, 2002.

[69] a Ahnesjö, “Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media.,” *Med. Phys.*, vol. 16, no. 4, pp. 577–592, 2013.

[70] B. R. B. Walters, R. Kramer, and I. Kawrakow, “Dose to medium versus dose to water as an estimator of dose to sensitive skeletal tissue.,” *Phys. Med. Biol.*, vol. 55, no. 16, pp. 4535–4546, 2010.

[71] J. V Siebers, P. J. Keall, a E. Nahum, and R. Mohan, “Converting absorbed dose to medium to absorbed dose to water for Monte Carlo based photon beam dose calculations.,” *Phys. Med. Biol.*, vol. 45, no. 4, pp. 983–995, 2000.

[72] C.-M. Ma and J. Li, “Dose specification for radiation therapy: dose to water or dose to medium?,” *Phys. Med. Biol.*, vol. 56, no. 10, pp. 3073–3089, 2011.

[73] P. Andreo, “Dose to ‘water-like’ media or dose to tissue in MV photons radiotherapy treatment planning: still a matter of debate.,” *Phys. Med. Biol.*, vol. 60, no. 1, pp. 309–37, 2015.

[74] W. H. Bragg and R. Kleeman, “XXXIX. On the α particles of radium, and their loss of range in passing through various atoms and molecules,” *Philos. Mag. Ser. 6*, vol. 10, no. 57, pp. 318–340, 1905.

[75] T. E. Burlin, “A general theory of cavity ionisation.,” *Br. J. Radiol.*, vol. 39, no. 466, pp. 727–734, 1966.

[76] M. J. Berger, J. S. Coursey, M. A. Zucker, and J. Chang, “ESTAR : Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions,” *NIST, http://www.nist.gov/pml/data/star/index.cfm*, 2011.

[77] W. R. Nelson, H. Hirayama, and D. W. Rogers, “The EGS4 Code System,” *System*, no. SLAC-265, p. 398, 1985.

[78] H. J. Griffiths, “Tissue Substitutes in Radiation Dosimetry and Measurement.,” *Radiology*, no. October, p. 1989, 1989.

[79] W. Scott, “THICK-TARGET BREMSSTRAHLUNG WHICH INCLUDES MULTIPLE by We Wuyne Scott,” *NASA Tech. Note TN D-4063*, no. August 1967.

[80] N. C. van der Voort van Zyp, M. S. Hoogeman, S. van de Water, P. C. Levendag, B. van der Holt, B. J. M. Heijmen, and J. J. Nuyttens, “Clinical introduction of Monte Carlo treatment planning: A different prescription dose for non-small cell lung cancer according to tumor location and size,” *Radiother. Oncol.*, vol. 96, no. 1, pp. 55–60, 2010.

[81] R. Jeraj and P. J. Keall, “Monte Carlo-based inverse treatment planning.,” *Phys. Med. Biol.*, vol. 44, no. 8, pp. 1885–1896, 1999.

[82] W. Laub, M. Alber, M. Birkner, and F. Nüsslin, “Monte Carlo dose computation for IMRT optimization.,” *Phys. Med. Biol.*, vol. 45, no. 7, pp. 1741–1754, 2000.

[83] P. Lambin, E. Roelofs, B. Reymen, E. R. Velazquez, J. Buijsen, C. M. L. Zegers, S. Carvalho, R. T. H. Leijenaar, G. Nalbantov, C. Oberije, M. Scott Marshall, F. Hoebers, E. G. C. Troost, R. G. P. M. van Stiphout, W. van Elmpt, T. van der Weijden, L. Boersma, V. Valentini, and A. Dekker, “‘Rapid Learning health care in oncology’ - an approach towards decision support systems enabling customised radiotherapy’.,” *Radiother. Oncol.*, vol. 109, no. 1, pp. 159–64, Oct. 2013.

[84] P. Lambin, R. G. P. M. van Stiphout, M. H. W. Starmans, E. Rios-Velazquez, G. Nalbantov, H. J. W. L. Aerts, E. Roelofs, W. van Elmpt, P. C. Boutros, P. Granone, V. Valentini, A. C. Begg, D. De Ruysscher, and A. Dekker, “Predicting outcomes in radiation oncology--multifactorial decision support systems.,” *Nat. Rev. Clin. Oncol.*, vol. 10, no. 1, pp. 27–40, 2013.

[85] E. Roelofs, L. Persoon, S. Nijsten, W. Wiessler, A. Dekker, and P. Lambin, “Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial,” *Radiother. Oncol.*, vol. 108, no. 1, pp. 174–179, 2013.

[86] E. Roelofs, A. Dekker, E. Meldolesi, R. G. P. M. van Stiphout, V. Valentini, and P. Lambin, “International data-sharing for radiotherapy research: an open-source based infrastructure for multicentric clinical data mining.,” *Radiother. Oncol.*, vol. 110, no. 2, pp. 370–4, Feb. 2014.

[87] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J.

Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M. M. Rietbergen, C. R. Leemans, A. Dekker, J. Quackenbush, R. J. Gillies, and P. Lambin, “Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach,” *Nat Commun*, vol. 5, p. 4006, 2014.

[88] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. W. L. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby, and R. J. Gillies, “Radiomics: The process and the challenges,” *Magn. Reson. Imaging*, vol. 30, no. 9, pp. 1234–1248, 2012.

[89] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. P. M. Van Stiphout, P. Granton, C. M. L. Zegers, R. Gillies, R. Boellard, A. Dekker, and H. J. W. L. Aerts, “Radiomics: Extracting more information from medical images using advanced feature analysis,” *Eur. J. Cancer*, vol. 48, no. 4, pp. 441–446, 2012.