



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2019

Patient-specific quality assurance in tomotherapy: independent dose calculation and on-line leaf open time measurement

Schopfer Mathieu

Schopfer Mathieu, 2019, Patient-specific quality assurance in tomotherapy: independent dose calculation and on-line leaf open time measurement

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_2D0B0A17BB926

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

**Département de radiologie médicale
Institut de radiophysique**

**Patient-specific quality assurance in tomotherapy:
independent dose calculation and on-line leaf open
time measurement**

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Mathieu SCHOPFER

Master ès sciences en physique de l'Université de Fribourg

Jury

Prof. Patric Hagmann, Président
Dr Raphaël Moeckli, Directeur de thèse
Prof. François Bochud, Co-directeur de thèse
Prof. Edmond Sterpin, Expert
Prof. Guy Kantor, Expert

Lausanne
(2019)

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur Prof. Patric Hagmann
Directeur·trice de thèse	Monsieur Dr Raphaël Moeckli
Co-directeur·trice	Monsieur Prof. François Bochud
Expert·e·s	Monsieur Prof. Guy Kantor
	Monsieur Prof. Edmond Sterpin

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Mathieu Schopfer

Master of Science en physique Université de Fribourg, Suisse

intitulée

**Patient-specific quality assurance in tomotherapy:
independent dose calculation and on-line leaf open
time measurement**

Lausanne, le 12 avril 2019

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Patric Hagmann



Contents

Remerciements	v
Résumé en français	vi
Abstract	viii
Acronyms	xi
1 Introduction	1
1.1 Tomotherapy system	2
1.1.1 Multileaf collimator and sinogram	3
1.1.2 Jaw collimator	4
1.1.3 Pitch	4
1.1.4 On-board detector	4
1.1.5 Data acquisition system	5
1.2 Quality assurance	5
1.2.1 State of the art in patient-specific quality assurance	6
1.2.2 γ -index dose comparison metric	8
1.3 Quality assurance in tomotherapy	9
1.4 Research introduction and motivation	10
1.4.1 Independent dose calculation	11
1.4.2 On-line leaf open time measurement	12
1.4.3 Delivered dose accuracy investigation through forward calculation	13
2 Overview of the results	15
2.1 Independent dose calculation	15
2.2 On-line leaf open time measurement	16
2.3 Delivered dose accuracy investigation through forward calculation	16

3 Discussion	19
3.1 Independent dose calculation	19
3.2 On-line leaf open time measurement	21
3.3 Delivered dose accuracy investigation through forward calculation	22
4 Conclusion and perspectives	27
Bibliography	29
A Implementation of TomoEDGE in the independent dose calculator CheckTomo	39
B In air and in vivo measurement of the leaf open time in tomotherapy using the on-board detector pulse-by-pulse data	49
C A delivery quality assurance tool based on the actual leaf open times in tomotherapy	61

Remerciements

Une thèse est évidemment un gros investissement personnel, avec des moments de grande motivation, d'autres de découragement. C'est un travail abattu essentiellement en solitaire, mais c'est aussi le fruits de nombreux échanges, de partage d'expérience et d'idées. En ce sens, je souhaite remercier les personnes qui m'ont aidé, soutenu, ou simplement entouré pendant les quatre années que j'ai passées en thèse.

Je souhaite dans un premier temps remercier mes collègues, à commencer par les techniciens en radiologie médicale (TRM). Merci à eux d'avoir manuellement collecté une quantité importante de données sur les installations de tomothérapie. J'ai conscience de la charge, notamment mentale, que ceci représentait, en sus de toute la charge qu'ils supportent déjà. Merci également à mes collègues physiciens du groupe de radiothérapie (GRT) pour les échanges sur la physique médicale et pour les bons moments passés autour d'un café, d'un verre, ou d'un repas dans et hors du CHUV.

Je souhaite également remercier mon codirecteur de thèse, François Bochud, grâce à qui j'ai beaucoup appris, notamment pour rédiger et présenter mes résultats clairement. Merci également à mon directeur de thèse, Raphaël Moeckli, pour avoir lancé mon projet de recherche avec Accuray et d'en avoir géré le côté "politique". J'ai compris grâce à lui l'importance de bien connaître et d'avoir une bonne vision d'ensemble de son domaine de recherche.

Un grand merci à tous les doctorants dont j'ai régulièrement partagé le bureau (Julien, Alexandre, Damien, Anaïs, Marie, Siria, Joshua). C'était très précieux de pouvoir s'entraider sur nos travaux respectifs et de pouvoir simplement passer du temps ensemble. Merci à eux pour leur bonne humeur et leur amitié. Et merci en particulier à Marie d'avoir relu attentivement ce manuscrit et d'avoir pointé certaines de ses faiblesses, me permettant ainsi de les corriger.

Je tiens également à remercier ma compagne, Mélanie, de m'avoir moralement soutenu dans les moments de découragement et de doute et d'avoir toléré que j'aie été parfois peu présent durant les trois derniers mois qui ont précédé le dépôt de ce manuscrit.

Finalement, merci à tous ceux qui ont été présents (ou auraient voulu l'être) lors de ma soutenance publique et ont contribué à la réussite de ce moment.

Résumé en français

La radiothérapie par modulation d'intensité (IMRT) est une technique actuelle de traitement des tumeurs cancéreuses, en constante expansion et en constant développement. La radiothérapie guidée par image (IGRT) permet de vérifier que le patient est correctement positionné avant de délivrer une fraction de dose. La tomothérapie est une technique d'IMRT consistant à délivrer la dose hélicoïdalement avec un faisceau fin (1 à 5 cm). Elle est commercialisée par Accuray sous la marque TomoTherapy[®]. Ce système possède un détecteur de tomодensitométrie mégavolt (MVCT) lui permettant d'implémenter l'IGRT.

Les recommandations internationales préconisent de vérifier l'exactitude de la dose de traitement. Une dose de qualité maximise la probabilité de contrôler la tumeur tout en minimisant le risque de dommages aux tissus sains, donc maximise les chances de rémission sans complication du patient. La pratique la plus courante consiste à mesurer systématiquement une fraction de dose avec un film radiochromique ou une matrice de détecteurs placée dans un fantôme homogène avant de commencer à traiter le patient. Ce protocole d'assurance qualité (QA) vérifie dans un fantôme l'exactitude avec laquelle l'unité de traitement délivre la dose. Cependant, il ne vérifie pas l'exactitude avec laquelle le système de planification du traitement (TPS) a calculé la dose dans les différentes structures anatomiques du patient. De plus, ce protocole ne permet pas de prédire l'écart entre dose délivrée et planifiée dans les organes à risque et volumes cible du patient. Afin de répondre à ces besoins, ce travail de thèse s'est focalisé sur la vérification de la dose calculée par le TPS et sur l'assurance qualité de la délivrance (DQA) de la dose en tomothérapie.

Dans un premier temps, un logiciel de calcul indépendant de la dose, CheckTomo, a été mis à jour suite à la commercialisation de TomoEDGE[™]. TomoEDGE[™] permet une meilleure conformation longitudinale de la dose, préservant ainsi les tissus sains situés en avant et en arrière de la tumeur. CheckTomo a été testé sur des plans TomoPhant (utilisés pour des mesures routinières de QA sur fantôme) avec mâchoires dynamiques et sur 30 plans cliniquement acceptés. Sur les plans TomoPhant, des erreurs de calcul de dose allant jusqu'à 5 % ont été observées. Les plans cliniques ont été soumis à des tests de taux de succès du γ -index (γ -test). Avec des tolérances de 3 %/2 mm (normalisation globale), le taux de succès était inférieur à 95 % dans 53 % des cas. Les résultats des γ -tests sur les plans avec mâchoires dynamiques étaient en moyenne les mêmes que sur des plans avec mâchoires statiques. Ceci tend à indiquer que la faible exactitude globale de CheckTomo ne dépend pas – ou pas uniquement – du mode des mâchoires. Finalement, une erreur de dose globale de 3 % a été appliquée aux plans. Dans ce cas, tous les plans ont échoué le γ -test à 3 %/ 2 mm avec un seuil de succès de 95 %. On conclut à une haute sensibilité de CheckTomo dans le cas d'une erreur globale.

Dans un second temps, les temps d'ouverture des lames (LOT) du collimateur ont été mesurés en ligne puis la dose délivrée a été calculée dans les images de planification du patient. Sur les unités de tomothérapie, le détecteur MVCT mesure la fluence de photons sortant du patient pendant les traitements. Un algorithme de mesure des LOTs à partir des données du détecteur a été développé. Pour le calcul de dose, un calculateur autonome fourni par Accuray a été utilisé. L'algorithme de mesure des LOTs a été testé sur des données de la procédure Daily QA de

l'outil TomoTherapy Quality Assurance (TQA) et sur les données de 25 plans cliniques. Les données cliniques ont été collectées une fois dans l'air et une fois *in vivo* (*i.e.* avec le patient sur la table). L'intensité du signal décroissant fortement avec la fermeture des mâchoires dans le cas d'un faisceau d'une largeur nominale de 5.0 cm, il n'a pas été possible de mesurer les LOTs en dessous d'un écartement des mâchoires de 13 mm. Pour les écartements de taille supérieure et pour les faisceaux avec d'autres largeurs nominales (1 et 2.5 cm), l'algorithme s'est révélé être robuste. Il a permis de mesurer les LOTs sur des plans à mâchoires statiques ou dynamiques sans que l'incertitude ne varie avec la taille du faisceau. De même, l'incertitude de mesure des LOTs n'était pas significativement supérieure *in vivo* que dans l'air. Pour tester la faisabilité de calculer la dose à partir des LOTs mesurés, des "erreurs" aléatoires de LOT (suivant une distribution gaussienne) ont été générées et introduites dans un plan TomoPhant. En tout, six plans avec des erreurs moyennes respectives de -6% , -4% , -2% , 2% , 4% , et 6% ont été générés. La dose de chacun de ces plans a été mesurée avec des chambres d'ionisation placées dans un fantôme au niveau du centroïde des volumes cibles. Les LOTs ont été mesurés à partir des données du détecteur et la dose délivrée a été calculée. Doses mesurées et doses calculées étaient compatibles à 0.5% , indiquant une bonne fiabilité des LOTs mesurés. Finalement, sur des plans cliniques, une corrélation de 0.84 a été observée entre la médiane des erreurs relatives de LOT et le changement de dose dans le volume cible. La médiane des erreurs relatives de LOT pourrait donc être un indicateur de la qualité de la délivrance facile à mesurer.

Un protocole de DQA basé sur la mesure des LOTs pourrait permettre de réduire la charge de travail liée à l'assurance qualité en IMRT en supprimant la nécessité de mesurer la dose délivrée dans un fantôme. Par ailleurs, mesurer les LOTs *in vivo* pendant la délivrance de chaque fraction de dose permettrait de faire un DQA de chaque fraction individuellement. Combiner la mesure en ligne des LOTs à un algorithme indépendant de calcul de dose permettrait de réaliser conjointement un QA de la planification et de la délivrance. De plus, calculer la dose dans les images journalières du patient permettrait de mettre en place un protocole de DQA très proche de la dosimétrie *in vivo*.

Abstract

Intensity modulated radiation therapy (IMRT) is a current technique for the treatment of cancerous tumours, constantly expanding and developing. Image guided radiation therapy (IGRT) is used to verify that the patient is correctly positioned before delivering a dose fraction. Tomotherapy is an IMRT technique that consists of delivering the dose helically with a fine beam (1 to 5 cm). It is marketed by Accuray under the TomoTherapy[®] brand. This system has a megavoltage CT detector (MVCT) to implement the IGRT.

International guidelines recommend to verify the accuracy of the treatment dose. A dose of quality maximises the probability of controlling the tumour while minimising the risk of damage to healthy tissues, thus maximises the chances of complication free remission for the patient. The commonest practice is to systematically measure a dose fraction with a radiochromic film or a detector array placed in a homogeneous phantom before starting to treat the patient. This quality assurance (QA) protocol verifies in a phantom the accuracy with which the treatment unit delivers the dose. However, it does not verify the accuracy with which the treatment planning system (TPS) calculated the dose in the patient's various anatomical structures. Additionally, this protocol does not allow the clinicians to predict the difference between delivered and planned dose in the patient's organs at risk and target volumes. In order to meet these needs, the work done in the context of this thesis focused on the verification of the dose calculated by the TPS and on the dose delivery quality assurance (DQA) in tomotherapy.

Firstly, an independent dose calculation software, CheckTomo, was upgraded following the launch of TomoEDGE[™]. TomoEDGE[™] allows a better longitudinal conformation of the dose, thus preserving healthy tissues located at the tumour's front and back. CheckTomo was tested on TomoPhant plans (used for routine QA measurements in a phantom) with dynamic jaws and on 30 clinically accepted plans. In the target volume of the TomoPhant plans, dose calculation errors up to 5 % were observed. The clinical plans were subjected to γ -index pass rate tests. With tolerances of 3 %/2 mm (global normalisation), the pass rate was less than 95 % in 53 % of the cases. The γ -index pass rates on plans with dynamic jaws were on average the same as on plans with static jaws. This suggests that CheckTomo's overall low accuracy does not depend – or not only – on the jaw mode. Finally, an overall dose error of 3 % was applied to the plans. In this case, all plans failed the γ -index pass rate test with tolerances of 3 %/ 2 mm and a threshold of 95 %. We conclude that CheckTomo is highly sensitive to global errors.

Secondly, the open times of the collimator leaves (LOT) were measured and the delivered dose was calculated in the patient's planning CT images. On tomotherapy units, the MVCT detector measures the photon fluence exiting the patient during treatment. An algorithm for measuring the LOTs based on detector data has been developed. For the dose calculation, a stand-alone calculator provided by Accuray was used. The LOT measurement algorithm was tested using data from the TomoTherapy Quality Assurance (TQA) Daily QA procedure and data from 25 clinical plans. Clinical data were collected once in air and once in vivo (*i.e.* with the patient on the treatment couch). As the signal strength decreases sharply as the jaws narrow for a beam with a nominal width of 5.0 cm, it was not possible to measure the LOTs below a jaw aperture of 13 mm. For larger apertures and for beam with

other nominal widths (1 and 2.5 cm), the algorithm proved to be robust. It allowed LOTs to be measured on static or dynamic jaw plans without any variation of the uncertainty in function the beam width. Similarly, the LOTs measurement uncertainty was not significantly greater *in vivo* than in air. To test the feasibility of calculating the dose from the measured LOTs, random LOT "errors" (following a Gaussian distribution) were generated and introduced into a TomoPhant plan. In total, six plans with average errors of -6 %, -4 %, -2 %, 2 %, 2 %, 4 %, and 6 %, respectively, were generated. The dose of each of these plans was measured with ionisation chambers placed in a phantom at the centroid of the target volumes. The LOTs were measured from the detector data and the delivered dose was calculated. Measured doses and calculated doses corresponded within 0.5 %, indicating a good reliability of the measured LOTs. Finally, on clinical cases, a correlation of 0.84 was observed between the median relative LOT error and the dose change in the target volume. The median relative LOT error could therefore be an easily measurable indicator of the dose delivery quality.

An DQA protocol based on the LOT measurement could reduce the workload related to IMRT quality assurance by eliminating the need to measure the delivered dose in a phantom. In addition, measuring *in vivo* the LOTs during the delivery of each dose fraction would allow the medical physicists to perform a DQA of each fraction individually. Combining the on-line LOT measurement with an independent dose calculation algorithm would allow dose calculation and delivery QA to be performed jointly. In addition, calculating the dose in the patient's daily images would make it possible to set up a DQA protocol very close to the *in vivo* dosimetry.

Acronyms

2D two-dimensionnal.

3D three-dimensionnal.

AAPM American Association of Physicists in Medicine.

ASN Autorité de Sûreté Nucléaire.

CBCT cone beam computed tomography.

CHUV Centre hospitalier universitaire Vaudois.

CS convolution superposition.

CT computed tomography.

DAS data acquisition system.

DD dose difference.

DQA delivery quality assurance.

DTA distance to agreement.

DVH dose volume histogram.

EPID electronic portal imaging device.

ESTRO European Society for Radiotherapy and Oncology.

FOV field of view.

FTP File Transfer Protocol.

FWHM full width at half maximum.

H&N head and neck.

ICRU International Commission on Radiation Units and Measurements.

IGRT image guided radiotherapy.

IMRT intensity modulated radiotherapy.

JAM jaw accelerator machine.

LINAC linear accelerator.

LOT leaf open time.

MLC multileaf collimator.

MOSFET metal-oxide-semiconductor field-effect transistor.

MVCT megavoltage computed tomography.

OAR organ at risk.

OBC on-board computer.

OBD on-board detector.

PDD percentage depth dose.

PTV planning target volume.

QA quality assurance.

SSD source skin distance.

TLD thermoluminescent dosimeter.

TPS treatment planning system.

TQA TomoTherapy Quality Assurance.

Chapter 1

Introduction

Radiation therapy is a therapy technique that intends to treat cancerous tumours with ionising radiations [1]. On common, modern radiation therapy treatment units, electrons are accelerated in a linear accelerator (LINAC) and used either as the treatment beam or, more commonly, to generate a photon beam.

In modern radiation therapy practice, the patient care spans over several weeks and splits in two main phases [1]: planning and treatment. The planning starts by acquiring computed tomography (CT) images of the patient. On these, the planning target volume (PTV) and organs at risk (OARs) are delineated. The PTV is the region that encompass the cancerous tumour and its microscopic extensions. The OARs are organs that can be damaged by the ionising radiations and thus need particular attention, like the heart, the lungs, or the spinal cord. The physicians issue a dose prescription to the PTV, dose constraints on the OARs, and a dose fractionation (the dose is delivered in several fractions). Dose prescription and fractionation depend on the target region [2] and dose constraints are organ-specific [3]. A commercial treatment planning system (TPS) is used to plan the dose delivery. The aim is to cover the PTV with the prescribed dose while not exceeding the constraints on the OARs. The plans must then pass dedicated quality assurance (QA) tests before it is cleared for treatment.

In most situations, the treatment phase lasts over five to six weeks. The patient comes daily to receive one – sometimes two – dose fraction. Minutes before the dose delivery, the patient is positioned on the treatment couch and a new set of images is acquired on-line. The daily images are compared to the planning images to verify target and organ positions. This process is called image guided radiotherapy (IGRT). The images are acquired in the treatment room by the radiation therapy unit itself, often using a cone beam computed tomography (CBCT) scanner. Daily imaging also allows the clinicians to monitor notable changes of the patient's anatomy like inflammations, reduction of the tumour size, or weight loss [4].

The treatment safety and quality have always been major concerns in radiation therapy. A dose accurately delivered to the patient maximises the tumour control potential while minimising the dose delivered to the OARs. It maximises the probability of a successful, complication free outcome of the radiation therapy treatment. Safety intends to avoid hazards with potentially severe adverse effects to the patients and the staff. Radiation therapy centres have dedicated safety protocols and QA programmes. The work presented in this manuscript concentrated



Figure 1.1 – TomoTherapy[®] H[™] series. (Courtesy of Accuray.)

on the latter.

In modern practice, inverse-planned intensity modulated radiotherapy (IMRT) has become the norm [5]. It allows delivering dose distributions of complex shapes. Because of this complexity, IMRT requires dedicated QA programmes and techniques.

The tomotherapy IMRT concept was proposed 26 years ago [6]. It was designed to be a fully integrated IGRT system. It led to the TomoTherapy[®] system owned and commercialised by Accuray (Accuray Inc., Sunnyvale CA, USA).

Section 1.1 presents the tomotherapy system, section 1.2 some general IMRT QA basics, and section 1.3 the QA specificities of the tomotherapy system. Section 1.4 introduces and motivates the contributions that the present thesis brought to the patient-specific QA in tomotherapy. It is divided in three subsections that correspond to a dedicated scientific article in appendix.

In the context of this thesis, an independent dose calculation software was upgraded to keep up with the last tomotherapy evolutions (§ 1.4.1). Additionally, an algorithm to measure on-line the leaf open times (LOTs) was developed (§ 1.4.2). This algorithm was subsequently used to forward calculate the delivered dose and define a dose accuracy predictor based on the LOT errors (§ 1.4.3). Hence, the work done in the context of this thesis covers two QA aspects: the QA of the planned dose and the delivery quality assurance (DQA).

1.1 Tomotherapy system

The tomotherapy concept was inspired by the CT scanner. Figure 1.1 shows a modern tomotherapy treatment unit. The patient lies on the treatment table, usually head first in supine position. The fluence is modulated by a jaw collimator and a multileaf collimator (MLC). There exists two main delivery modalities: helical and topographic, commercialised as the TomoHelical[™] and TomoDirect[™] products, respectively. A helical delivery consists of one passage of the couch through the bore while the gantry rotates at constant speed. A topographic delivery consists of multiple passages (usually four) through the bore. During each passage, the gantry is positioned at a specific angle. The topographic mode is mostly used for tangential breast irradiation [7].

The radiation oncology department of the Centre hospitalier universitaire Vaudois (CHUV) possesses two tomotherapy units (SN137 and SN290), hereafter called TOMO1 and TOMO2, respectively.

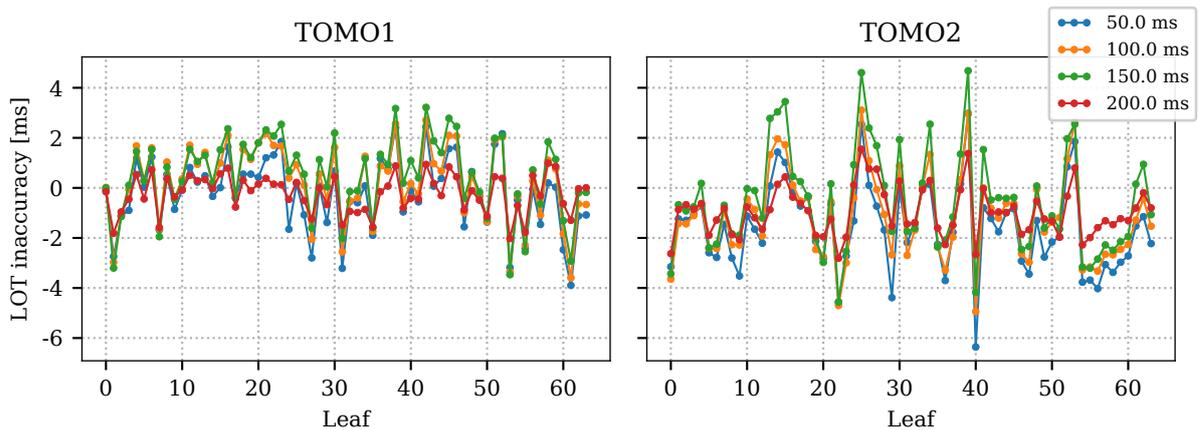


Figure 1.2 – Leaf open time (LOT) inaccuracies measured by the Daily QA module of TomoTherapy Quality Assurance (TQA) on May 15, 2018.

1.1.1 Multileaf collimator and sinogram

Each tomotherapy unit is equipped with a MLC that modulates and collimates laterally the treatment beam. The MLC is composed of 64 binary leaves which are driven by compressed air. Each leaf may open or close independently of the others. It takes less than 20 ms for a leaf to switch from one position to the other.

For helical planning, the gantry rotations are split into 51 sectors called sinogram projections. To modulate the fluence, the optimiser assigns an open time to each leaf at each projection. The LOTs may range from a minimum threshold to the projection duration. The minimum LOT is defined in the jaw accelerator machine (JAM) settings. On CHUV’s units, it was set at 18 ms.

The MLC controller relies on a set of optical sensors to check on-line the leaves state (open, close, or switching) [8]. The controller interrupts the treatment if a leaf takes more than 40 ms to activate (*i.e.* to start moving) and if a leaf takes more than 30 ms to switch state (*i.e.* to move completely from its close to its open position and *vice versa*).

Activation and switching times are both subject to uncertainty. As a result, the actual LOTs differ from the commended LOTs [9–12]. The LOT inaccuracies at one projection are impacted by the number of moving leaves at previous projections, because of pressure variations of the compressed air used to drive the leaves [12]. Additionally, Lissner *et al.* [12] repeated a specific procedure over six months and monitored the LOT inaccuracy of each individual leaf. They reported a maximum drift of the inaccuracy of 3 ms over the six months period.

The LOT inaccuracy is measured every morning during the checks performed by the Daily QA module of the TomoTherapy Quality Assurance (TQA) tool (§ 1.3). The LOTs are measured by opening the leaves 8 by 8 for 50, 100, 150, and 200 ms, once at a jaw aperture of 7 mm and once at 20 mm. Figure 1.2 shows the mean LOT measured on May 15, 2018, during the morning machine checks.

The LOT inaccuracy is referred to as the *leaf latency* within the tomotherapy system. It is a parameter of the machine model in the TPS. When planning a treatment, the TPS first calculates the ideal, theoretical LOTs. At the end of the planning process, it applies the leaf latency correction. This way, the actual LOTs are as close as

possible to the ideal LOTs.

1.1.2 Jaw collimator

To complement the MLC, tomotherapy units are equipped with a jaw collimator which collimates the beam longitudinally. In the first version of the tomotherapy system, the treatment could be delivered at fixed jaw apertures only, *i.e.* at fixed longitudinal field widths.

In 2012, Accuray released the TomoEDGE™ product. It allows the jaws to move during the treatment to improve the dose longitudinal conformity [13–15]. The possibilities offered by the dynamic jaws are likely to increase in the future, *e.g.* for motion management [16].

1.1.3 Pitch

The pitch is the ratio of the table advance per gantry rotation and the field width,

$$\text{pitch} = \frac{\text{distance per rotation}}{\text{field width}}. \quad (1.1)$$

It quantifies how much the beam overlaps with itself over successive rotations. If the pitch is smaller than 1, then the beam will overlap with itself. In tomotherapy, the pitch is always small enough to ensure the complete coverage of the target area. Kissick *et al.* [17] found that the ideal pitch in tomotherapy equals

$$\text{pitch} = \frac{0.86}{N}, \quad N = 2, 3, 4, \dots \quad (1.2)$$

These values minimise the thread effect, thus ensure the best dose uniformity.

1.1.4 On-board detector

The tomotherapy system uses a megavoltage computed tomography (MVCT) on-board detector (OBD) array for on-line imaging. The OBD is made of 640 ionisation chambers filled with high pressure xenon gas. CHUV's treatment units have different detector models, a GE (General Electric, Boston, MA, USA) on TOMO1 and a Hitachi (Hitachi Ltd., Chiyoda, Tokyo, JP) on TOMO2. The differences between both models are the chamber size (1.219 and 1.250 mm, respectively), the surface to centre distance (20 and 4 mm, respectively), and the radius of curvature (110.99 and 99.8 cm, respectively).

On the one hand, the OBD is used for imaging. On the other hand, it measures the fluence exiting the patient during the delivery of a treatment dose fraction. The OBD raw data are collected and temporarily stored by the treatment unit.

1.1.5 Data acquisition system

The data acquisition system (DAS) collects the values of various machine parameters at each LINAC pulse (*i.e.* at a frequency of 80 Hz while imaging and of 300 Hz while delivering a dose fraction). The data logged by the DAS are the OBD output profile, the monitor chambers reading, the couch position, the gantry angle, and the values of various pressure and temperature sensors. These data are stored on the on-board computer (OBC) and can be retrieved to the operator station using the File Transfer Protocol (FTP). On the OBC, these data have a limited lifetime. They are erased whenever a new delivery procedure is loaded on the operator station. Note that the DAS logs neither the reading of the MLC controller optical sensors (*i.e.* the state of the MLC leaves), nor the value of the jaw position encoder.

1.2 Quality assurance

In order to ensure quality and safety of radiation therapy sites and treatments, international organisations recommend performing step-by-step verifications of IMRT devices and treatments [18, 19]. This consists in acceptance tests, commissioning, routine machine QA tests, and patient-specific planning QA and DQA tests. The work done in the context of this thesis, introduced in section 1.4 hereafter, concentrated on the patient-specific QA. The other aspects of IMRT verification are out of the scope of the present manuscript.

Quality assurance aims to prevent errors of all types, from minor incidents to severe accidents. Ford *et al.* [20] proposed to use a clinical scale to score the adverse events ranging from 0 (no harm) to 10 (death). This scale encompasses events of all severity, classifying them by the event's outcome. The event may have minor and temporary, minor but permanent, disabling, life threatening, or lethal outcomes. Ford *et al.* [20] recommend to classify near-misses by the severity of the harm that would have been caused had the adverse event occurred. They also warn about distinguishing between the deterministic effects resulting from an over-irradiation and the non-deterministic effects that may result from underdosing the PTVs. In fact, a PTV underdosage may lead to a failure of the local tumour control, thus to a disease recurrence, and eventually to the patient's death. They recommend to rate underdosage with a score of 9 (*i.e.* consider it as a life-threatening event).

Arnold [21] pointed out two notable properties of radiation therapy errors. Firstly, the large majority of incidents are minor. Secondly, a major accident usually impacts several patients; a major accident is usually not a one-time event, but arises from a systematic failure. To prevent systematic failure, extensive acceptance tests, commissioning, and routine tests of the TPS and treatment unit are required [18]. In addition, the manufacturers have built several interlocks into their treatment systems to automatically detect and prevent possible, known failures (*e.g.* output divergence, leaf positioning errors, etc.). In this context, *known* is important, because it is difficult to anticipate all sources of errors. In that sense, Ford *et al.* [20] recommend to systematically apply *safety barriers*. In other words, any patient plan should systematically go through a dedicated set of critical control points. The choice of the control points is a topic open to debate [22, 23].

Huq *et al.* [23] have reported that the entire patient's IMRT care splits into 91 steps. These steps include all

activities needed to take the patient in charge, from entering the patient’s personal information in the hospital database to delivering each treatment dose fraction. From that, they have identified 216 potential failure sources.

Ford and Terezakis [24] have estimated that error rate in radiation therapy was of 1 in 600, all severity combined, over the period going from 2001 to 2008 in the US. They have reported that 94% of the errors were judged of no clinical significance. This still left them with an error rate resulting in severe injury of 1 in 10,000 patients. Ford and Terezakis [24] noted that the death rate in civil aviation, which is considered to be the gold standard in safety management, is of 1 in 10 million. In Switzerland, the injury rate in road traffic accidents in 2017 was of 1 in 2,400 (seriously injured) and 1 in 475 (lightly injured). The harm caused by radiation therapy incidents during the 2001 and 2008 showed that there was room for practices improvement. It should be noted that the most severe accidents in IMRT occurred while the corresponding relatively new techniques were being implemented throughout western countries and that the US were the forerunners in providing IMRT at a large scale. Finally, to the author knowledge, no severe radiation incidents have occurred (respectively been reported) in the 2010’s. This may indicate a consolidation of the QA practices and that lessons were learned from the mistakes of the preceding decade.

1.2.1 State of the art in patient-specific quality assurance

Patient-specific QA consists of all the actions taken to ensure the quality of the patient’s treatment plan and the safety of the patient’s during their entire care. The most common checks applied in the radiation therapy clinics world-wide were described in many publications (*e.g.* Alber *et al.* [18], Hartford *et al.* [19], Ford *et al.* [22] and Huq *et al.* [23]). Most of the patient-specific QA checks are performed manually by the staff. The checks start to roll out as soon as the patient enters the clinic. As an example, verifying the correctness of the patient’s personal information is already a first QA check. Many QA checks aim to ensure that a plan of good quality is being prepared, *i.e.* that the plan presents an acceptable balance between PTV coverage and OARs sparing. In particular, the dosimetrists, physicians, and physicists verify that the structures are correctly delineated, that the dose prescription is respected, and that the dose is homogeneous. Then, the dose calculated by the TPS should be checked using an independent dose calculation software and the delivered dose is physically measured. These checks are meant to ensure that planned and delivered dose correspond acceptably. The QA of the planned dose intends to ensure by independent calculation means that the dose calculated by the TPS is accurate. Inaccuracies could *e.g.* occur in heterogeneous media or in the presence of a metal implant. The DQA intends to ensure that the treatment unit is able to deliver accurately the planned dose. Delivery inaccuracies can be caused *e.g.* by inaccurate positioning of the MLC. This thesis work was focused on the planned dose QA and dose DQA checks. It does not address the planification steps.

In that sense, the ideal QA method would measure the dose directly where it is delivered, *i.e.* at a specific position within a patient’s OAR or target site. Such an *in vivo* dosimetry method would be the most direct way to ensure the quality of the delivered dose. Yet, it is complicated – if not impossible – to achieve in most treatment situations. Some groups have successfully placed thermoluminescent dosimeter (TLD) or metal-oxide-semiconductor field-effect transistor (MOSFET) dosimeters in the patient’s natural cavities, *e.g.* the mouth or the

oesophagus [25–27]. Though, many target sites are not accessible through a cavity, detector positioning is prone to uncertainty, and the method is highly invasive for the patient. Various indirect *in vivo* dosimetry alternatives exist.

Some groups measured the dose at entrance or exit points on the patient’s skin using silicon diodes, TLD, MOSFET detectors, or radiochromic films, amongst others [28–30]. Others have implemented transit dosimetry based on the fluence measured by an electronic portal imaging device (EPID) [31–34]. With this method, the dose inside the patient must be reconstructed using a back-projection algorithm. A dose reconstruction algorithm accounts for attenuation and scatter within the patient, scatter from the patient to the EPID and scatter within the EPID itself. From the fluence measured on the EPID, the algorithm back-projects the dose in the patient planning or daily CT images. Recent studies showed the development of commercial solutions and spreading of EPID-based *in vivo* dosimetry [35].

The same results may be achieved by forward calculating the delivered dose. For this method, a fluence map – or machine parameters allowing a fluence map to be rebuilt – must be collected on the machine while the dose is delivered. Then, the dose can be forward calculated on the patient’s planning or daily images using a dose calculation model, ideally independent of the TPS (*e.g.* a Monte-Carlo or convolution superposition (CS) algorithm).

Not all centres have implemented *in vivo* dosimetry for IMRT treatments, but most – if not all – have a dedicated pretreatment DQA protocol. The commonest method consists in measuring the delivered dose with a detector array in a water equivalent phantom. The measured dose is compared to the dose distribution calculated in the phantom by the TPS. The accuracy of the measured dose is evaluated using a specific metric, often a γ -index pass rate (§ 1.2.2).

Phantom-based QA protocols have some drawbacks. First, the phantoms are cumbersome and need to be positioned accurately on the treatment couch. This implies acquiring and registering a set of positioning images for each patient QA, which is time consuming. Also, the array detector response may be angular dependent [36, 37]. Moreover, a phantom-based QA provides no information of the discrepancies between planned and delivered dose in the patient’s OARs and PTVs. Finally, such protocols constitute *in fine* DQA tests. They do not verify the accuracy of the dose calculated by the TPS in the patient’s anatomy. They do verify the accuracy of dose calculated in the phantom, though calculating the dose in a homogeneous medium does not much challenge the TPS algorithm. Ideally, independent dose calculations should be conducted in parallel to phantom-based DQA to ensure the accuracy of the dose calculated by the TPS.

By consensus, pretreatment IMRT DQA is considered to be a safety barrier [20]. It should prevent errors to propagate through the workflow. Huq *et al.* [23] noted that the actual QA practice is primarily focused on the treatment device and does not address the question comprehensively. In particular, the author notes that many of the accidents that occurred during the 2000’s were related to a lack of staff training and communication. In that sense, the appropriateness of the current phantom-based DQA practice is subject to debate [22, 23]. Huq *et al.* [23] recommend to address the QA more comprehensively. They propose a comprehensive framework to analyse

the radiation therapy workflow, identify the critical points, and quantify the related risk.

Ford *et al.* [22] showed that the pretreatment phantom-based DQA has a poor effectiveness in catching errors. They analysed a batch of reported high-severity near-misses. The author graded the incidents following the Autorité de Sureté Nucléaire (ASN) scale and considered incidents with a score greater or equal to 4 (thus “accidents” in the ASN scale). The IMRT DQA caught only around 1.4 % of them. This should be no surprise, because the pretreatment DQA cannot catch an error occurring during the planning process, like a PTV being poorly delineated. Ford *et al.* [22] publication should not lead the reader to conclude of the uselessness the pretreatment DQA. Firstly, Ford *et al.* [22] analysed only its effectiveness in detecting high-severity incidents (grader of 4 or 5 on the ASN scale). Though, it is also desirable to avoid incidents of grade 2 or 3. Secondly, they did not address the PTV underdosage problem, because target underdosage does not lead to deterministic outcomes, thus does not enter the ASN scale. Yet, underdosage is considered to be a potentially lethal incident by some of the same authors [20]. As we will see latter, the tomotherapy system can deviate from the planned dose by more than the 5 % recommend in the International Commission on Radiation Units and Measurements (ICRU) report 50 [38] when suboptimally used.

Beyond this, both Ford *et al.* [22] and Huq *et al.* [23] have pointed out that the patient-specific QA consumes a lot of resources and puts a lot of pressure on the staff.

1.2.2 γ -index dose comparison metric

The γ -index is a widely used dose evaluation metric proposed by Low *et al.* [39] in 1998. It is used to compare a measured or calculated dose distribution against a reference. Considering a point of coordinates \mathbf{r}_e and of dose D_e of the evaluated distribution, the problem is to know whether there is a point of the reference distribution that is in the *vicinity* of \mathbf{r}_e with simultaneously a dose *close* to D_e . This implies combining dose difference (DD) ΔD and distance to agreement (DTA) Δd tolerances. The problem can be interpreted as determining whether a point of the reference distribution is located within the ellipse centred in (\mathbf{r}_e, D_e) and of radii Δd and ΔD (formally, a four-dimensional hyper-ellipsoid). This condition can be written as

$$\gamma(\mathbf{r}_e, D_e) = \min \left(\sqrt{\frac{|\mathbf{r}_e - \mathbf{r}_r|^2}{\Delta d^2} + \frac{|D_e - D_r|^2}{\Delta D^2}}, \forall \mathbf{r}_r \right) \leq 1. \quad (1.3)$$

\mathbf{r}_r is a point of the reference dose distribution. This equation is the γ -index test of point \mathbf{r}_e . $\gamma(\mathbf{r}_e, D_e)$ is the point’s γ -index. If the condition is fulfilled, then \mathbf{r}_e is considered to have passed the γ -index test. The whole distribution is then evaluated by calculating the γ -index pass rate, *i.e.* the relative number of points with a γ -index smaller than 1.

The γ -index evaluation of dose distributions is a balance between the DD and DTA tolerances. If the distance to agreement is large, then the dose difference must be small. And *vice versa*. This approach allows the medical physicists to evaluate jointly flat dose and dose gradient regions. In a flat dose region, the γ -index tests predominantly the dose level accuracy through the DD tolerance. Reciprocally, in a dose gradient region, it tests

predominantly the gradient correct location through the DTA tolerance.

The difficulty with the γ -index resides in choosing the tolerances and interpreting the results. The American Association of Physicists in Medicine (AAPM) Task Group 218 issued recommendations [40] concerning the tolerances and pass rate thresholds to apply to measurement-based QA verifications. Tolerances and pass rates depend on the type of test (physical or clinical) and of detector (ionisation chamber, array, *etc.*). In particular, the test will have a different sensitivity depending whether the DD is defined locally (in function of the dose D_e of the point to evaluate) or globally (in function of the maximum dose of the evaluated or reference distribution). Also, points in the low dose region will bias the pass rate if the dose tolerance is too high. They should be excluded from the statistics, the recommendation being to neglect all points below 10 % of the dose distribution maximum [40]. Finally, the acceptable pass rate threshold must be determined to accept/reject a plan.

The performance of tests based on the γ -index has been extensively discussed in the literature [41–49]. Some authors have studied the correlation (or lack of) between the γ -index pass rate and the plan quality (*e.g.* the dose coverage of the PTV). Others have studied the sensitivity and specificity of the γ -index pass rate test. They reported failures to detect or highlight clinically relevant errors. Therefore, alternatives to the γ -index were proposed. In particular, Zhen *et al.* [41] proposed to turn to metrics based on the dose volume histogram (DVHs). Stojadinovic *et al.* [48] proposed a divide-and-conquer approach of the γ -index by analysing high dose, dose gradient, and low dose regions separately.

1.3 Quality assurance in tomotherapy

Booklet no. 9 of the European Society for Radiotherapy and Oncology (ESTRO) provides specific guidelines for the verification of the helical tomotherapy system [18]. The AAPM set up a task group who issued recommendations for machine and patient-specific QA in helical tomotherapy [50]. Machine QA tests were also extensively presented and discussed by Fenwick *et al.* [10].

TQA is a tool to check the tomotherapy system [51]. It may run health and accuracy check of any machine component or subsystem. Daily QA is one module of TQA for daily routine checks of the machine. In particular, it measures daily the leaf latencies (§ 1.1.1).

For routine TPS QA, Accuray provides the synthetic “TomoPhant” plans. Point dose measurements are performed with ionisation chambers in the Cheese Phantom in gradient and flat dose regions. There is a TomoPhant plan for each delivery mode (helical or topographic) and each jaw mode (static or dynamic). The plans have two off-axis cylindrical targets of 6 cm.

The original patient-specific DQA recommended by Accuray was to measure a two-dimensionnal (2D) dose distribution with a radiochromic film and an absolute point dose with an ionisation chamber in the Cheese Phantom. At least four commercial systems have been validated to replace the film measurements. They use a diodes or ionisation chambers array placed in a phantom. Amongst the validated systems, one may find Octavius (PTW, Freiburg, Germany) [52], Delta4 (ScandiDos, Uppsala, Sweden) [53], MapCHECK (Sun Nuclear, Melbourne FL,

USA) [54], and ArcCHECK (Sun Nuclear, Melbourne FL, USA) [55].

The first *in vivo* dosimetry attempts in tomotherapy were conducted by Kron *et al.* [56]. They placed dosimetric films on the treatment couch under the patient. Qi *et al.* [57] used MOSFET detectors during the treatment of naso-pharyngeal carcinomas. They used moulded oral plates to position the MOSFET dosimeters and measured *in vivo* the dose to the patient's tongue. Alnaghy *et al.* [58] measured *in vivo* the dose to the anterior rectal wall using MOSFET dosimeters during prostate boost treatments.

Three groups proposed dose reconstruction algorithms for tomotherapy [59–63]. There exists also a commercial *in vivo* and pretreatment DQA dosimetry software, Dosimetry Check (Math Resolutions, Columbia MD, USA), which was recently bought and incorporated into RadCalc (LAP, Lüneburg, Germany). It uses either the primary fluence or the exit fluence (if a phantom or a patient is on the treatment couch) measured by the OBD (§ 1.1.4) to reconstruct three-dimensional (3D) dose distributions in the phantom's or patient's CT images. It relies on a pencil beam collapsed cone convolution algorithm. Some authors reported to have implemented clinically RadCalc on tomotherapy at their satisfaction [64, 65]. Mezzenga *et al.* [64] reconstructed “*in vivo*” the dose from the exit fluence of three plans on the Cheese Phantom. They placed ionisation chambers in the target regions and reported a mean dose deviation between the ionisation chamber measurements and RadCalc calculation of $-1.4 \% \pm 1.6 \%$.

1.4 Research introduction and motivation

In vivo dosimetry is still not widely spread in tomotherapy. The common patient-specific DQA consists in dose measurements in a water equivalent phantom, implying tedious repetition of DQA measurements and absence of patient-specific QA on the TPS side.

In the context of this thesis, the upgrade of an independent dose calculation software was developed and validated (§ 1.4.1). In addition, to set the basis of an automated DQA tool, an algorithm was developed to measure on-line the LOTs (§ 1.4.2) and was used to forward calculate the dose distribution in the patient's CT images (§ 1.4.3). This work aims to eventually alleviate the load of patient-specific DQA for medical physicists and to propose a more dose-oriented acceptance criterion of the treatment plans.

To ensure the accurate delivery of the dose, one must ensure the accuracy of:

- the accelerator output,
- the patient's position relatively to the radiation source,
- the jaw collimator and MLC operation.

The parameters of a majority of the machine components are logged by the DAS (§ 1.1.5). Amongst others, the DAS data contain the monitor chambers reading, the couch position, and the gantry angle. During the treatment delivery, these values are logged at each LINAC pulse, *i.e.* at a frequency of 300 Hz.

The dose output is controlled by a servomechanism. Thanks to that, Moutrie *et al.* [66] reported a stable dose output both intrafractions and interfractions. The intrafraction output standard deviation was of 0.1 %.

Concerning the couch position and the gantry angle, Handsfield *et al.* [67] reported that the values logged by the DAS agreed with what was expected from the treatment plan. Neither did they report any quantified error, nor did they specify whether they had applied particular tolerances. Though, couch position and gantry angle are measured with uncertainties of 0.1 mm and 0.1°, respectively.

The jaw positions and the states of the MLC leaves (open, close, or switching) are not logged by the DAS. Though, Lee *et al.* [68] verified the accuracy of jaw motion and position during nine months. They monitored the jaw encoder error during a procedure with sweeping jaws, reporting a maximum error of 0.4 mm (in projection in the isocentre plane). In the opposite, many authors have reported dosimetric errors related to the MLC inaccuracy [63, 67, 69–71]. The pitch was identified as the problematic planning parameter. Pitches of 0.215 and smaller, *i.e.* $N \geq 4$ (§ 1.1.3), caused significant mean relative LOT discrepancies [69, 72]. The tighter the pitch, the more the beam overlaps with itself between successive rotations. Thus, the smaller the mean LOT required to deliver a specific dose and the higher the relative LOT inaccuracy.

In summary, during normal operation of the machine, all components are reliable and operate with the desired accuracy, except the MLC. Therefore, the work presented in sections 1.4.2 and 1.4.3 focuses on the effect of the MLC inaccuracies on the delivered dose. Obviously, no QA check but DQA measurements can detect significant LOT discrepancy. In section 1.4.3, a method to detect such errors either in air for pretreatment DQA or *in vivo* is presented. The method relies on the detector raw data (§ 1.1.4) and could thus be fully automated. This is very interesting, principally for two reasons. Firstly, as Ford *et al.* [22] calculated, the pretreatment DQA catches only 1.4 % of severe potential errors. From this perspective, there is a strong interest in rationalising the efforts nowadays put into manually performing pretreatment DQA checks. Secondly, as Arnold [21] pointed out, incidents usually result from systematic failures. In that sense, it is interesting to have tools that could systematically capture errors, before and while the dose is delivered.

1.4.1 Independent dose calculation

Thomas *et al.* [73] developed in MATLAB (MathWorks Inc, Natick MA, USA) an independent dose calculation software for tomotherapy and called it CheckTomo. The software reads in patient’s CT images and treatment plans from DICOM-RT files exported from the TPS. It allows the dose to be calculated at several points on a regular 3D grid. Gibbons *et al.* [74] had previously proposed an independent dose calculation software, but the calculation was limited to a single point.

The calculation model of CheckTomo is based on empirical factors measured in fixed source skin distance (SSD) condition, namely dose rate, field output factors, beam intensity profiles at different depths, and percentage depth dose (PDD). The dose calculation is patient-specific. The software calculates the water equivalent depth (the radiological path length) of the grid points with respect to the beam incidence angles and the electronic density along the beam path. The software was originally released before the dynamic jaw mode of tomotherapy. Therefore, CheckTomo did not offer the possibility to calculate the dose of dynamic jaw plans.

In the context of this thesis, it was investigated whether CheckTomo could be upgraded to calculate the dose

of dynamic jaw plans. A method was eventually developed and specific tests were performed. This work has been reported in article A, published in 2017.

CheckTomo algorithm models only longitudinally on-axis beams with longitudinally symmetric intensity profiles. Yet, the very purpose of tomotherapy's dynamic jaw mode is to produce off-axis beams. Additionally, it turns out that these beams have longitudinally asymmetric intensity profiles. The main development challenge was thus to reach an acceptable symmetric approximation of the profiles of the off-axis beams. The off-axis beam profiles were measured at field widths of 10, 18, and 25 mm. A symmetric curve was fit on each profile (article A, figure 2). (The off-axis beam profiles at other field widths are calculated on the fly by spline interpolation.)

The dose calculation accuracy of the dynamic jaw plans was evaluated on 5 simple synthetic plans and on 30 real clinical plans that had been approved for treatment. First, the dose of dynamic jaw TomoPhant plans (§ 1.3) with various PTV lengths was calculated. The error between CheckTomo and TPS longitudinal dose profiles was calculated. Then, the dose of clinical dynamic jaw treatment plans was calculated. There were 10 abdomen, 10 head and neck (H&N), and 10 breast plans. CheckTomo and TPS dose distributions were compared with γ -index pass rates (§ 1.2.2). Various DD (global normalisation) and DTA tolerances were used. To calculate the pass rates, the minimum dose threshold was fixed at 50 % of the dose maximum (*i.e.* points below the 50 % isodose level were excluded from the pass rate calculations). Finally, some errors were simulated on the treatment plans that had had a pass rate above 95 % for tolerances of 3 %/2 mm. It was checked whether these errors could be detected by a γ -index pass rate test with tolerances of 3 %/ 2 mm and a pass rate threshold of 95 %.

1.4.2 On-line leaf open time measurement

As mentioned in the introduction of this section, there is a strong interest in measuring the actual LOTs to calculate more accurately the delivered dose. An algorithm to measure the LOTs from the OBD pulse-by-pulse data was developed. It was extensively presented in article B.

The OBD is active during the treatment dose delivery and collects the fluence exiting the patient (§ 1.1.4). It was possible to map each leaf to a specific detector channel. Then, determining the leaf state consisted in observing the signal in the leaf's channel. The period of time during which the leaf is open corresponds to a peak in the detector signal. The related LOT was defined as the peak's full width at half maximum (FWHM).

The principal difficulty was to ensure the algorithm accuracy without having the opportunity to rely on an external reference. The state of the MLC leaves is checked on-line by optical sensors [8], but the sensors output is not logged by the DAS. As a workaround, the LOTs measured by the algorithm were checked step-by-step, starting in simple conditions and adding complexity layers:

1. The LOT errors were compared to those reported by the Daily QA module of TQA. This meant measuring the LOTs without attenuation material in the beam path and under static jaws.
2. The effect of the field width on the LOT measurement was investigated in air on dynamic jaw procedures.

3. On the same dynamic jaw procedures, the effect of the beam attenuation by the patient on the measured LOTs was investigated.

1.4.3 Delivered dose accuracy investigation through forward calculation

The aim of the work presented in article C was to propose and to validate clinically a DQA method with a dose-oriented approach. In particular, it was investigated whether the LOT discrepancies related to the dose in the PTV, with the idea to propose an indicator of the dose delivery quality based on the measured LOTs.

The principle is to measure the LOTs with the algorithm presented in article B and to calculate the dose actually delivered by replacing the planned LOTs with the measured LOTs in the treatment plan. In this work, the calculations were performed with Accuray's standalone dose calculator [75].

It was first tested whether the method would successfully highlight LOT discrepancies and accurately calculate the delivered dose in synthetic conditions. Six plans with intentional random LOT errors were generated from the TomoPhant plan with a dynamic jaw field of 2.5 cm. The random errors followed Gaussian distributions with means of -6% , -4% , -2% , 2% , 4% , and 6% , respectively. Each modified plan was delivered. The dose was measured with two ionisation chambers, one at the centroid of each of the TomoPhant plan targets. The OBD raw data were collected and the LOTs subsequently calculated. The delivered dose was calculated using the measured LOTs and compared to the dose measured with the ionisation chambers.

60 clinical treatment plans were selected randomly, 30 on each of CHUV's treatment unit. The plans were delivered in air with the couch retracted out of the bore (static couch DQA procedures). For each plan, the delivered dose was calculated using the measured LOTs. The correlation between the mean relative LOT discrepancy and the change of the dose covering 95 % of the PTV (D_{95}) was calculated. The plans pitch were of 0.287 or 0.215 ($N = 3$ or $N = 4$, *c.f.* § 1.1.3). 0.215 is a value known to cause significant mean relative LOT discrepancies (§ 1.4).

Chapter 2

Overview of the results

In this chapter, we briefly summarise the results of the three scientific articles in the appendix.

2.1 Independent dose calculation

First, the dose calculation accuracy of CheckTomo on plans with dynamic jaws is reported. The calculation accuracy was evaluated on TomoPhant synthetic plans and on clinical plans. Then, the sensitivity of CheckTomo to simulated errors is reported.

The error committed when symmetrically modelling the off-axis beam asymmetrical intensity profiles was estimated. Measured (asymmetrical) and fitted (symmetrical) beam profiles are represented in figure 2 of article A. The maximum error, relatively to the on-axis beam output, was below 3 %. It is not obvious how this relates to the inaccuracy of the calculated dose.

It was more interesting to investigate whether CheckTomo calculated accurately the dose in synthetic conditions, *i.e.* on dynamic jaw TomoPhant plans. The results are reported in figure 4 of article A and reproduced here with

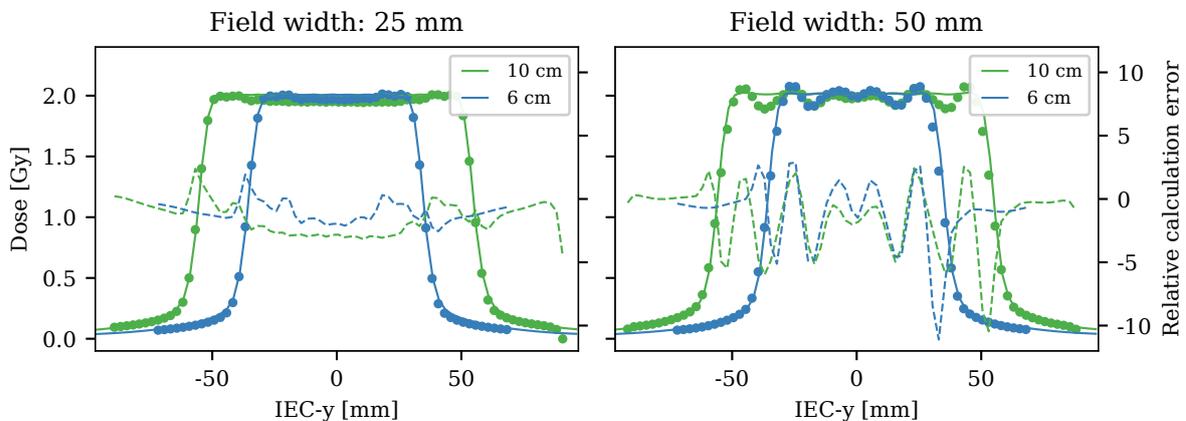


Figure 2.1 – Longitudinal dose profiles of the dynamic jaw TomoPhant plans for two PTV lengths, computed by the TPS (plain lines) and CheckTomo (dots), with the relative calculation error of CheckTomo (dashed lines). In this case, the relative jaw penumbral filter were sampled at 10, 18, and 25 mm.

additionally the relative error (figure 2.1). The error was normalised by the planned dose at the centre of the profile (2.01 Gy). On plans with a field width of 2.5 cm, the dose error ranged from -3.2% to 2.4% . With a field width of 5.0 cm, it increased to a range from -11.1% to 2.9% . The largest errors were at the edges of the dose plateau.

CheckTomo and TPS dose distributions were compared on 10 abdomen, 10 H&N, and 10 breast plans with γ -index pass rates (article A tables 2, 3, and 4, respectively). The DD (global normalisation) tolerance had the most influence on the pass rates. Plans in the abdomen region were the most accurately calculated. In the opposite, plans in the H&N region had the lowest pass rates.

The accuracy of the γ -index pass rate test was calculated (article A table 5). In this context, “accuracy” is the ability of the test to detect correctly if plans are clinically acceptable or not. With tolerances of $3\%/2\text{ mm}$ and a pass rate threshold of 95% , the accuracy of the γ -index pass rate test was of 47% . This means that the pass rate of 53% of the analysed clinical plans was below 95% , even though all plans were clinically acceptable. The accuracy was of 100% with a DD tolerance of 6% .

Finally, the pass rate (tolerances of $3\%/2\text{ mm}$) of all plans fell below the 95% threshold when either a dose shift of 3% or a spatial shift of 4 mm was applied (article A table 6).

2.2 On-line leaf open time measurement

First, the agreement between the LOTs measured by the algorithm and by TQA is reported. Then, the effects of the field width and of the beam attenuation on the LOT measurement are reported.

The agreement between the LOTs measured by the algorithm and those measured by TQA was of $0.0 \pm 0.3\text{ ms}$ (one standard deviation).

The standard deviation and mean of the discrepancies between measured and planned LOTs did not depend on the jaw aperture in dynamic J20 procedures (article B figure 5). Yet, the detector data were not exploitable at the beginning and at the end of dynamic jaw J42 procedures, below a jaw aperture of 12.4 mm (article B table 2). This is because the beam is off the source axis and the detector is in the beam umbra. Finally, the discrepancy distributions were the same in air and *in vivo* (article B figure 6).

LOT discrepancies were measured up to 45 ms . These large values were related to either a short (below 30 ms) or a large (within 30 ms of the projection time) planned LOT. Planned LOTs may range from 18 ms to the projection time (§ 1.1.1).

2.3 Delivered dose accuracy investigation through forward calculation

First, the output of the measurements on the modified TomoPhant plans are reported. The discrepancies between measured and planned LOTs were calculated. The dose agreement between ionisation chamber measurements and calculations based on the measured LOTs was established. Then, correlations calculated between the mean relative LOT discrepancy and the dose change in the PTV (D_{95}) on two sets of 30 clinical plans are reported.

On the modified TomoPhant plans, the intentional LOT discrepancies could be highlighted. On each of the TomoPhant plans (one reference and six modified plans), the mean relative discrepancy between measured and planned LOTs was of 1.10 ± 0.05 % on TOMO1 and of 0.02 ± 0.03 % on TOMO2 ($N = 14$, one standard deviation, article C figures 2 and 3).

The agreement between the dose calculated from the measured LOTs and the dose measured with ionisation chambers in the Cheese Phantom on the TomoPhant plans was of 0.2 ± 0.3 % on TOMO1 and of 0.1 ± 0.3 % on TOMO2 ($N = 28$, one standard deviation). The maximum difference was of 1 % (article C figure 4).

On the clinical plans, the correlation between the mean relative LOT discrepancy and the dose deviation in the PTV was of 0.76 ($p \approx 10^{-15}$) on TOMO1 and 0.65 ($p \approx 10^{-10}$) on TOMO2 (article C figure 6). There was no such correlation between the γ -index pass rate and the PTV dose (article C table 1 and figure 7).

Chapter 3

Discussion

Independently calculating the planning dose tests the TPS algorithm in heterogeneous media. It constitutes a patient-specific QA on the TPS side. Measuring on-line the machine parameters is a patient-specific DQA. In that sense, the work presented in the article A and that presented in articles B and C are complementary, because patient-specific DQA and TPS QA are complementary.

Admittedly, we were misunderstanding this when we wrote article A:

“Performing an independent dose calculation with CheckTomo is not as comprehensive as actually measuring it during a QA procedure, in that sense that it performs no control on the machine side.”
(Article A, § 3.B.2, 98)

Measuring the dose in a water equivalent phantom is indeed a machine DQA, but the measured dose is compared to a dose calculated in a homogeneous medium. In that sense, it lacks a true verification of the calculated dose accuracy in the patient’s anatomical structures.

The extent and type of IMRT verifications to perform is an open topic. In particular, whether or not to base the IMRT QA on softwares and/or machine logs have already been heavily debated [76, 77]. Here, the independent dose calculation software CheckTomo relies on independently measured beam data and an independent algorithm. The LOT measurement algorithm is based on machine logged *raw* values. The algorithm itself is independent.

3.1 Independent dose calculation

An upgrade of an existing independent dose calculation software for tomotherapy, CheckTomo, was implemented. The upgraded was required to follow the TomoEDGE™ evolution of the tomotherapy system. Article A presented upgrade implementation in the calculation model of CheckTomo and presented how the accuracy of the calculated dose distributions was investigated.

The inaccuracy of the dose calculated on the synthetic TomoPhant plans was high, up to 11 % with a field width of 5.0 cm (figure 2.1). This inaccuracy could be reduced by increasing the sampling of the off-axis beam profiles (figure 3.1). The beam profiles were originally sampled at 10, 18, and 25 mm. Meanwhile, four samples have been

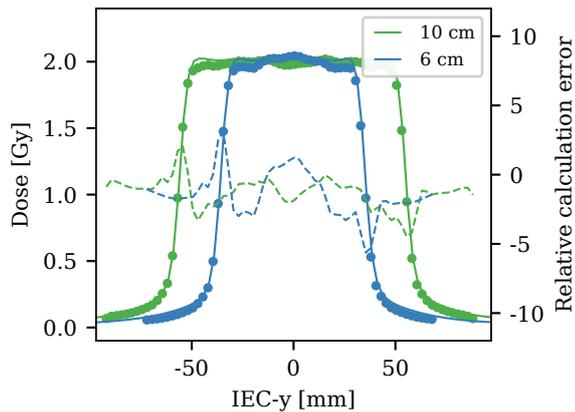


Figure 3.1 – Longitudinal dose profiles of the 5 cm dynamic jaw TomoPhant plans for two PTV lengths, computed by the TPS (plain lines) and CheckTomo (dots), with the relative calculation error of CheckTomo (dashed lines). In this case, the relative jaw penumbral filter had been sampled at 10, 18, 25, 30, 35, 40, and 45 mm.

added at 30, 35, 40, and 45 mm. With this modification, the error committed when calculating the 5 cm dynamic jaw TomoPhant plans ranged from -5.6% to 3.1% (previously from -11.1% to 2.9%). Yet, only 2 of the 30 clinical treatments analysed in article A were planned with a field width of 5.0 cm. Of those, only one (article A table 2, plan A06) had low γ -pass rates. Thus, the sampling modification introduced here would not change the overall results and conclusions of article A. Anyway, a calculation error greater than 5 % in simple conditions like the TomoPhant plan is not satisfactory.

Thomas *et al.* [73] reported mean γ -index pass rates (3 %/3 mm) of 96.3 % and 89.2 % on prostate and H&N static jaw plans, respectively. In article A, mean pass rates (3 %/3 mm) of 95.9 % and 90.3 % were reported on abdomen and H&N dynamic jaw plans, respectively. This indicates that the software calculates the dose as accurately on static jaw plans as on dynamic jaw plans. In that sense, the upgrade was successful.

The γ -index pass rate test with tolerances of 3 %/2 mm and a threshold of 95 % detected all the errors simulated by applying a global dose shift of 3 % or position shift of 4 mm. This means that CheckTomo had a high sensitivity towards global errors. Though, regarding the high dose inaccuracy near the dose plateau on the TomoPhant plans (figure 3.1), the sensitivity to errors of other types should be investigated (*e.g.* cold spots [78]).

Which is more, with tolerances of 3 %/2 mm, the software had a low overall accuracy. Only 47 % of the plans passed the γ -index test. As there are no measurement or positioning uncertainty when comparing two *calculated* dose distributions, tolerances of 3 %/2 mm seems reasonable. Thus, CheckTomo would yield too many false alarms to be a reliable, useful QA tool.

Karlsson *et al.* [79] claimed in ESTRO booklet 10 that an independent dose calculation is a useful QA procedure if an advanced algorithm is used. The question is to define what “advanced” means in this context. The booklet recommends to use kernel-based models. CheckTomo relies on an empirical factor model measured in a water tank, so not on an “advanced” algorithm in the recommendations’ sense.

The results reported and discussed here support ESTRO recommendations. In practice, CheckTomo should be (and will likely be) superseded by more advanced algorithms.

3.2 On-line leaf open time measurement

An on-line LOT measurement algorithm was developed. The results showed that the LOTs could be measured for the various jaw and field width settings, except at the beginning and at the end of dynamic J42 procedures. In this situation, the LOTs could not be measured at jaw apertures narrower than 12.4 mm. The results also showed that the LOTs could be measured in air during a DQA procedure or *in vivo* with the patient on the treatment couch during the delivery of a dose fraction. The main issue was to ensure the accuracy of the measured LOTs without having the possibility to compare them to an external reference.

To test the algorithm, the measurement conditions went from simplest to most complex. First, the LOTs were measured in air with static jaws, then in air with dynamic jaws, and finally *in vivo* with dynamic jaws. The LOTs measured by the algorithm agreed closely to the LOTs measured by TQA. This indicates a low measurement uncertainty in simple conditions. Then, successively adding complexity to the measurement conditions did not impede the LOT measurements, *i.e.* the measurement uncertainty did not increase. In particular, in air and *in vivo* discrepancy distributions had the same means and standard deviations (article B figure 6).

As mentioned, the LOT errors measured on Daily QA raw data agreed very well with those reported by TQA. Thus, if the LOTs measured by TQA are accurate, those measured by the algorithm are also accurate. As complexifying the measurement conditions did not impede the LOT measurement, the LOTs should be as accurately measured in the simplest setting as in the more complex ones.

As mentioned in section 1.3, the machine daily checks are performed with the Daily QA module of TQA. During the checks, the leaf opens 8 by 8 during 50, 100, 150, and 200 ms. This pattern is repeated twice, once at a jaw opening of 7 mm and once at 20 mm. Thus, the LOT error measured for a specific leaf on a specific day is the average of only two values. This should neither have artificially masked the actual, intrinsic LOT variability, nor the uncertainty of the measured LOTs. In other words, measured LOTs were compared to other measured LOTs, not means of repeated measurements to means of repeated measurements.

Additionally, Chen *et al.* [72] used a similar algorithm to measure the LOTs in air and *in vivo*. They had the opportunity to compare the output of their algorithm to the optical sensors of the MLC controller on static jaw plans. (TomoEDGE™ had not been released at that time.) Thus, their algorithm was validated against an external reference on static jaw data. The algorithm presented here parses data of static jaw procedures very similarly to that of Chen *et al.* [72]. Both algorithms should yield the same LOTs on static jaw data. Therefore, if one algorithm measured reliably the LOTs of static jaw procedures, the other must have done as well.

One last problem must be addressed: the lack of consensus regarding the very definition of the LOT in tomotherapy. In article B, the LOT was defined as the FWHM of the open leaf signal regions (article B, § II.B.3). Taking the signal FWHM, *i.e.* choosing $\tau_{OC} = 0.5$, was arbitrary. When looking at the detector signal, it is obvious when a leaf is either fully closed or fully open. Though, there is no unique way of handling the signal gradients (corresponding to the leaf switching between either states).

Amongst the other authors who also measured the LOTs from the OBD pulse-by-pulse data, Chen *et al.* [72]

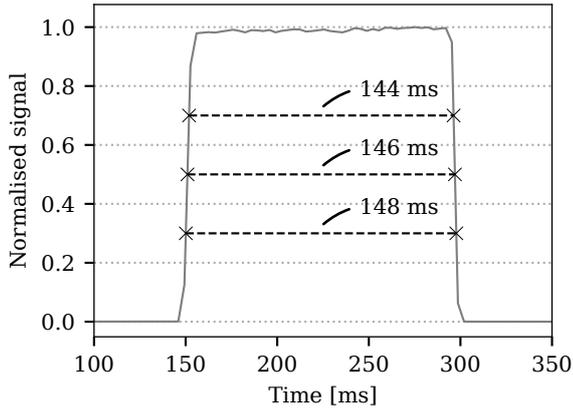


Figure 3.2 – Normalised signal under leaf 26 at projection 69 measured *in vivo* during a mediastinal treatment on TOMO2 (article B table 1, patient 1). The measured leaf open time (LOT) is given at different values of the open-close threshold τ_{OC} (article B, § II.B.3).

also took the signal FWHM. Proceeding differently, Sevillano *et al.* [11] calculated an estimate of the normalised primary leaf fluence from the OBD raw signal. They defined the LOT as the integral over time of the primary fluence. In the context of article B, this would roughly consist in taking the integral over time of the deconvolved and normalised OBD signal (article B, § II.B.2). Note that this would not work as such *in vivo*.

To estimate how the choice of the threshold τ_{OC} influences the LOT measurement, the LOT in an example case was measured with different values of τ_{OC} (figure 3.2). Lowering or raising τ_{OC} by 0.2 yielded LOT variations of 2 ms. Therefore, the measurement inaccuracy must range from -2 to 2 ms.

In summary, the LOT measurement was impeded neither by the jaw aperture, nor by the beam attenuation. Yet, the relation between the LOT and the leaf position has never been formally defined. It is not clear if a leaf should be considered open once it has moved half way through the beam, or when it starts moving, or when it is finished moving, *etc.* To choose the optimal threshold, the measured LOTs must be related to the measurable quantity they control: the photon fluence. The first half of article C concentrated on assessing whether the measured LOTs led to accurate dose calculation.

3.3 Delivered dose accuracy investigation through forward calculation

The work presented in article C brought back to the clinic the LOT measurement algorithm presented in article B. An indirect validation of the measured LOTs accuracy through dose measurements and calculations in synthetic TomoPhant plans was provided. The correlation between the relative LOT discrepancy and the dose change in the PTV was investigated, allowing a LOT-based indicator of the dose delivery quality to be proposed.

On the synthetic TomoPhant plans, the measurements supported the feasibility of detecting LOT errors from the OBD pulse-by-pulse data (article C figure 3). Figure 3 of article C shows the mean agreement between measured and planned LOTs for all TomoPhant plans. On TOMO1, the measured LOTs were in average 1 % greater than expected. This phenomenon did not happen on TOMO2. The mean *LOT* accuracy was of 2.0 ± 0.1 ms on TOMO1 and of 0.1 ± 0.1 ms on TOMO2. (Note that this tells nothing about the *measurement* accuracy itself.) No leaf latency correction was applied to the TomoPhant plans (§ 1.1.1). Interestingly, the latency correction that is normally applied to clinical plans is of 2.4 ms on TOMO1 and 0.5 ms on TOMO2. If the latency correction had

been applied to the TomoPhant plans, the mean LOT accuracy would be within 0.4 ms on both units. The aim of measuring the LOTs on synthetic TomoPhant plans was to assess whether the measured LOTs allowed the delivered dose to be accurately calculated. The aim was not to measure the accuracy of the leaf latency correction.

The intentional discrepancies introduced on the TomoPhant plans were chosen to be Gaussian because most of the LOT error distributions reported in article B were bell-shaped (article B figure 6). Thus, the simulated errors were comparable to the real LOT errors caused by the intrinsic MLC inaccuracy. Gross errors, like a leaf being stuck open or closed, were not tested. They should not occur in the first place because the leaf controller would interrupt the treatment in such situations. This work concentrated on testing uncaptured, yet clinically relevant discrepancies.

Using the measured LOTs as input for the dose calculation led to an accurate prediction of the delivered dose. The agreement between the ionisation chambers and the calculated dose was high (0.2 ± 0.3 % on TOMO1 and 0.1 ± 0.3 % on TOMO2, article C figure 4). In the TomoPhant plans used in the study, a mean LOT difference of 2 ms would lead to a dose difference of 1 %. This indicates that the choice of the threshold τ_{OC} within the LOT measurement algorithm, as discussed in section 3.2, is appropriate.

With correlations of 0.76 (TOMO1) and 0.65 (TOMO2), the mean relative LOT discrepancy was a good indicator of the dose deviation in the PTV (article C figure 6). This result differs from the previous findings of Deshpande *et al.* [71]. They reported a steeper relation between mean LOT error and the dose change in the PTV. One would expect the relation between fluence and LOT to be linear. Thus, the relation reported here (figure 3.4) seems to make more sense. Additionally, Deshpande *et al.* [71] used a different LOT measurement algorithm. They validated it by comparing the dose measured with ionisation chambers to the dose calculated based on the measured LOTs (Deshpande *et al.* [71] table 3). They reported higher differences between measured and calculated dose than what is reported in article C (article C figure 4), indicating that their LOT measurement algorithm has room for finer calibration.

Observing the relation between the *mean* LOT discrepancy and the dose deviation in the PTV came from the assumption that the LOT discrepancy distributions are symmetric. Figure 3.3 shows the dispersion of the relative LOT discrepancy as a function of the planned LOT on thorax treatment plan. The LOTs were measured during a static couch DQA procedure. Figure 3.3 also shows the planned LOTs and the LOT discrepancies histograms. The LOT discrepancy distribution was not symmetrical in this case. The distribution had a secondary peak near 7 %. Thus discrepancies mean and median were not equal. On this plan, the difference was low (0.5 %). It was higher (up to 2 %) in other of the clinical plans analysed in article C. As it turned out, the median (figure 3.4) was a better predictor than the mean (article C figure 6) of the dose deviation in the PTV. The correlation between median LOT discrepancy and dose deviation in the PTV was of 0.86 ($p \approx 10^{-23}$) on both plan sets.

Figure 3.4 show that the median relative LOT discrepancy is a good indicator of the dose delivery quality. Yet, a plan evaluation cannot be limited to observing a single value. Therefore, the spatial dose difference was also investigated in the plan that showed the greatest PTV dose deviation (figure 3.5). It appears that the dose difference was not homogeneous over the whole target volume. In particular, the difference is above 3 % close to

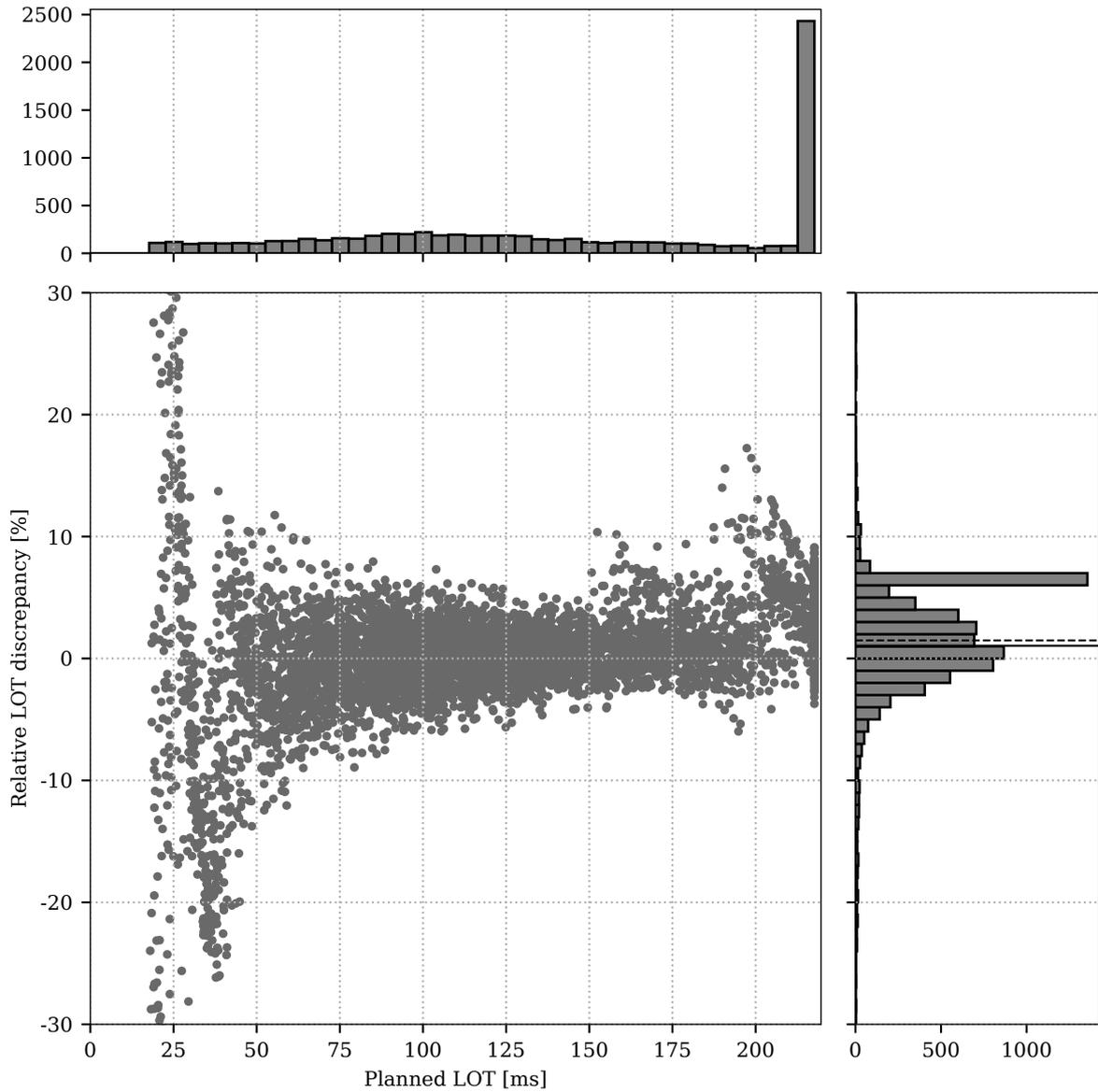


Figure 3.3 – Planned leaf open times (LOTs) and LOT discrepancy of thorax lesion treatment plan on TOMO2. The LOTs were measured during a static couch DQA procedure. The horizontal lines in the LOT discrepancies histogram represent the distribution mean (plain) and median (dashed).

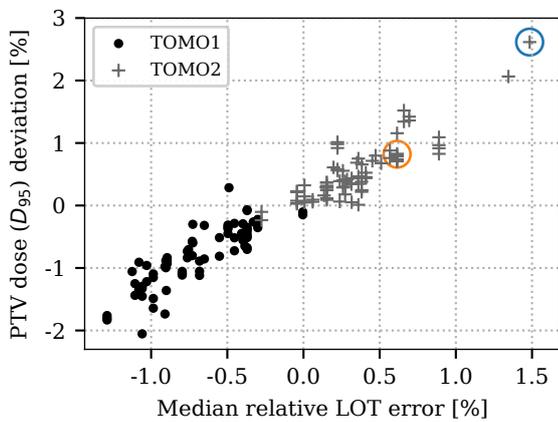


Figure 3.4 – Median relative leaf open time (LOT) discrepancy and deviation between planned and delivered dose to the PTV of the plans delivered on TOMO1 (black) and TOMO2 (grey). The circled points correspond to the thorax (blue) and the H&N (orange) plans represented in figures 3.3 and 3.7, respectively.

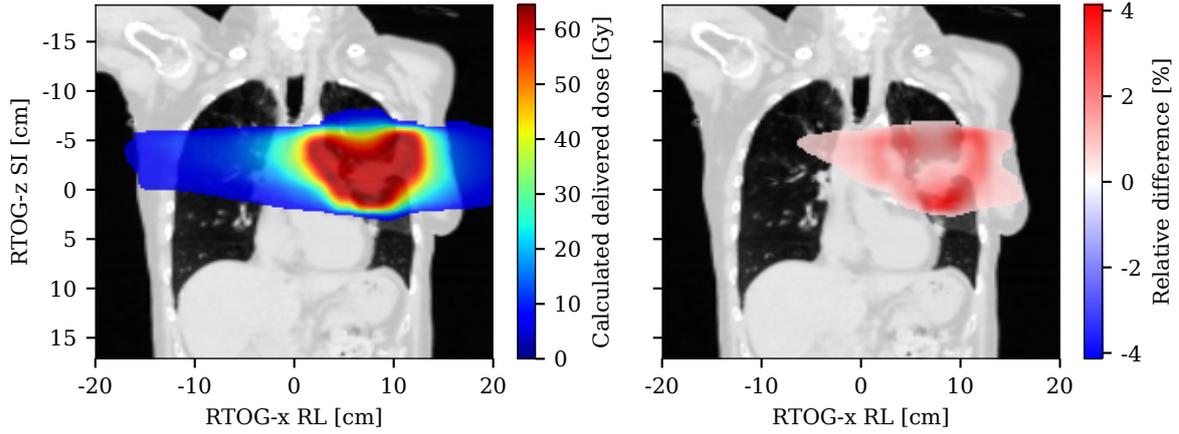


Figure 3.5 – Calculated delivered dose distribution and relative difference between planned and delivered dose of a thorax lesion treatment plan on TOMO2. The related leaf open time (LOT) discrepancy distribution is represented in figure 3.3. The median relative LOT discrepancy was of 1.5 % and the dose deviation in the pTV (D_{95}) was of 2.4 % (figure 3.4).

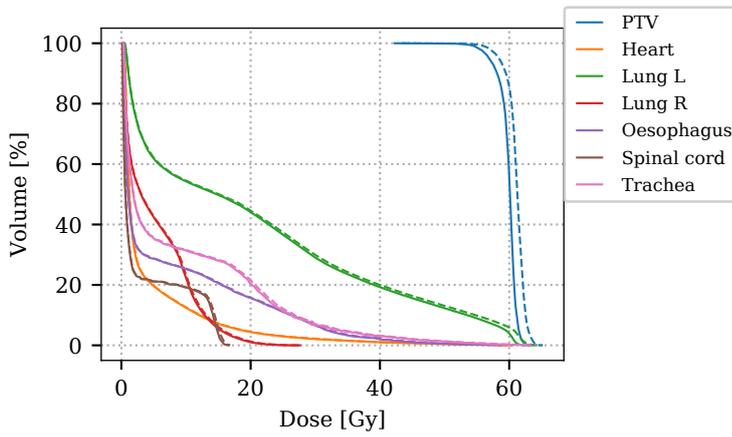


Figure 3.6 – Dose volume histogram (DVHs) of the planned (plain) and calculated delivered (dashed) dose distributions of a thorax lesion treatment plan on TOMO2. The dose distribution is represented on patient's CT images in figure 3.5.

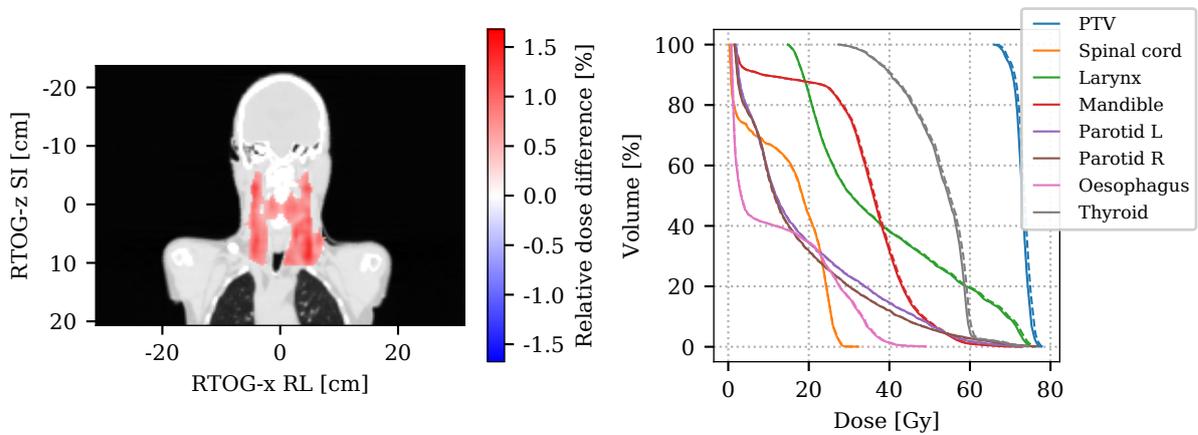


Figure 3.7 – Relative difference between planned and actual dose (left) and planned (plain) and actual (dashed) DVHs in the PTV and OARs (right).

the heart. Consequently, the dose deviation in the OARs was verified through DVH calculation and verification (figure 3.6). The maximum dose to the heart and to the left lung increased by 0.5 and 1.3 Gy, respectively. The median dose in the same OARs increased by 0.0 and 0.4 Gy, respectively. In an H&N plan, the dose difference distribution allowed us to see hotspots along the carotid arteries (figure 3.7). Delivered dose DVHs and 3D dose difference distribution will provide valuable information for the clinical evaluation of the treatment plans.

To fully validate the use the LOT measurement algorithm in conjunction with a standalone calculator, finer measurements should be performed. Here, the accuracy of the measured LOTs was assessed by comparing calculated and measured dose. The measurements were performed on the Cheese Phantom for synthetic TomoPhant plans. Though, the TomoPhant plans do not require complicated LOT modulation patterns. The TPS is required to cover PTVs with a single dose level (2 Gy) and is not required to protect any OAR. Comparing the calculated dose of clinical plans (as the dose distribution shown in figure 3.5) to measurements (*e.g.* with films in the Cheese Phantom) would be a finer validation.

Finally, the forward calculation based on the measured LOTs is not a comprehensive *in vivo* DQA tool. Firstly, even though the LOTs can be measured with the patient on the treatment couch, the patient anatomy at the day of the treatment is not taken into account. A first step would be to recalculate the dose in the patient's daily images. Though, the field of view (FOV) of the tomotherapy MVCT scanner is of 40 cm. In many situations, the patient is partly out of FOV. Secondly, the independence of the standalone calculator versus the TPS algorithm should be ensured. Alternatively, any robust independent dose calculation algorithm could be used.

Chapter 4

Conclusion and perspectives

In the context of this thesis, an independent dose calculation algorithm was upgraded to keep up with the TomoEDGE™ evolution of the tomotherapy treatment system. The results indicate that the upgrade was acceptable in the sense that the software had the same overall accuracy on static jaw plans (which did not require the upgrade) and on dynamic jaw plans (which did require the upgrade). Though, the overall software accuracy was low. This questioned the software's usefulness as QA tool. On the other hand, an on-line LOT measurement algorithm was developed. The algorithm proved to be robust against the beam attenuation, thus allowing the LOTs to be measured *in vivo*. The accuracy of the measured LOTs was assessed from dose measurements. The dose calculated with the measured LOTs as input agreed very well with ionisation chamber measurements. Finally, measuring the LOTs allowed dose deviations to be highlighted. The median (or the mean) relative LOT discrepancy showed a correlation with the dose deviation in the PTV. Thus, the median discrepancy could serve as an indicator of the dose delivery quality that has the advantage to be easy to measure.

Performing LOT-based DQA could help lighten the workload caused by the common IMRT DQA protocols. As already mentioned, these protocols consist in measuring a dose fraction in a phantom for each patient. A LOT-based method would only require the fluence to be delivered in air (with or without treatment couch in the gantry bore). Moreover, the subsequent dose calculation in the patient's planning CT images would provide more valuable information than a dose difference measured in a phantom and evaluated with the γ -index metric. The calculation of the delivered dose in the patient's images allow the dose deviations to be evaluated directly over the patient's anatomy, thus revealing immediately the severity (or insignificance) of hot and cold spots. Finally, the delivered dose can be evaluated through DVH calculation, making the severity of the overall difference between planned and delivered dose directly assessable.

As the LOT measurement algorithm works *in vivo*, the LOT-based DQA could be automatically performed at each treatment fraction. This would be a clear improvement compared to the current practice. Nowadays, no DQA of each individual dose fraction is performed. The patient-specific QA merely stops at the pretreatment step. From that point, the machine reliability is fully assumed. For the sake of completeness, and other perspectives detailed hereafter, the opportunity to cheaply perform a DQA of each dose fraction is highly desirable. The LOT

measurement algorithm presented here offers this possibility.

The LOT measurement algorithm could also make the patient flow more flexible by dropping pretreatment DQA measurements. Of course, an in-depth risk analysis must precede such a decision. Some insights were provided here. Firstly, the tomotherapy system has all the necessary interlocks to prevent severe accidents. Secondly, the LOTs are the only parameters adjusted by the TPS on which the tolerance is high enough to cause clinically relevant dose discrepancies. Other authors have reported it. In other words, accidents on a tomotherapy system originating from hardware defaults are highly improbable thanks to the builtin interlocks, but PTV underdosage has a non-negligible probability of occurrence. Though, it has been shown that large LOT errors originate from requesting an over-tight pitch during planification. This fact sustains the recommendations of Huq *et al.* [23] to place the QA checks in a broader framework and to insist on the staff training. The staff should be aware of the possible “traps” of the system.

If the patient-specific QA is placed in the correct framework and if there is a possibility to perform *in vivo* DQA of each dose fraction, then it seems reasonable to drop pretreatment DQA measurements for those planning parameters known to not cause dose discrepancies. The dose discrepancies measured on the data collected for article C were in an acceptable $\pm 2\%$ range (figure 3.4). This indicates that a correctly trained and experienced staff produces clinically acceptable plans. Let us finally remark that, would an error happen, the *in vivo* DQA would highlight it. As it is possible to calculate the delivered dose from the measured LOTs, it would be possible to quantify the dose discrepancy, and it would thus be possible to adapt the following dose fractions to correct for the original error. This, of course, is no primary option. It makes more sense to let no error occur in the first place, because replanning would be more costly and more inaccurate than performing a pretreatment DQA measurement. In addition, some indicators could help the staff decide whether or not they should perform a pretreatment DQA. In particular, the mean planned LOT correlates with the pitch. The tighter the pitch, the lower the mean planned LOT. Thus, the mean planned LOT (inversely) relates to the probability of delivering an inaccurate dose.

Combining the LOT measurement algorithm with an independent dose calculation algorithm would gather together the TPS QA and the dose DQA. To constitute a fully comprehensive *in vivo* dosimetry method, the combined LOT measurement and independent dose calculation approach should incorporate the patient’s daily images. The delivered dose would be more accurately calculated on the patient’s daily images than on the planning images. This should be made possible in a near future. Accuray recently added a fan-beam CT scanner to its tomotherapy system.

Then, it would presumably be possible to observe the drift of the dose error as the patient’s anatomy changes (weight loss, tumour size reduction, *etc.*). The delivered dose could be summed over the treatment fractions, thus allowing the clinicians to calculate more accurate adaptive treatment plans.

Another extension of an *in vivo* DQA method would be to check the patient position during the treatment, *e.g.* by reconstructing images during the delivery from the treatment beam or by identifying the patient’s contour. It would go even further by merging a patient position verification with a tumour tracking feature and a respiration correlated dose accumulation algorithm.

Finally, the LOT measurement algorithm could also be used for machine QA purpose. The medium term (over several months) drift of the LOT inaccuracies could be monitored and related to MLC breakdowns. As for now, TQA monitors the maximum daily LOT inaccuracy. Accuray recommends preventive servicing when the maximum error is greater than 20 ms. This does not fully prevent MLC breakdowns, so either the tolerance is too low or the monitored value is inadequate. The per leaf drift of the LOT inaccuracy could be measured in real usage conditions thanks to the ability of the algorithm to measure the LOTs *in vivo*. In contrast, TQA measures the LOTs in very synthetic conditions during the daily checks of the units. The inaccuracy drift could be an indicator of the MLC degradation. A predictor of MLC breakdowns would be of high interest. It could help reducing the machine down time by preventively servicing or replacing the MLC.

The work done and published in the context of this thesis is a valuable basis for the development of handy and more accurate DQA and machine QA tools for tomotherapy. It offers the opportunity to automatise patient-specific DQA, thus to reduce the considerable, human resources put into systematic performing pretreatment DQA in tomotherapy. Within a comprehensive QA framework and in parallel with exhaustive knowledge of the treatment system, it offers the opportunity to make the DQA practice in tomotherapy more flexible, namely by dropping pretreatment measurements in most situations. Finally, it offers the opportunity to monitor the delivery of each dose fraction individually, without demanding more efforts from the staff. This is an obvious improvement with respect to the current practice.

Bibliography

- [1] Charles M. Washington and Dennis T. Leaver. *Principles and Practice of Radiation Therapy*. Mosby, 2015.
- [2] Suzanne B. Evans *et al.* ‘Standardizing dose prescriptions: An ASTRO white paper’. In: *Practical Radiation Oncology* 6.6 (Nov. 2016), e369–e381. DOI: 10.1016/j.prrro.2016.08.007.
- [3] G.G. Hanna *et al.* ‘UK Consensus on Normal Tissue Dose Constraints for Stereotactic Radiotherapy’. In: *Clinical Oncology* 30.1 (Jan. 2018), pp. 5–14. DOI: 10.1016/j.clon.2017.09.007.
- [4] Stephanie Lim-Reinders, Brian M. Keller, Shahad Al-Ward, Arjun Sahgal and Anthony Kim. ‘Online Adaptive Radiation Therapy’. In: *International Journal of Radiation Oncology*Biography*Physics* 99.4 (Nov. 2017), pp. 994–1003. DOI: 10.1016/j.ijrobp.2017.04.023.
- [5] Cai Grau *et al.* ‘Radiotherapy equipment and departments in the European countries: Final results from the ESTRO-HERO survey’. In: *Radiotherapy and Oncology* 112.2 (Aug. 2014), pp. 155–164. DOI: 10.1016/j.radonc.2014.08.029.
- [6] T. Rock Mackie *et al.* ‘Tomotherapy: A new concept for the delivery of dynamic conformal radiotherapy’. In: *Medical Physics* 20.6 (Nov. 1993), pp. 1709–1719. DOI: 10.1118/1.596958.
- [7] Harumitsu Hashimoto *et al.* ‘Tangent field technique of TomoDirect improves dose distribution for whole-breast irradiation’. In: *Journal of Applied Clinical Medical Physics* 16.3 (May 2015), pp. 225–232. DOI: 10.1120/jacmp.v16i3.5369.
- [8] T R Mackie. ‘History of tomotherapy’. In: *Physics in Medicine and Biology* 51.13 (June 2006), R427–R453. DOI: 10.1088/0031-9155/51/13/r24.
- [9] John Balog, Gustavo Olivera and Jeff Kapatoes. ‘Clinical helical tomotherapy commissioning dosimetry’. In: *Medical Physics* 30.12 (Nov. 2003), pp. 3097–3106. DOI: 10.1118/1.1625444.
- [10] J D Fenwick *et al.* ‘Quality assurance of a helical tomotherapy machine’. In: *Physics in Medicine and Biology* 49.13 (June 2004), pp. 2933–2953. DOI: 10.1088/0031-9155/49/13/012.
- [11] David Sevillano, Cristina Mínguez, Alicia Sánchez and Alberto Sánchez-Reyes. ‘Measurement and correction of leaf open times in helical tomotherapy’. In: *Medical Physics* 39.11 (Oct. 2012), pp. 6972–6980. DOI: 10.1118/1.4762565.

- [12] Steffen Lissner, Kai Schubert, Sebastian Klüter, Dieter Oetzel and Jürgen Debus. ‘A method for testing the performance and the accuracy of the binary MLC used in helical tomotherapy’. In: *Zeitschrift für Medizinische Physik* 23.2 (May 2013), pp. 153–161. DOI: 10.1016/j.zemedi.2012.08.001.
- [13] Florian Sterzing *et al.* ‘Dynamic Jaws and Dynamic Couch in Helical Tomotherapy’. In: *International Journal of Radiation Oncology*Biography*Physics* 76.4 (Mar. 2010), pp. 1266–1273. DOI: 10.1016/j.ijrobp.2009.07.1686.
- [14] Yu Chen, Quan Chen, Mingli Chen and Weiguo Lu. ‘Dynamic tomotherapy delivery’. In: *Medical Physics* 38.6Part1 (May 2011), pp. 3013–3024. DOI: 10.1118/1.3584198.
- [15] Yi Rong, Yu Chen, Lu Shang, Li Zuo, Weiguo Lu and Quan Chen. ‘Helical tomotherapy with dynamic running-start-stop delivery compared to conventional tomotherapy delivery’. In: *Medical Physics* 41.5 (Apr. 2014), p. 051709. DOI: 10.1118/1.4870987.
- [16] Eric Schnarr *et al.* ‘Feasibility of real-time motion management with helical tomotherapy’. In: *Medical Physics* 45.4 (Feb. 2018), pp. 1329–1337. DOI: 10.1002/mp.12791.
- [17] M. W. Kissick *et al.* ‘The helical tomotherapy thread effect’. In: *Medical Physics* 32.5 (Apr. 2005), pp. 1414–1423. DOI: 10.1118/1.1896453.
- [18] Markus Alber *et al.* *Guidelines for the verification of IMRT*. ESTRO. 2008.
- [19] Alan C. Hartford *et al.* ‘American College of Radiology (ACR) and American Society for Radiation Oncology (ASTRO) Practice Guideline for Intensity-modulated Radiation Therapy (IMRT)’. In: *American Journal of Clinical Oncology* 35.6 (Dec. 2012), pp. 612–617. DOI: 10.1097/coc.0b013e31826e0515.
- [20] E. C. Ford, L. Fong de Los Santos, T. Pawlicki, S. Sutlief and P. Dunscombe. ‘Consensus recommendations for incident learning database structures in radiation oncology’. In: *Medical Physics* 39.12 (Nov. 2012), pp. 7272–7290. DOI: 10.1118/1.4764914.
- [21] Anthony Arnold. ‘Building a learning culture and prevention of error - to near miss or not’. In: *Journal of Medical Radiation Sciences* 64.3 (Sept. 2017), pp. 163–164. DOI: 10.1002/jmrs.242.
- [22] Eric C. Ford, Stephanie Terezakis, Annette Souranis, Kendra Harris, Hiram Gay and Sasa Mutic. ‘Quality Control Quantification (QCQ): A Tool to Measure the Value of Quality Control Checks in Radiation Oncology’. In: *International Journal of Radiation Oncology*Biography*Physics* 84.3 (Nov. 2012), e263–e269. DOI: 10.1016/j.ijrobp.2012.04.036.
- [23] M. Saiful Huq *et al.* ‘The report of Task Group 100 of the AAPM: Application of risk analysis methods to radiation therapy quality management’. In: *Medical Physics* 43.7 (June 2016), pp. 4209–4262. DOI: 10.1118/1.4947547.
- [24] Eric C. Ford and Stephanie Terezakis. ‘How Safe Is Safe? Risk in Radiotherapy’. In: *International Journal of Radiation Oncology*Biography*Physics* 78.2 (Oct. 2010), pp. 321–322. DOI: 10.1016/j.ijrobp.2010.04.047.

- [25] Serge Marcié, Elisabeth Charpiot, René-Jean Bensadoun, Gaston Ciais, Joel Héroult, André Costa and Jean-Pierre Gérard. ‘In vivo measurements with MOSFET detectors in oropharynx and nasopharynx intensity-modulated radiation therapy’. In: *International Journal of Radiation Oncology*Biophysics* 61.5 (Apr. 2005), pp. 1603–1606. DOI: 10.1016/j.ijrobp.2004.12.034.
- [26] Per E. Engström, Pia Haraldsson, Torsten Landberg, Hanne Sand Hansen, Svend Aage Engelholm and Håkan Nyström. ‘In vivo dose verification of IMRT treated head and neck cancer patients’. In: *Acta Oncologica* 44.6 (Jan. 2005), pp. 572–578. DOI: 10.1080/02841860500218983.
- [27] F M Gagliardi, K J Roxby, P E Engström and J C Crosbie. ‘Intra-cavitary dosimetry for IMRT head and neck treatment using thermoluminescent dosimeters in a naso-oesophageal tube’. In: *Physics in Medicine and Biology* 54.12 (May 2009), pp. 3649–3657. DOI: 10.1088/0031-9155/54/12/003.
- [28] Raj Varadhan, John Miller, Brenden Garrity and Michael Weber. ‘In vivoprostata IMRT dosimetry with MOSFET detectors using brass buildup caps’. In: *Journal of Applied Clinical Medical Physics* 7.4 (Sept. 2006), pp. 22–32. DOI: 10.1120/jacmp.v7i4.2278.
- [29] A J Vinall, A J Williams, V E Currie, A Van Esch and D Huyskens. ‘Practical guidelines for routine intensity-modulated radiotherapy verification: pre-treatment verification with portal dosimetry and treatment verification within vivodosimetry’. In: *The British Journal of Radiology* 83.995 (Nov. 2010), pp. 949–957. DOI: 10.1259/bjr/31573847.
- [30] Nils Kadesjö, Tufve Nyholm and Jörgen Olofsson. ‘A practical approach to diode based in vivo dosimetry for intensity modulated radiotherapy’. In: *Radiotherapy and Oncology* 98.3 (Mar. 2011), pp. 378–381. DOI: 10.1016/j.radonc.2010.12.018.
- [31] Sebastiaan M.J.J.G. Nijsten, Ben J. Mijnheer, André L.A.J. Dekker, Philippe Lambin and André W.H. Minken. ‘Routine individualised patient dosimetry using electronic portal imaging devices’. In: *Radiotherapy and Oncology* 83.1 (Apr. 2007), pp. 65–75. DOI: 10.1016/j.radonc.2007.03.003.
- [32] Mathilda van Zijtveld, Maarten Dirkx, Marcel Breuers, Hans de Boer and Ben Heijmen. ‘Portal dose image prediction for in vivo treatment verification completely based on EPID measurements’. In: *Medical Physics* 36.3 (Feb. 2009), pp. 946–952. DOI: 10.1118/1.3070545.
- [33] Markus Wendling, Leah N. McDermott, Anton Mans, Jan-Jakob Sonke, Marcel van Herk and Ben J. Mijnheer. ‘A simple backprojection algorithm for 3D in vivo EPID dosimetry of IMRT treatments’. In: *Medical Physics* 36.7 (June 2009), pp. 3310–3321. DOI: 10.1118/1.3148482.
- [34] A. Mans *et al.* ‘Catching errors with in vivo EPID dosimetry’. In: *Medical Physics* 37.6Part2 (May 2010), pp. 2638–2644. DOI: 10.1118/1.3397807.
- [35] Niall D MacDougall, Michael Graveling, Vibeke N Hansen, Kevin Brownsword and Andrew Morgan. ‘In vivo dosimetry in UK external beam radiotherapy: current and future usage’. In: *The British Journal of Radiology* 90.1072 (Apr. 2017), p. 20160915. DOI: 10.1259/bjr.20160915.

- [36] Christopher Neilson, Michael Klein, Rob Barnett and Slav Yartsev. ‘Delivery quality assurance with Arc-CHECK’. In: *Medical Dosimetry* 38.1 (Mar. 2013), pp. 77–80. DOI: 10.1016/j.meddos.2012.07.004.
- [37] Syam Kumar, Aswathi Cheruparambil and Anjana Parakkat Thokkayil. ‘Clinically evaluating directional dependence of 2D seven²⁹ ion-chamber array with different IMRT plans’. In: *International Journal of Cancer Therapy and Oncology* 3.4 (Dec. 2015), p. 348. DOI: 10.14319/ijcto.34.8.
- [38] T. Landberg, J. Chavaudra, J. Dobbs, G. Hanks, K.-A. Johansson, T. Möller and J. Purdy. ‘Report 50’. In: *Journal of the International Commission on Radiation Units and Measurements* os26.1 (Sept. 1993), NP–NP. DOI: 10.1093/jicru/os26.1.report50.
- [39] D. A. Low, W. B. Harms, S. Mutic and J. A. Purdy. ‘A technique for the quantitative evaluation of dose distributions’. In: *Medical Physics* 25.5 (May 1998), pp. 656–661.
- [40] Moyed Miften *et al.* ‘Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218’. In: *Medical Physics* 45.4 (Mar. 2018), e53–e83. DOI: 10.1002/mp.12810.
- [41] Heming Zhen, Benjamin E. Nelms and Wolfgang A. Tomé. ‘Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA’. In: *Medical Physics* 38.10 (Sept. 2011), pp. 5477–5489. DOI: 10.1118/1.3633904.
- [42] M. Stasi, S. Bresciani, A. Miranti, A. Maggio, V. Sapino and P. Gabriele. ‘Pretreatment patient-specific IMRT quality assurance: A correlation study between gamma index and patient clinical dose volume histogram’. In: *Medical Physics* 39.12 (Nov. 2012), pp. 7626–7634. DOI: 10.1118/1.4767763.
- [43] Simon J. Thomas and Ian R. Cowley. ‘A Comparison of Four Indices for Combining Distance and Dose Differences’. In: *International Journal of Radiation Oncology*Biophysics* 82.5 (Apr. 2012), e717–e723. DOI: 10.1016/j.ijrobp.2011.11.009.
- [44] Sara Bresciani, Amalia Di Dia, Angelo Maggio, Claudia Cutaia, Anna Miranti, Erminia Infusino and Michele Stasi. ‘Tomotherapy treatment plan quality assurance: The impact of applied criteria on passing rate in gamma index method’. In: *Medical Physics* 40.12 (Nov. 2013), p. 121711. DOI: 10.1118/1.4829515.
- [45] Benjamin E. Nelms, Maria F. Chan, Geneviève Jarry, Matthieu Lemire, John Lowden, Carnell Hampton and Vladimir Feygelman. ‘Evaluating IMRT and VMAT dose accuracy: Practical examples of failure to detect systematic errors when applying a commonly used metric and action levels’. In: *Medical Physics* 40.11 (Oct. 2013), p. 111722. DOI: 10.1118/1.4826166.
- [46] Stephen F. Kry *et al.* ‘Institutional Patient-specific IMRT QA Does Not Predict Unacceptable Plan Delivery’. In: *International Journal of Radiation Oncology*Biophysics* 90.5 (Dec. 2014), pp. 1195–1201. DOI: 10.1016/j.ijrobp.2014.08.334.

- [47] Elizabeth M. McKenzie, Peter A. Balter, Francesco C. Stingo, Jimmy Jones, David S. Followill and Stephen F. Kry. ‘Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans’. In: *Medical Physics* 41.12 (Nov. 2014), p. 121702. DOI: 10.1118/1.4899177.
- [48] Strahinja Stojadinovic, Luo Ouyang, Xuejun Gu, Arnold Pompoš, Qinan Bao and Timothy D. Solberg. ‘Breaking bad IMRT QA practice’. In: *Journal of Applied Clinical Medical Physics* 16.3 (May 2015), pp. 154–165. DOI: 10.1120/jacmp.v16i3.5242.
- [49] M. Stasi *et al.* ‘EP-1587: Sensitivity and specificity of gamma index method for Tomotherapy plans’. In: *Radiotherapy and Oncology* 119 (Apr. 2016), S737. DOI: 10.1016/s0167-8140(16)32837-7.
- [50] Katja M. Langen *et al.* ‘QA for helical tomotherapy: Report of the AAPM Task Group 148’. In: *Medical Physics* 37.9 (Aug. 2010), pp. 4817–4853. DOI: 10.1118/1.3462971.
- [51] John Balog, Tim Holmes and Richard Vaden. ‘Helical tomotherapy dynamic quality assurance’. In: *Medical Physics* 33.10 (Sept. 2006), pp. 3939–3950. DOI: 10.1118/1.2351952.
- [52] Ann Van Esch, Christian Clermont, Magali Devillers, Mauro Iori and Dominique P. Huyskens. ‘On-line quality assurance of rotational radiotherapy treatment delivery by means of a 2D ion chamber array and the Octavius phantom’. In: *Medical Physics* 34.10 (Sept. 2007), pp. 3825–3837. DOI: 10.1118/1.2777006.
- [53] M. Geurts, J. Gonzalez and P. Serrano-Ojeda. ‘Longitudinal study using a diode phantom for helical tomotherapy IMRT QA’. In: *Medical Physics* 36.11 (Oct. 2009), pp. 4977–4983. DOI: 10.1118/1.3238153.
- [54] Paul A. Jursinic, Renu Sharma and Jim Reuter. ‘MapCHECK used for rotational IMRT measurements: Step-and-shoot, Tomotherapy, RapidArc’. In: *Medical Physics* 37.6Part1 (May 2010), pp. 2837–2846. DOI: 10.1118/1.3431994.
- [55] David Chapman, Rob Barnett and Slav Yartsev. ‘Helical tomotherapy quality assurance with ArcCHECK’. In: *Medical Dosimetry* 39.2 (2014), pp. 159–162. DOI: 10.1016/j.meddos.2013.12.002.
- [56] T. Kron, S. Yartsev and T. R. Mackie. ‘Verification dosimetry during treatment for helical tomotherapy using radiographic film’. In: *Australasian Physics & Engineering Sciences in Medicine* 28.4 (Dec. 2005), pp. 232–237. DOI: 10.1007/bf03178723.
- [57] Zhen-Yu Qi *et al.* ‘Real-Time In Vivo Dosimetry With MOSFET Detectors in Serial Tomotherapy for Head and Neck Cancer Patients’. In: *International Journal of Radiation Oncology*Biography*Physics* 80.5 (Aug. 2011), pp. 1581–1588. DOI: 10.1016/j.ijrobp.2010.10.063.
- [58] Sarah J. Alnaghy, Shrikant Deshpande, Dean L. Cutajar, Kemal Berk, Peter Metcalfe and Anatoly B. Rosenfeld. ‘In vivo endorectal dosimetry of prostate tomotherapy using dual MOSkindetectors’. In: *Journal of Applied Clinical Medical Physics* 16.3 (May 2015), pp. 107–117. DOI: 10.1120/jacmp.v16i3.5113.

- [59] T. R. McNutt, T. R. Mackie and B. R. Paliwal. ‘Analysis and convergence of the iterative convolution/superposition dose reconstruction technique for multiple treatment beams and tomotherapy’. In: *Medical Physics* 24.9 (Sept. 1997), pp. 1465–1476. DOI: 10.1118/1.598035.
- [60] J. M. Kapatoes, G. H. Olivera, K. J. Ruchala, J. B. Smilowitz, P. J. Reckwerdt and T. R. Mackie. ‘A feasible method for clinical delivery verification and dose reconstruction in tomotherapy’. In: *Medical Physics* 28.4 (Apr. 2001), pp. 528–542. DOI: 10.1118/1.1352579.
- [61] J M Kapatoes, G H Olivera, K J Ruchala and T R Mackie. ‘On the verification of the incident energy fluence in tomotherapy IMRT’. In: *Physics in Medicine and Biology* 46.11 (Oct. 2001), pp. 2953–2965. DOI: 10.1088/0031-9155/46/11/313.
- [62] J M Kapatoes, G H Olivera, J P Balog, H Keller, P J Reckwerdt and T R Mackie. ‘On the accuracy and effectiveness of dose reconstruction for tomotherapy’. In: *Physics in Medicine and Biology* 46.4 (Mar. 2001), pp. 943–966. DOI: 10.1088/0031-9155/46/4/303.
- [63] Ke Sheng *et al.* ‘3D Dose Verification Using Tomotherapy CT Detector Array’. In: *International Journal of Radiation Oncology*Biography*Physics* 82.2 (Feb. 2012), pp. 1013–1020. DOI: 10.1016/j.ijrobp.2010.12.043.
- [64] E Mezzenga, E Cagni, A Botti, M Orlandi and M Iori. ‘Pre-treatment and in-vivo dosimetry of Helical Tomotherapy treatment plans using the Dosimetry Check system’. In: *Journal of Instrumentation* 9.04 (Apr. 2014), pp. C04039–C04039. DOI: 10.1088/1748-0221/9/04/c04039.
- [65] Eunah Chung, Dongyeol Kwon, Taeyang Park, Hyeri Kang and Yoonsun Chung. ‘Clinical implementation of Dosimetry Check™ for TomoTherapy® delivery quality assurance’. In: *Journal of Applied Clinical Medical Physics* 19.6 (Oct. 2018), pp. 193–199. DOI: 10.1002/acm2.12480.
- [66] Zoë R. Moutrie, Craig M. Lancaster and Litang Yu. ‘First experiences in using a dose control system on a TomoTherapy Hi-Art II’. In: *Journal of Applied Clinical Medical Physics* 16.3 (May 2015), pp. 277–284. DOI: 10.1120/jacmp.v16i3.5489.
- [67] Lydia L. Handsfield, Ryan Jones, David D. Wilson, Jeffery V. Siebers, Paul W. Read and Quan Chen. ‘Phantomless patient-specific TomoTherapy QA via delivery performance monitoring and a secondary Monte Carlo dose calculation’. In: *Medical Physics* 41.10 (Sept. 2014), p. 101703. DOI: 10.1118/1.4894721.
- [68] Francis Kar-Ho Lee, Simon Kar-Yiu Chan and Ricky Ming-Chun Chau. ‘Dosimetric verification and quality assurance of running-start-stop (RSS) delivery in tomotherapy’. In: *Journal of Applied Clinical Medical Physics* 16.6 (Nov. 2015), pp. 23–29. DOI: 10.1120/jacmp.v16i6.5336.
- [69] David C. Westerly, Emilie Soisson, Quan Chen, Katherine Woch, Leah Schubert, Gustavo Olivera and Thomas R. Mackie. ‘Treatment Planning to Improve Delivery Accuracy and Patient Throughput in Helical Tomotherapy’. In: *International Journal of Radiation Oncology*Biography*Physics* 74.4 (July 2009), pp. 1290–1297. DOI: 10.1016/j.ijrobp.2009.02.004.

- [70] Shrikant Deshpande, Mark Geurts, Philip Vial, Peter Metcalfe, Mark Lee and Lois Holloway. ‘Clinical significance of treatment delivery errors for helical TomoTherapy nasopharyngeal plans – A dosimetric simulation study’. In: *Physica Medica* 33 (Jan. 2017), pp. 159–169. DOI: 10.1016/j.ejmp.2017.01.006.
- [71] Shrikant Deshpande, Aitang Xing, Peter Metcalfe, Lois Holloway, Philip Vial and Mark Geurts. ‘Clinical implementation of an exit detector-based dose reconstruction tool for helical tomotherapy delivery quality assurance’. In: *Medical Physics* 44.10 (2017), pp. 5457–5466. DOI: 10.1002/mp.12484.
- [72] Quan Chen, David Westerly, Zhenyu Fang, Ke Sheng and Yu Chen. ‘TomoTherapy MLC verification using exit detector data’. In: *Medical Physics* 39.1 (Dec. 2011), pp. 143–152. DOI: 10.1118/1.3666762.
- [73] Simon J. Thomas, Katie R. Eyre, G. Samuel J. Tudor and Jamie Fairfoul. ‘Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube’. In: *Medical Physics* 39.1 (Dec. 2011), pp. 160–167. DOI: 10.1118/1.3668061.
- [74] John P. Gibbons, Koren Smith, Dennis Cheek and Isaac Rosen. ‘Independent calculation of dose from a helical TomoTherapy unit’. In: *Journal of Applied Clinical Medical Physics* 10.1 (Feb. 2009), pp. 103–119. DOI: 10.1120/jacmp.v10i1.2772.
- [75] Quan Chen, Weiguo Lu, Yu Chen, Mingli Chen, Douglas Henderson and Edmond Sterpin. ‘Validation of GPU based TomoTherapy dose calculation engine’. In: *Medical Physics* 39.4 (Mar. 2012), pp. 1877–1886. DOI: 10.1118/1.3693057.
- [76] Ramon Alfredo C. Siochi, Andrea Molineu and Colin G. Orton. ‘Patient-specific QA for IMRT should be performed using software rather than hardware methods’. In: *Medical Physics* 40.7 (May 2013), p. 070601. DOI: 10.1118/1.4794929.
- [77] Nathan Childress, Quan Chen and Yi Rong. ‘Parallel/Opposed: IMRT QA using treatment log files is superior to conventional measurement-based method’. In: *Journal of Applied Clinical Medical Physics* 16.1 (Jan. 2015), pp. 4–7. DOI: 10.1120/jacmp.v16i1.5385.
- [78] Wolfgang A. Tomé and Jack F. Fowler. ‘On cold spots in tumor subvolumes’. In: *Medical Physics* 29.7 (June 2002), pp. 1590–1598. DOI: 10.1118/1.1485060.
- [79] Mikael Karlsson, Anders Ahnesjö, Dietmar Georg, Tufve Nyholm and Jörgen Olofsson. *Independent dose calculations concepts and models*. ESTRO. 2010.

Appendix A

Implementation of TomoEDGE in the independent dose calculator CheckTomo

Mathieu Schopfer, Simon J. Thomas, G. Samuel J. Tudor, Jean Bourhis, François Bochud, and Raphaël Moeckli

In: *Journal of Applied Clinical Medical Physics* 18.2 (March 2017), pp. 92–99

DOI: 10.1002/acm2.12048

Implementation of TomoEDGE in the independent dose calculator CheckTomo

Mathieu Schopfer¹ | Simon J. Thomas² | G. Samuel J. Tudor² | Jean Bourhis³ |
François Bochud¹ | Raphaël Moeckli¹

¹Institute of Radiation Physics, CHUV and University of Lausanne, CH-1007 Lausanne, Switzerland

²Department of Medical Physics, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK

³Department of Radio-Oncology, CHUV and University of Lausanne, CH-1007 Lausanne, Switzerland

Author to whom correspondence should be addressed. Raphaël Moeckli
E-mail: Raphael.Moeckli@chuv.ch
Telephone: (+41) 213144618;
Fax: (+41) 213148299.

Abstract

Purpose: CheckTomo is an independent dose calculation software for tomotherapy. Recently, Accuray (Accuray Inc., Sunnyvale, CA, USA) released an upgrade of its tomotherapy treatment device, called TomoEDGE Dynamic Jaws, which improves the quality of treatment plans by enhancing the dose delivery with the help of jaws motion. This study describes the upgrade of CheckTomo to that new feature.

Methods: To account for the varying width and off-axis shift of dynamic jaws fields, the calculation engine of CheckTomo multiplies the treatment field profile by a penumbral filter and shifts the dose calculation grid. Penumbral filters were obtained by dividing the edge field profiles by that of the corresponding nominal field. They were sampled at widths 1.0, 1.8, and 2.5 cm at isocenter in the edges of the 2.5 and 5 cm treatment field.

Results: The upgrade of CheckTomo was tested on 30 patient treatments planned with dynamic jaws. The gamma pass rate averaged over 10 abdomen plans was 95.9%, with tolerances of 3 mm/3%. For 10 head and neck plans, the mean pass rate was 95.9% for tolerances of 4 mm/4%. Finally, misplacement and overdosage errors were simulated. In each tested cases, the 2 mm/3% gamma pass rate fell below 95% when a 4 mm shift or 3% dose difference was applied.

Conclusions: These results are equivalent to what CheckTomo achieves in static jaws cases. So, in terms of dose calculation accuracy and errors detection, the upgraded version of CheckTomo is as reliable for dynamic jaws plans as the former release was for static cases.

PACS

87.55.kd, 87.55.Qr

KEY WORDS

independent, quality assurance, tomotherapy

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, Inc. on behalf of American Association of Physicists in Medicine.

1 | INTRODUCTION

Independent dose verification is considered to be important to ensure patient safety.¹ It can be performed through an independent calculation with commercial softwares for three-dimensional conformal radiation therapy (3DCRT), image-modulated radiation therapy (IMRT), and volume-modulated arc therapy (VMAT) treatments. For tomotherapy, as far as we know, there exists a commercial tool, Mobius 3D (Mobius Medical Systems, Houston, TX, USA), and a single-point dose verification software.² Additionally, an open source solution, CheckTomo, was released in 2011.³ That software independently generates a three-dimensional point-based dose distribution, using patient CT images and delivery plan, and compares it against the dose volume calculated by the tomotherapy treatment planning system (TPS).

Accuray (Accuray Inc., Sunnyvale, CA, USA) released an upgrade of its tomotherapy device called TomoEDGE Dynamic Jaws.⁴ The purpose of this upgrade is to reduce the field penumbra along the patient longitudinal (inferior–superior) axis by the mean of jaws motion. The way the dose is delivered is hence modified and the dose calculation engine of CheckTomo needed to be upgraded consequently.

This study aims to present the work done to develop and implement the upgrade of CheckTomo and the tests that were performed to assess that the dose calculation carried out with the upgrade is as reliable as it was with the previous version. It does not suggest any improvement of the core calculation engine.

2 | MATERIALS AND METHODS

2.A | TomoEDGE dynamic jaws

In tomotherapy, the field is delimited in the longitudinal (IEC-y) direction by a pair of collimators, called jaws. A non-TomoEDGE direct or helical tomotherapy treatment is delivered with static jaws, i.e., at fixed field width during the whole treatment procedure, either 1, 2.5, or 5 cm at isocenter. This implies that the field penumbra in the longitudinal direction is of approximately the field size on both cranial and caudal sides of the target. To limit the extra dose to organs at risk (OAR) and other healthy tissues, the treatment can be

delivered with a smaller field width, but this usually increases the irradiation time.

To overcome this poor trade-off, TomoEDGE introduced jaws motion during treatment delivery.⁵ At treatment start, the jaws delimit at isocenter an asymmetrical 1 cm wide field, located off the source axis toward the patient's feet. Then as the couch moves forward, the cranial jaw sweeps toward the patient's head to keep the field edge 5 mm ahead of the planning target volume (PTV), until the jaws delimit a symmetrical field (respectively to the beam axis) of the nominal treatment size, either 2.5 or 5 cm at isocenter. Similarly, the caudal jaw closes behind the PTV as it exits the beam, until the field is 1 cm wide again.⁴ In a TomoEDGE treatment, the penumbra on the cranial and caudal sides of the PTV is reduced to 1 cm. See Fig. 1 for a graphical depiction.

For clarity, the fields will be denominated “nominal” when delimited by symmetrically positioned static jaws and “edge” otherwise.

2.B. | CheckTomo

2.B.1 | Software basics

CheckTomo is a software written in MATLAB (The MathWorks Inc., Natick, MA, USA) that computes a three-dimensional point-based dose distribution using CT data and treatment plan on the patient side and independently acquired beam data on the machine side.

Patient data are read from DICOM CT and RT-plan files where beam geometry and patient position during treatment are described.

Beam data are provided with CheckTomo for each nominal treatment field in text files with a homemade structure. They consist of a reference dose point, tissue-phantom ratios (TPRs), output factors (OFs), and off-axis ratios (OARs) measured for various field shapes. The $5 \times 40 \text{ cm}^2$ field at isocenter was taken as the reference one and the dose reference point was measured isocentrically at depth 10 cm. All machine data were independently acquired on a tomotherapy unit using an ionization chamber at different depths in a water tank.

CheckTomo dose distribution is usually calculated on a grid of $15 \times 15 \times 15$ points, with a 1 to 1.5 cm spacing. Grid resolution and size can be adapted if needed. For each sinogram projection (or control point), the dose deposited at a particular location is the

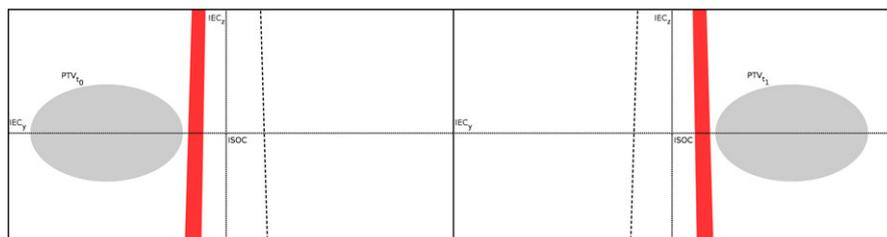


Fig. 1. Schematic representation of a TomoEDGE treatment beam at two moments. Dashed lines represent the nominal field width. Edge fields (in red) are represented at treatment start (right) and end (left). At treatment start, the jaws delimit a 1 cm wide field on the negative IEC-y side of the beam axis. During treatment (not represented), as the PTV moves forward, the cranial jaw opens to keep the superior field edge ahead of the PTV superior limit. Then, the caudal jaw closes to keep the inferior field edge behind the PTV inferior limit. Finally, when the treatment ends, the jaws delimit a 1 cm wide field again, but on the positive IEC-y side of the beam axis.

product of the projection time, the dose rate, TPR, OF, and OAR. The fluence is considered to arise from the mean angle of the projection arc, which, regarding the tomotherapy standard of defining 51 control points per gantry rotation, extends over 7.29°. To increase the number of control points and thus improve the dose calculation accuracy, CheckTomo offers the option to split each projection into multiple subprojections.⁶

CheckTomo dose distribution can be compared to that calculated by the tomotherapy treatment planning system (TPS) by means of a gamma⁷ or box comparison index.⁸ Required patient data, beam data collection, dose calculation model, and comparison indices were explained in more detail in the original release paper of CheckTomo.³

2.B.2 | Beam profile model

In CheckTomo, the longitudinal profile of a nominal field is calculated by multiplying the field OAR, the TPR, and the OF. CheckTomo handles OARs expressed in angular distance respectively to the beam source, instead of Cartesian coordinates. It follows the tomotherapy naming conventions of field size, calling the longitudinal dimension the width and the in-plane dimension the length (width and length are always given at isocenter). Which is more, the OF of the tomotherapy beam, hereafter S_{cp} , is not a function of the equivalent square field size but depends independently on both the field width and length.³ In CheckTomo, it is therefore considered to be a function S_{cp,w_0} of the field length specific to the nominal field of width w_0 .

Thus, the longitudinal profile at angular coordinate θ_y and depth d of a nominal field of width w_0 and length L is given by

$$P_N(w_0, L, \theta_y, d) = OAR_y(\theta_y, d) \cdot TPR(A_{sq}, d) \cdot S_{cp,w_0}(L). \quad (1)$$

A_{sq} is the equivalent square field size.

2.C. | Implementation of a dynamic jaws beam profile model in CheckTomo

Jaws motion induces changes in the field shape and OF that have to be accounted for in the profile model. Theoretically, the longitudinal profile of an edge field is obtained by multiplying Eq. (1) by a jaw penumbral filter and by correcting the OF. But as mentioned in section 2.B.2, the OF function S_{cp} was not designed to account for a varying field width. To overcome this limitation, the relative jaw penumbral filter (RJPF) was introduced, defined as the ratio of the edge and nominal longitudinal profiles P_E and P_N ,

$$RJPF(w, w_0, \theta_y, d) = \frac{P_E(w, w_0, L, \theta_y, d)}{P_N(w_0, L, \theta_y, d)}. \quad (2)$$

Here P_E is the edge field profile given in angular coordinates respectively to the beam source. The transformation consists in first applying a coordinates shift along the longitudinal axis so that the field maximum is at $IEC-y = 0$. Then, the shifted Cartesian coordinates are converted in angular distances.

The edge field profile equation is obtained by inverting relation (2) and replacing P_N with equation (1), namely

$$P_E(w, w_0, L, \theta_y, d) = OAR_y(\theta_y, d) \cdot TPR(A_{sq}, d) \cdot S_{cp,w_0}(L) \cdot RJPF(w, w_0, \theta_y, d). \quad (3)$$

Note that it yields a profile originating at the source axis. To account for the edge field off-axis nature, the dose calculation grid is shifted longitudinally — toward head or feet depending on the edge side — by half the field width.

In practice, P_E and P_N were sampled at field widths and depths specified in section 2.D, normalized, respectively, to P_N peak maxima and converted into angular coordinates. RJPFs were then calculated from Eq. (2) by interpolating P_E and P_N over for a set of arbitrary points. These data were stored in new text files structured like the existing CheckTomo beam data files. Note that the RJPFs were not sampled at different field lengths because it was checked that this parameter only has a slight influence on the longitudinal profiles. Lastly, for widths and depths falling between sampling values, the RJPF is interpolated on the fly at run time.

2.D. | Measurements of edge beam data

Profiles measurements were performed with an Exradin A1SL ionization chamber (Standard Imaging, Middleton, WI, USA) in a water tank at SSD 85 cm, all MLC leaves open and depths 1.5, 5, 10, 15, and 20 cm. They were all run successively for the nominal and edge fields.

Measurements of both the edge and nominal profiles were needed to calculate the RJPF from Eq. (2). The edge field width varying continuously between 1 cm and the nominal field size, it was necessary to pick some sampling values. During the TomoEDGE acceptance test procedure (ATP), field data were measured for widths 1.0, 1.8, and 2.5 cm in both edges of 5 cm nominal field. We decided to perform profile measurements for that same set of values. Due to the flattening filter free (FFF) beam of tomotherapy units, the profile of an edge field depends also on its distance to the source axis. So, similar measurements were performed in the edge of the 2.5 cm nominal field as well. Obviously, it was sufficient to realize them only on one side of the source axis.

2.E. | Dose calculation verification and tests of accuracy

2.E.1 | Gradient check

Five plans were generated using the images of the Cheese Phantom and the 5.0 cm plans structures set provided with the TomoPhant IMRT verification patient, which is usually available in the tomotherapy TPS. Three PTVs of 2 cm, 6 cm, and 10 cm were created by shrinking or extending the original target volume. Plans were calculated for the 2.5 cm field on these three PTVs and for the 5 cm field on the 6 cm and 10 cm PTVs. All plans were calculated in dynamic jaw mode. The PTVs were centered on the machine isocenter, the

prescription dose was of 2 Gy and the pitch was 0.287. To force some field modulation, a constraint was applied on a structure of the same size as the target located 2 cm beneath it.

All five plans were calculated in CheckTomo with a 2.5 mm longitudinal spacing and global 2 mm/3% and 3 mm/4% gamma indices were calculated. Additionally, the dose profiles along the longitudinal axis in the isocenter plane were extracted from both the CheckTomo and tomotherapy TPS dose volume so that they could be compared visually.

2.E.2 | Dose verification in real patient cases

The upgrade of CheckTomo was tested on 30 patient cases planned and treated with dynamic jaws. All plans had successfully passed a clinical quality assurance (QA) test which consisted in comparing the TPS dose distribution to a measurement performed with an Octavius 729 detector array in an Octavius II phantom (PTW, Freiburg, Germany). Dose comparison was done in VeriSoft (PTW, Freiburg, Germany) using a 3 mm/3% gamma comparison index⁷ for points within the 10% isodose and considering a 95% pass rate threshold.

The independent calculation of the dose distributions was performed with the upgraded version of CheckTomo using the original patients CT images, a $31 \times 31 \times 31$ calculation grid with a longitudinal spacing of 6 mm (8 mm in two cases) and one subprojection

per projection. These grid settings ensured us to cover in each case a major part of the PTV and to get a reasonably high dose point resolution in the field edges. Note that in some cases, the PTV was too large to fit entirely in the dose calculation grid. PTV length and field width for each patient are given in Tables 2–4.

The calculation accuracy was assessed for each of the 30 plans by computing the mean dose difference and performing global gamma comparison tests between the CheckTomo and the tomotherapy TPS dose distributions. The gamma index was calculated for various tolerances over the points located within the 50% isodose and at least 5 mm deep in the patient's body. A test is considered successful if its gamma pass rate is above 95%.

2.E.3 | Errors simulation

Finally, in order to test the ability of the upgrade of CheckTomo to detect errors, 15 cases that had passed a 2 mm/3% gamma comparison test were selected, independently of the treatment location. Then longitudinal misplacements and overdosages were simulated over them by applying a 2 mm and 4 mm coordinate shift and a 3% dose offset to the TPS dose distribution.

3 | RESULTS AND DISCUSSION

3.1 | Edge beam data and profiles model

Figure 2 shows field profiles measured and calculated on the positive IEC-y side of the gantry. All profiles were normalized to the maximum of the corresponding nominal field. The difference in relative intensity between a profile in the edge of the 2.5 cm and 5 cm nominal field is visible, particularly for the 1 cm field.

Note that Eq. (3) yields a symmetric approximation of the edge field profiles, which are actually asymmetric (because the position of the jaws compared to the axis of the beam generates asymmetric penumbra). This approximation is inherent to the beam model of CheckTomo, which was not designed to handle asymmetric fields. Though, as can be seen in Fig. 2, the calculated edge profiles (plain

TABLE 1 Gamma pass rate (γ) for two tolerances and average mean dose difference ($\overline{\Delta D}$) of the five plans calculated in the TomoPhant. Points within the 50% isodose and at least 5 mm depth were considered in the calculation of the gamma index.

Field width [cm]	PTV length [cm]	γ 2 mm, 3% [%]	γ 3 mm, 4% [%]	$\overline{\Delta D}$ [%]
2.5	2	91.2	96.4	2.2
2.5	6	100.0	100.0	0.0
2.5	10	99.8	100.0	-1.0
5.0	6	88.0	96.0	-0.1
5.0	10	86.7	95.3	-0.9

TABLE 2 Geometrical setup, gamma pass rate (γ) for various tolerances and average mean dose difference ($\overline{\Delta D}$) of the 10 abdomen plans. Points within the 50% isodose and at least 5 mm depth were considered in the calculation of the gamma index.

Case	PTV dose [Gy]	Field width [cm]	PTV length [cm]	γ 2 mm, 3% [%]	γ 3 mm, 3% [%]	γ 4 mm, 4% [%]	$\overline{\Delta D}$ [%]
A01	2	5.054	30.2	97.8	99.2	100.0	-0.9
A02	1.8	2.51	29.2	83.7	88.6	96.4	-1.8
A03	2	2.51	9.75	100.0	100.0	100.0	0.27
A04	1.8	2.51	17.6	97.0	98.2	99.2	-0.5
A05	7	2.51	3.8	97.5	99.1	100.0	1.93
A06	1.8	5.054	20.4	84.2	89.6	94.5	1.38
A07	2	2.51	13	99.5	99.8	100.0	-0.2
A08	2	2.51	12.2	83.6	85.8	93.0	2.56
A09	3	2.51	19.2	97.9	98.6	99.7	-0.8
A10	2.3	2.51	33.4	99.6	99.9	100.0	-0.7

TABLE 3 Geometrical setup, gamma pass rate (γ) for various tolerances and average mean dose difference ($\overline{\Delta D}$) of the 10 head and neck plans. Points within the 50% isodose and at least 5 mm depth were considered in the calculation of the gamma index.

Case	PTV dose [Gy]	Field width [cm]	PTV length [cm]	γ 2 mm, 3% [%]	γ 3 mm, 3% [%]	γ 4 mm, 4% [%]	$\overline{\Delta D}$ [%]
HN01	2	2.51	12.75	69.2	74.6	88.9	3.7
HN02	2.12	2.51	7.75	65.6	76.6	89.6	3.1
HN03	2.12	2.51	14.75	100.0	100.0	100.0	0.3
HN04	2.12	2.51	15	88.1	91.8	97.9	2.3
HN05	2	2.51	12.4	90.0	92.5	97.5	2.2
HN06	2.12	2.51	15.6	93.5	96.1	98.5	1.6
HN07	2	2.51	13.8	92.3	93.3	96.8	0.9
HN08	2	2.51	11.2	75.0	82.0	90.6	3.0
HN09	2.12	2.51	16	95.2	96.7	99.0	0.8
HN10	2.12	2.51	18.75	99.2	99.3	100.0	0.1

TABLE 4 Geometrical setup, gamma pass rate (γ) for various tolerances and average mean dose difference ($\overline{\Delta D}$) of the 10 breast plans. Points within the 50% isodose and at least 5 mm depth were considered in the calculation of the gamma index.

Case	PTV dose [Gy]	Field width [cm]	PTV length [cm]	γ 2 mm, 3% [%]	γ 3 mm, 3% [%]	γ 4 mm, 4% [%]	$\overline{\Delta D}$ [%]
B01	2	2.51	23.6	91.5	95.2	98.2	2.0
B02	2	2.51	14.4	73.4	81.8	91.5	2.8
B03	2	2.51	21.8	87.4	92.6	96.8	2.0
B04	1.8	2.51	21.4	98.1	99.1	99.7	1.3
B05	2	2.51	25.4	95.8	97.0	99.7	1.5
B06	2	2.51	24.6	98.2	98.8	99.7	0.5
B07	2	2.51	20.6	90.2	93.7	97.8	2.0
B08	2	2.51	20.4	85.1	91.0	96.5	2.4
B09	2	2.51	23.8	96.3	97.3	99.3	1.1
B10	2.65	2.51	20.2	71.1	88.2	91.7	2.8

TABLE 5 Number of successes to the gamma comparison test ($N_{\gamma>95\%}$, i.e., pass rate above 95%) and mean gamma pass rate ($\bar{\gamma}$) for various tolerances for the three different regions investigated. Ten treatment plans were tested in each region. Points within the 50% isodose and at least 5 mm depth were considered in the calculation of the gamma index.

	Abdomen and pelvis		Head and neck		Breast	
	$N_{\gamma>95\%}$	$\bar{\gamma}$ [%]	$N_{\gamma>95\%}$	$\bar{\gamma}$ [%]	$N_{\gamma>95\%}$	$\bar{\gamma}$ [%]
2 mm, 3%	7	94.1	3	86.8	4	88.7
2 mm, 4%	7	97.2	7	93.1	5	92.3
3 mm, 3%	7	95.9	4	90.3	5	93.5
3 mm, 4%	8	97.9	7	95.0	8	94.9
3 mm, 5%	10	98.9	7	97.5	9	98.0
3 mm, 6%	10	99.4	10	98.5	10	98.9
4 mm, 4%	8	98.3	7	95.9	8	97.1
$\overline{\Delta D}$ [%]	1.1		1.8		1.8	

lines) show a good agreement with the measurements (dots). The maximal error induced by the approximation of Eq. (3) is of respectively 2.7% and 1.5% for the 1 cm and 2.5 cm edge fields. Also note

that in Eq. (3), the spatial coordinate is the angular distance at the source, but that the field profiles are represented in Fig. 2 in Cartesian coordinates.

Figure 3 shows the relative jaw penumbral filters of the 2.5 cm and 5 cm nominal field, defined by Eq. (2) and calculated using the measured profiles shown in Fig. 2. One can see that the RJPFs are depth-dependent, as are the field profiles. Also, one should note that they do not converge toward 0 when reaching the field limit, as would be expected. This is a numerical artifact: obviously, both the nominal and edge field profiles also tend toward 0 at the field boundary, and dividing two small values one with another [in Eq. (2)] may result in large numbers. In other words, the RFJPFs are hardly calculable outside the field. Though, this is not an issue because the product of the profile and the RJPF [in Eq. (3)] converges toward 0 at the field limit. As can be seen on Fig. 2, the calculated edge field profiles (plain lines) match the measurements (dotted lines).

Finally, CheckTomo upgrade was designed having in mind that TomoEDGE could in the future evolve and perform more complex dose sculpting. One can think of sharpening the edges of a simultaneous integrated boost (SIB) or tracking a tumor.

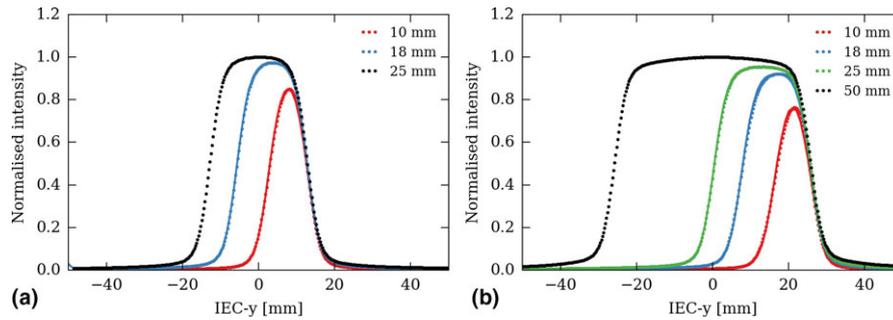


FIG. 2. Longitudinal profiles, along the machine IEC-y axis, of the 2.5 cm (a) and 5 cm (b) on-axis nominal fields (black) and their related off-axis edge fields (colored). Dots represent measurements, plain lines the edge profiles calculated from Eq. (3). Measurements were performed at 1.5 cm depth, all leaves open, on the positive IEC-y side of the gantry. Values in the legend correspond to the field width at isocenter.

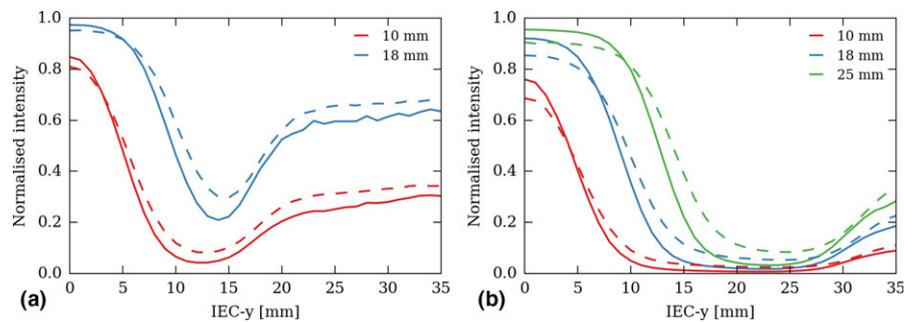


FIG. 3. Relative jaw penumbral filters of the 2.5 cm (a) and 5 cm (b) nominal fields for each of the three off-axis edge field widths sampled, along the machine longitudinal (IEC-y) axis. RJPF are shown at depth 1.5 cm (solid lines) and 10 cm (dashed lines).

3.B | Dose calculation gamma pass rate

3.B.1 | Gradient verification

Gamma index pass rates for all five plans calculated in the TomoPhant are given in Table 1. With the 2.5 cm field, the pass rate is high (99.8%) for the 6 cm and 10 cm target. For the 2 cm target, the index tolerance must be increased to 3 mm/4%. Note that this case was designed for testing purposes. In clinical practice, it would not make sense to try to cover a 2 cm long PTV with the 2.5 cm wide field and the 1 cm field would have been used instead.

The gamma pass rates of the plans calculated with the 5 cm field are lower, below 90% for the 2 mm/3% tolerance. As can be seen in Fig. 4 (b), the dose calculation is perturbed over 5 cm by the approximation of the varying field width profile. Though, this figure also shows that calculation of the field gradient by CheckTomo matches well that of the TPS both in space and dose.

3.B.2 | Real patient cases and errors detection

The calculation of 29,791 dose points for one case takes between 2 and 3 minutes on Intel Core i5 3.4 GHz processor, depending on the size of the region of interest considered.

CheckTomo was tested on 10 abdomen and pelvis, 10 head and neck (H&N), and 10 breast plans. For each case, the mean dose difference ($\overline{\Delta D}$) and global gamma comparison tests between the CheckTomo and the tomotherapy TPS dose distributions were calculated. Individual results of all cases are provided in Tables 2–4. For each location and tolerances, the number of plans that succeed the gamma test ($N_{\gamma > 95\%}$, i.e., pass rate above 95%) and the mean gamma pass rate ($\bar{\gamma}$) over the 10 plans are given in Table 5. One can see that plans in the abdominal and pelvic region are the most accurately calculated with at least 7 plans out of 10 succeeding the gamma comparison test. In the opposite, dose calculation for the H&N cases is

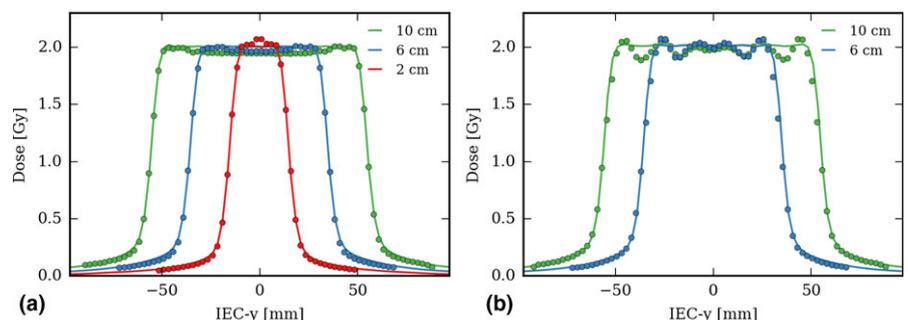


FIG. 4. Dose profiles of the plans calculated in the TomoPhant for varying PTV length, for the 2.5 cm field width (a) and 5 cm (b). Plain line corresponds to the dose calculated by the tomotherapy TPS and dots to CheckTomo dose points.

more prone to errors and requires the gamma index dose tolerance to be increased to 4% to have a majority of plans succeeding the test. This can be explained by the fact that PTVs in the abdominal and the pelvic area usually encompass large homogenous tissue volumes, while bones and air cavities can be found in the H&N region. The difference of calculation accuracy between those two kinds of location comes from the fact that the dose calculation in CheckTomo relies on a water-based model, which is obviously more reliable in tissues with densities close to water. Note that the scope of this manuscript is to describe the implementation of TomoEDGE in CheckTomo, not to suggest improvements of its calculation engine.

Concerning the breast cases, where the target volumes are often off-axis, increasing the number of subprojections per projection from 1 to 3 or 5 could improve the results accuracy.⁶

Table 6 shows the results of the error simulation tests. It concerns 15 cases that had passed the 2 mm/3% gamma comparison test. As one can see, all plans failed the 2 mm/3% gamma test when a 4 mm shift was applied longitudinally to the calculation grid or if the dose was offset by 3%.

The results of the gamma comparison tests presented here for dynamic cases are similar to what had been obtained for static jaws plans with the original release of CheckTomo.³ In other words, the overall dose calculation accuracy and sensitivity to errors is equivalent for both TomoEDGE and non-TomoEDGE plans.

Performing an independent dose calculation with CheckTomo is not as comprehensive as actually measuring it during a QA procedure, in that sense that it performs no control on the machine side. Though, CheckTomo successfully detected simulated errors exceeding tolerances. In other words, it is conservative of the quality assurance, thus can provide a good indicator of the accuracy of the dose calculation. Nonetheless, the way CheckTomo could be used in

TABLE 6 Number of cases succeeding the gamma test ($N_{\gamma>95\%}$), mean pass rate ($\bar{\gamma}$), and average mean dose difference ($\overline{\Delta D}$) for 15 treatment plans on which was applied a longitudinal shift of 2 and 4 mm and a dose offset of 3%. Only plans which had passed (without simulated error) a 2 mm/3% gamma test were considered.

	Unshifted	2 mm shift	4 mm shift	Overdosage 3 %
$N_{\gamma>95\%}$ [%]	15	13	0	0
$\bar{\gamma}$ [%]	98	96.7	83.7	63.2
$\overline{\Delta D}$ [%]	0.79	0.77	0.91	3.49

practice (e.g., replace a patient QA measurements) remains the responsibility of the local medical physicist.

3.C | Occasional edge dose calculation error

In some cases, the dose is over or under estimated in the target volume edges, as shown in Fig. 5 left-hand side. The occurrence of such errors seems random and is caused by rounding mistakes in the calculation of the dose grid coordinates. Even a submillimetric registration error between the CheckTomo and tomotherapy TPS dose distributions could lead to a dose miscalculation of several Gy within the high gradient region. Though, such a problem can be easily addressed by shifting longitudinally the TPS dose volume, using a manual registration tool included in CheckTomo since the first version. As it happens, the error appearing in Fig. 5 was corrected by applying a 1 mm shift. The result is shown on the figure right-hand side.

Even if such an error is not accounted for, it does not much impact the overall gamma pass rate of the plan (0.3% in the case of Fig. 5). The relative dose difference does usually not exceed 10% and concerns only the points located in the field edges, hence a small portion of the PTV. However, one should note that CheckTomo was not specifically designed to be a dose gradient verification tool and should not try to use it as such. CheckTomo cannot isolate a particular region of interest and lacks analysis tools dedicated to conformity verification.⁹

4 | CONCLUSION

CheckTomo software for independent dose calculation in tomotherapy was upgraded for TomoEDGE treatments by introducing the RPJF in its profile calculation model. It was noted that this method implies that a slight inaccuracy in the edge field profiles calculation has to be tolerated. The results of the gamma comparison tests demonstrated that, in terms of dose calculation accuracy and errors detection, the upgraded version of CheckTomo is as reliable for dynamic jaws plans as the former release was for static cases. This leads us to conclude that, from now on, CheckTomo offers the opportunity to perform independent dose calculation equivalently for both static and dynamic jaws tomotherapy plans.

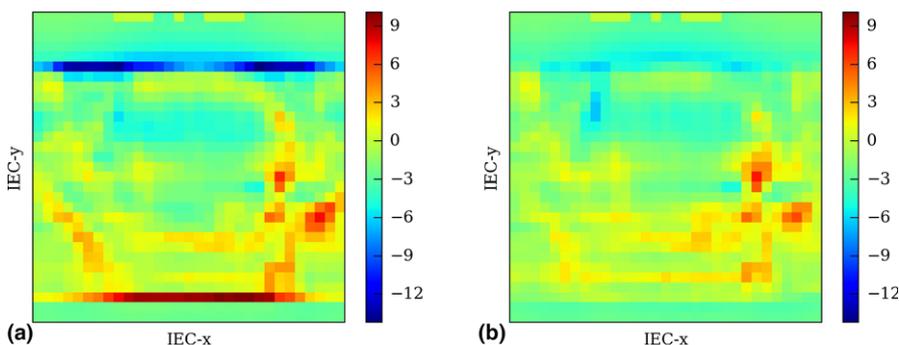


FIG. 5. Coronal view of the relative difference, given in percent, between the CheckTomo and the tomotherapy TPS dose of a pelvis plan. As one can see in figure (a), the dose is miscalculated in both edges of the target volume. The error is corrected by applying a 1 mm shift in the longitudinal direction (IEC-y) to the TPS dose distribution, as shown in figure (b).

CONFLICT OF INTEREST

R. Moeckli is holding a grant from Accuray for a research project in tomotherapy. However, the present work is not directly related to that grant.

REFERENCES

1. IAEA. *Technical Reports Series No. 430: Commissioning and Quality Assurance of Computerized Planning Systems for Radiation Treatment of Cancer*; 2004.
2. Gibbons JP, Smith K, Cheek D, Rosen I. Independent calculation of dose from a helical TomoTherapy unit. *J Appl Clin Med Phys*. 2009;10:2772.
3. Thomas SJ, Eyre KR, Tudor GSJ, Fairfoul J. Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. *Med Phys*. 2012;39:160–167.
4. Accuray Inc. *Physics training on the TomoEDGE feature ETT.700064.D*; 2013.
5. Katayama S, Haefner MF, Mohr A, et al. Accelerated tomotherapy delivery with TomoEdge technique. *J Appl Clin Med Phys*. 2015;16:4964.
6. Tudor GSJ, Thomas SJ. Impact of the fixed gantry angle approximation on dosimetric accuracy for helical tomotherapy plans. *Med Phys*. 2013;40:011711.
7. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys*. 1998;25:656–661.
8. Thomas SJ, Cowley IR. A comparison of four indices for combining distance and dose differences. *Int J Radiat Oncol Biol Phys*. 2012;82:e717–e723.
9. Feuvret L, Noël G, Mazeron J-J, Bey P. Conformity index: a review. *Int J Radiat Oncol Biol Phys*. 2006;64:333–342.

Appendix B

In air and in vivo measurement of the leaf open time in tomotherapy using the on-board detector pulse-by-pulse data

Mathieu Schopfer, François Bochud, Jean Bourhis, and Raphaël Moeckli

In: *Medical Physics* (2019)

DOI: 10.1002/mp.13459

In air and *in vivo* measurement of the leaf open time in tomotherapy using the on-board detector pulse-by-pulse data

Mathieu Schopfer, and François O. Bochud

Institute of Radiation Physics, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Jean Bourhis

Radiation-oncology department, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Raphaël Moeckli^{a)}

Institute of Radiation Physics, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

(Received 20 September 2018; revised 11 February 2019; accepted for publication 19 February 2019; published xx xxxx xxxx)

Purpose: We developed an algorithm to measure the leaf open times (LOT) from the on-board detector (OBD) pulse-by-pulse data in tomotherapy. We assessed the feasibility of measuring the LOTs in dynamic jaw mode and validated the algorithm on machine QA and clinical data. Knowledge of the actual LOTs is a basis toward calculating the delivered dose and performing efficient phantom-less delivery quality assurance (DQA) controls of the multileaf collimator (MLC). In tomotherapy, the quality of the delivered dose depends on the correct performance of the MLC, hence on the accuracy of the LOTs.

Materials and methods: In the detector signal, the period of time during which a leaf is open corresponds to a high intensity region. The algorithm described here locally normalizes the detector signal and measures the FWHM of the high intensity regions.

The Daily QA module of the TomoTherapy Quality Assurance (TQA) tool measures LOT errors. The Daily QA detector data were collected during 9 days on two tomotherapy units. The errors yielded by the method were compared to these reported by the Daily QA module.

In addition, clinical data were acquired on the two units (25 plans in total), in air without attenuation material in the beam path and *in vivo* during a treatment fraction. The study included plans with fields of all existing sizes (1.05, 2.51, 5.05 cm). The collimator jaws were in dynamic mode (TomoED-GETM). The feasibility of measuring the LOTs was assessed with respect to the jaw aperture.

Results: The mean discrepancy between LOTs measured by the algorithm and those measured by TQA was of 0 ms, with a standard deviation of 0.3 ms. The LOT measured by the method had thus an uncertainty of 1 ms with a confidence level of 99%.

In 5.05 cm dynamic jaw procedures, the detector is in the beam umbra at the beginning and at the end of the delivery. In such procedures, the algorithm could not measure the LOTs at jaw apertures between 7 and maximum 12.4 mm. Otherwise, no measurement error due to the jaw movement was observed. No LOT measurement difference between air and *in vivo* data was observed either.

Conclusion: The method we propose is reliable. It can equivalently measure the LOTs from data acquired in air or *in vivo*. It handles fully the static procedures and the 2.51 cm dynamic procedures. It handles partially the 5.05 cm dynamic procedures. The limitation was evaluated with respect to the jaw aperture. © 2019 American Association of Physicists in Medicine [https://doi.org/10.1002/mp.13459]

Key words: leaf open time, machine QA, tomotherapy

1. INTRODUCTION

Intensity-modulated radiation therapy (IMRT) is a widely used treatment modality.¹ It enables the delivery of highly inhomogeneous dose distributions. Because of this inherent complexity, international guidelines recommend performing systematic pretreatment delivery quality assurance (DQA) controls of IMRT plans.^{2,3} The most common DQA protocols involve measuring a dose distribution in a virtual water phantom using a film, ion-chambers, or a detector array. The quality of the

measured distribution is usually assessed through statistical criteria (e.g., a γ -index percent pass rate⁴). These protocols verify the overall correspondence between the planned and the delivered dose. Their advantage is that they provide a check of the overall performance of the treatment unit.

The most commonly used pretreatment DQA protocols present three major drawbacks. First, the phantom-based measurements are usually time consuming. Phantom setup, positioning, and other repetitive tasks are carried out manually. An automated phantom-less protocol would lighten the

workload. Second, the usual acceptance criteria may not be suitable for dose error detection, especially when incorrectly used.^{5–9} Another option would be to calculate the delivered dose on the patient planning images using the fluence measured inline by the treatment system. It would make it possible to use dose-oriented criteria, which are more reliable.⁵ Third, there is no control of the individual treatment fractions. A comprehensive DQA tool would work *in vivo*.

Tomotherapy (Accuray Inc., Sunnyvale, CA, USA) is a topographic and helical IMRT system. It provides image guided radiation therapy by megavoltage computed tomography (MVCT). The on-board MVCT detector (OBD) is mounted opposite to the LINAC on the gantry ring. It measures the exit fluence per LINAC pulse during imaging and treatment delivery. The binary multileaf collimator (MLC) modulates the beam. The open times of the MLC leaves are optimized to reach the desired dose distribution.

Delivered dose accuracy and patient safety depend on the MLC performance, i.e., on the correctness of the leaf open times (LOT). The manufacturer has implemented mechanisms to prevent severe MLC failures, yet, it tolerates smaller errors. These can lead to observable dose discrepancies,^{10–14} but phantom-based DQA procedures do not evaluate the actual dose distribution inside the patient.

Some authors have used the detector data to measure the LOTs.^{14–17} Others used the exit fluence to verify MLC operation^{18–20} or to reconstruct the delivered dose without measuring the LOTs.^{11,21–23} Handsfield *et al.*¹² and Deshpande *et al.*¹⁴ calculated the delivered dose from the measured LOTs. Handsfield *et al.*¹² used the method described by Chen *et al.*¹⁵ to measure the LOTs. They calculated the delivered dose of individual treatment fractions for some patients.

Around a decade ago, Accuray started developing a dynamic jaw mode. It allows changing the field size and shifting the field longitudinally during treatments.²⁴ In 2012, Accuray released the TomoEDGE™ feature. It reduces the penumbra on the target cranial and caudal edges. Dynamic jaw delivery is intended to offer more possibilities in the future, e.g., motion management.²⁵ However, when the field edge is too distant from the source axis, the beam is off the detector. Thus, only the scattered fluence is recorded. It may be impossible to measure the LOTs or to reconstruct the dose from the detector raw data at certain jaw apertures.

As stated above, in tomotherapy, the MLC performance is a major concern for dose accuracy and safety. What is more, dynamic jaw delivery is obviously becoming the standard. Therefore, there is a strong interest in measuring the LOTs in air and *in vivo*, for both the static and the dynamic jaw modes. As far as we know, only Chen *et al.*¹⁵ have measured the LOTs in air and *in vivo*. Their method was designed to handle static jaw procedures because TomoEDGE™ had not been released at that time.

As is highlighted in the present paper, detector data collected under dynamic jaws cannot be parsed as data collected under static jaws. Therefore, we developed an algorithm that aims to measure the LOTs from the detector pulse-by-pulse data of both static and dynamic jaw procedures. It is based on

the work of Chen *et al.*¹⁵ We assessed in which segments of the dynamic jaw procedures the algorithm could reliably measure the LOTs. We validated the algorithm on machine QA data and on a broad sample of clinical treatment data acquired both in air and *in vivo* under all jaw settings.

2. MATERIALS AND METHODS

2.A. Tomotherapy system

All data used in this study were collected in our institution on two TomoHDA™ units, hereafter called TOMO1 and TOMO2. Both these units offer every delivery mode, i.e., helical or topographic delivery with either static or dynamic jaws.

2.A.1. Multileaf collimator and leaf open time

Tomotherapy units use a binary MLC with 64 leaves for beam modulation. The leaves can only be in an open or a closed position, and the beam modulation consists of varying the LOTs. The leaves are activated by compressed air. This allows for a fast travel — around 10 ms — between the two positions.

In helical mode, the gantry rotations are split into 51 sectors. A sector is called a sinogram projection. It extends over 7.06°. The rotation period depends on the target dose, the pitch, and the field size. It is fixed between 12 and 60 s. Equivalently, the projection time is fixed between 235 and 1176 ms.

The LOTs are optimized across the projections. They range from an arbitrary minimum threshold to the projection time. The threshold is set in the jaw accelerator machine (JAM) settings. On our units, it is 18 ms.

2.A.2. Leaf latency, sinograms, and leaf motion advance

The leaves need some time to move from one position to the other once a motion command is issued. This phenomenon is called leaf latency. The maximum latencies admitted are specified in the machine calibration parameters. On our units, they are set to 70 ms. The latency causes variations of the effective LOTs. The average variation is measured and adjusted accordingly in the JAM settings. The treatment planning system (TPS) considers it.

The MLC sinogram is a matrix with a line per leaf and a row per projection. It provides the LOTs. The sinogram of any treatment plan exists in two versions. One is “machine specific”. It contains the LOTs corrected of the variation caused by the latency. It is meant for a correct delivery of the planned fluence. The other is “machine agnostic”. It contains the optimal LOTs used for dose calculation. When delivering, the actual LOTs should ideally be equal to the optimal LOTs.

The leaves are planned to open and close symmetrically (in time and angle) around the projection center. Yet in practice, the leaves move before their planned motion time. It is not clear to us why, but we suspect that it must help to reduce

the apparent latencies. Daily QA, a module of the TomoTherapy Quality Assurance (TQA) tool,²⁶ measures the leaf motion advance on a daily basis. It reports that the leaves effectively move ahead of their planned open or close point of time by 5–20 ms (2–6 LINAC pulses). The effective advance varies from leaf to leaf and projection to projection.

2.A.3. On-board MVCT detector and pulse-by-pulse data

Tomotherapy units employ MVCT detector arrays. These are made up of ionization chambers filled with high-pressure xenon, separated by tungsten septa. Our units have detectors of different generations and from different manufacturers: GE (General Electric, Boston, MA, USA) on TOMO1 and Hitachi (Hitachi Ltd., Chiyoda, Tokyo, JP) on TOMO2. Both detectors possess 640 chambers. They differ in chamber size (1.219 and 1.250 mm, respectively), distance from surface to center (20 and 4 mm, respectively), and radius of curvature (110.99 and 99.8 cm, respectively).

The detector measures the fluence exiting the patient per LINAC pulse, i.e., at 80 Hz during imaging and 300 Hz during treatment. These are the pulse-by-pulse data or uncompressed raw data. The detector measurements are stable in time.²⁷

The radiation delivery system (RDS) collects the raw data. It stores them temporarily on the machine's on-board computer in a binary file. It erases them when a new delivery procedure is loaded. The raw data file can be retrieved via FTP after the procedure delivery. It contains an array with a line-per-pulse and a column-per-channel, along with other information.

2.A.4. Field size and jaw mode

The collimator jaws delimit the field size along the patient's longitudinal direction. The jaw aperture (projected in the isocenter plane) is preset to either 7, 20, or 42 mm, resulting in a field FWHM of 1.05, 2.51, or 5.05 cm. The jaw presets are commonly referred to as J07, J20, and J42. TomoEDGE™ is the only feature based on the dynamic jaw mode to have been released so far. It is available for the J20 and the J42 presets. It is meant to reduce the dose penumbra on the target region cranial and caudal edges.²⁴

At the beginning and at the end of dynamic procedures, the jaws are located on the same side of the source axis. Under the J42 preset, the jaws cross up to 14 mm (in projection in the isocenter plane) over the axis. The detector is thus momentarily in the beam umbra at the beginning and at the end of these procedures.

2.A.5. Leaf spread functions and channel to leaf mapping

The leaf width at isocenter (6.25 mm) is greater than the detector channel width. The beamlet issued from a single leaf covers multiple detector channels. Which is more, the signal measured by the detector is the convolution product of the beam profile and the detector response to the incoming fluence.

We measured the beam spreading from each leaf at the three preset jaw apertures (7, 20, and 42 mm) and at the narrower apertures of 7, 9, 11, 13, 15, and 17 mm of the J20 and J42 presets on both sides of the source axis. From the detector signal (the beam profiles convoluted to the detector response), we determined a leaf spread function (LSF) for each leaf and each jaw positions. We fitted a sum of Gaussian functions on each signal. We measured and subtracted the background signal (caused by MLC leakage and detector dark current) at each jaw position individually.

A bijective leaf-to-channel map was determined by matching a leaf to the center of the beamlet it produced. To determine the leaves state, it is sufficient to observe the signal at these specific channels.^{14,15,17}

2.A.6. TomoTherapy Quality Assurance Daily QA module

The Daily QA module of TQA performs various machine tests. In particular, it runs two LOT tests. In the first, the leaves open 8 by 8 during 50, 100, 150, and 200 ms. In the second, all the 64 leaves open simultaneously during 100 and 150 ms. In both tests, the projection time is 200 ms and the jaws are static. The module calculates and reports the LOT error for each leaf and each test.

In Daily QA, the commended LOTs were set such that the leaf should operate normally. No leaf should fail to open and no leaf should fail to close between two successive projections. If the commended LOTs were shorter (less than 30 ms) or larger (within 30 ms of the projection time), large LOT errors (greater than 15 ms) could occur.^{15,28,29}

2.B. Leaf open time calculation from raw data

In the detector signal, the period during which a leaf is open corresponds to a distinct peak or a higher intensity region (Fig. 1). The related LOT can be calculated by determining the region width. In the following sections, we present our calculation algorithm. It relies on an in-house function written in Python to read the data from the binary files into NumPy arrays.³⁰ It outputs a sinogram of the plan sinogram size providing the measured LOTs.

The algorithm works in three steps (Fig. 2). First, it corrects the signal from arcings and deconvolve the signal spreading from leaves to leaves (Section 2.B.1). Then, it determines at each projection individually which leaves have opened and which have not, and it normalizes the signal locally (Section 2.B.2). Finally, it determines the edges of the open leaf regions and calculates the related LOTs (Section 2.B.3).

2.B.1. Raw data preprocessing

When an electrical arcing happens in the LINAC, the detector records a short and steep drop of the beam intensity.

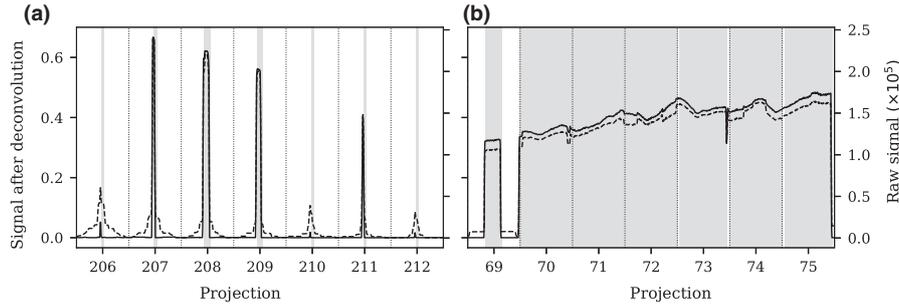


FIG. 1. Detector signal after deconvolution (solid lines) and raw signal (dashed lines) of leaf 40 (a) and leaf 26 (b). *In vivo* data from patient 1 treatment (c.f. Table I). The signal is plot over seven sinogram projections. The dotted vertical lines correspond to the projection boundaries. Areas in gray correspond to the periods during which the leaves were planned to be open. (a) The planned LOTs were short, between 19 and 61 ms. (b) The planned LOTs were of 475 ms (the projection time) at projections 70, 71, 72, and 74. They were of 433 and 436 ms at projections 73 and 75, respectively.

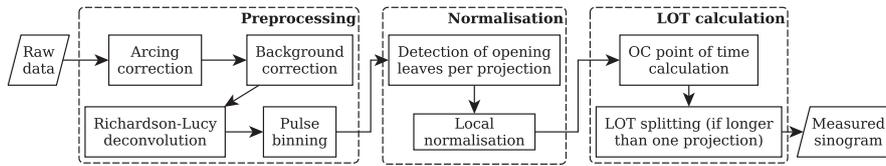


FIG. 2. Workflow of the LOT calculation algorithm.

That intensity drop can be confused with a leaf closing briefly before opening again. Therefore, the algorithm corrects the raw data whenever an arcing occurs. It interpolates temporally the impacted pulses with their neighbors. Arcings can be easily identified by observing the time derivative of the monitor chambers signal.

The raw data are interpolated at the 64 leaf specific channels according to the channel-to-leaf map (c.f., Section 2.A.5). Then, the algorithm calculates the background signal for each leaf over the machine warm-up pulses, and subtracts it.

The beamlet of a leaf spreads over the detector, i.e., an open leaf creates a parasitic signal in its neighbors' channel. To prevent this contamination from inducing LOT measurement errors, the algorithm corrects it by applying the iterative deconvolution of Richardson-Lucy,^{31,32} as proposed by Chen *et al.*¹⁵

Let us we write L_{ij} the value of leaf i spread function at the specific channel of leaf j (c.f., Section 2.A.5). L_{ij} is the fraction of leaf i signal measured in leaf j channel. Writing d_i the detector signal measured under leaf i during any specific pulse, the Richardson-Lucy algorithm yields

$$\hat{d}_i^{t+1} = \hat{d}_i^t \sum_j \frac{L_{ji}d_j}{\sum_k L_{jk}\hat{d}_k^t} \quad (1)$$

t is the iteration step and \hat{d} is the estimate of d under the deconvolution.

Finally, for calculation purpose, the data are binned per sinogram projection. They are handled as a three-dimensional matrix, written D_{lpi} hereafter. The indices l , p , and i represent the leaf, the projection, and the pulse, respectively. Data are resampled if the number of pulses per projection is not integer.

2.B.2. Signal normalization

The data corresponding to leaf l at projection p should be normalized locally only if leaf l has actually opened during projection p . Otherwise, one would just amplify the noise. Therefore, the algorithm determines at each projection the leaves that have opened and normalizes only the data parts related to opening leaves.

To detect the opening leaves, the algorithm relies on a projection-dependent threshold. It defines a thresholding value per projection. On static jaw procedures, a single threshold value applied on the entire signal was suitable.¹⁵ In dynamic jaw mode, however, the threshold must account for the signal intensity drop at the beginning and at the end of the procedures (Fig. 3). The detection works in two steps.

First, the algorithm detects the projections at which at least one leaf has opened. It calculates the maximum signal measured per projection

$$\tilde{D}_p = \max_i D_{lpi}. \quad (2)$$

When \tilde{D}_p is above a specific threshold τ_{proj} , we consider that at least one leaf has opened during projection p . We used $\tau_{proj} = 0.01$.

Second, the algorithm detects the opening leaves at each of the projections it has selected at the first step. It discards the projections without opening leaves. At a particular projection p , it calculates the maximum signal measured per leaf

$$\tilde{D}_{lp} = \max_i D_{lpi}. \quad (3)$$

When

$$\tilde{D}_{lp} > \tau_{leaf}(\tilde{D}_p) \quad (4)$$

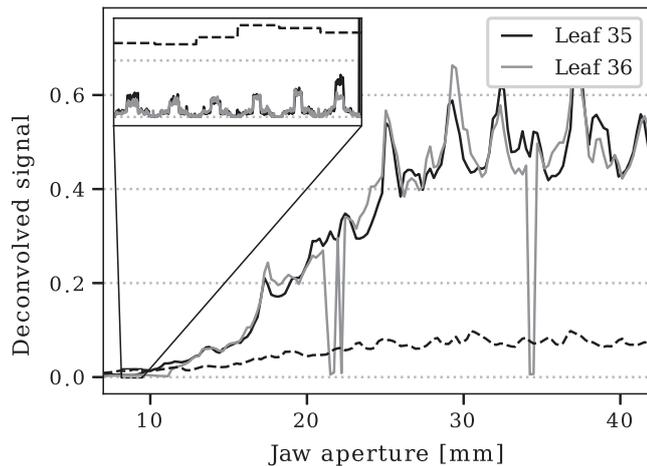


FIG. 3. Maximum detector signal per sinogram projection \tilde{D}_{lp} (3) after Richardson–Lucy deconvolution. *In vivo* data from patient 20 treatment (c.f. Table I). The data of two leaves are represented as a function of the jaw aperture, from the moment the jaws started opening (aperture of 7 mm) until they reached their maximal aperture (42 mm). Leaf 36 signal dropped near 21, 22, and 34 mm because leaf 36 was actually closed throughout the entire corresponding projections; it is not an artifact of the algorithm. The magnification shows the signal of the same leaves at all pulses, at a jaw aperture from 8.2 to 9.5 mm (projections 45–50). Leaf 35 opened at each of these six projections and leaf 36 stayed closed. The dashed line in both plots is the threshold τ_{leaf} used in (4).

we consider that leaf l has opened during projection p . The threshold τ_{leaf} is a piecewise linear function of \tilde{D}_{lp} .

Finally, the algorithm normalizes the raw data per leaf and per projection. If leaf l was detected to have opened during projection p , the corresponding data are normalized by \tilde{D}_{lp} . The algorithm accounts for the pulses possibly binned in the wrong projection (see Section 2.A.2).

2.B.3. Open time calculation

As stated in the introduction of Section 2.B, the period during which a leaf is open corresponds to a peak or a higher intensity region in the detector signal. To calculate the LOT, the algorithm determines the leaf opening and closing points of time. We have defined them as the two points at which the signal crosses a specific threshold τ_{oc} . We have fixed τ_{oc} at 0.5. In this case, the LOTs correspond to the FWHM of the open leaf regions. The signal is linearly interpolated to get a finer precision than the measurement sampling of 3.3 ms. Others took a similar approach.^{15,17}

A leaf may stay open across multiple projections [e.g., from projection 70–75 in Fig. 1(b)]. Because leaves move ahead of the planned time (c.f. Section 2.A.2), the open leaf region may span over the boundary of the first projection [in Fig. 1(b), the open leaf portion of the signal spans in projection 69]. In these cases, the algorithm shifts the signal region by the mean leaf motion advance observed over the whole procedure. Then, the region is split per projection.

2.C. Algorithm validation

We collected machine TQA data (Section 2.A.6) and clinical data to validate the LOT measurement algorithm. We collected the Daily QA raw data and test results on both our treatment units each Monday during 9 weeks. We collected clinical in air and *in vivo* (one treatment fraction) data from 25 patient treatments (Table I). The in air measurements consisted of delivering the DQA procedure without anything in the beam path (except the couch). *In vivo* measurements miss for plans 8 and 9. In total, 48 datasets were available. J20 and J42 procedures were always in dynamic jaw mode.

TABLE I. Summary of the treatment plans used to collect data, split by treatment unit. The projection and duration columns give the plan projection time and the procedure total duration, respectively.

TOMO1					TOMO2				
Patient	Localization	Jaw	Proj. (ms)	Duration (s)	Patient	Localization	Jaw	Proj. (ms)	Duration (s)
2	Thorax	J20	424	384.6	1	Mediastinum, left	J20	475	316.0
3	Lung, left	J20	333	249.0	9*	Mediastinum	J42	498	397.9
4	Thorax	J20	947	608.0	10	Breast, right	J42	614	562.8
5	Skull	J20	231	149.7	11	Lumbar vertebrae	J42	433	363.6
6	Skull	J20	294	191.5	13	Vertebra T12	J42	845	314.4
7	Rib, right	J07	690	540.4	14	Esophagus	J42	600	714.0
8*	Mediastinum	J42	314	292.7	15	Vertebrae L3 to L4	J42	535	171.3
18	Skull, peduncle	J07	418	460.7	16	Skull	J07	231	522.7
19	Skull, frontal	J07	633	270.4	17	Eye, right	J07	420	489.7
20	Vertebra T9	J42	357	149.5	24	Head & Neck	J20	231	310.7
21	Vertebrae L1 to L3	J42	412	261.1					
22	Vertebrae T3 to T6	J42	369	230.8					
23	Vertebra T11 to L3	J42	476	373.6					
26	Pelvis	J20	231	159.9					
27	Head & Neck	J20	288	309.0					

*No *in vivo* data.

We compared the LOT errors measured by our method to the errors reported by the Daily QA module to perform a basic algorithm verification (Section 2.C.1). We used clinical data to determine at which jaw aperture the LOT measurement was feasible in dynamic jaw mode and to verify whether it was independent of the jaw aperture (Section 2.C.2). Again using clinical data, we reviewed some parts of the raw signal to ensure that large discrepancies between planned and measured LOTs were no algorithmic errors and we verified that the LOTs were measured equivalently in air and *in vivo* (Section 2.C.3).

2.C.1. Comparison to Daily QA and measurement uncertainty

Comparing the LOT errors measured by our algorithm to the errors reported by Daily QA allowed us to validate the LOT measurement against an accepted reference in controlled conditions. Daily QA is performed in air with static jaws. The leaves should operate normally (see Section 2.A.6).

Additionally, we estimated the variability of the leaf motion advance (c.f., Section 2.A.2) from the Daily QA raw data and estimated the LOT measurement uncertainty.

2.C.2. Leaf open time measurement in dynamic jaw segments

We investigated whether measurement errors occurred at the beginning and at the end of dynamic procedures. In particular, the J42 data measured when the beam was off the detector (Section 2.A.4) might be unexploitable. The unexploitable data segments were determined by looking at the differences between planned and measured LOTs. As actual discrepancies are reported up to 45 ms,¹⁵ we considered that absolute differences greater than 50 ms were likely errors of the algorithm. We looked for the maximal jaw aperture at which such errors occurred on each J20 and J42 datasets. We measured the relative length of the unexploitable segments respectively to the total treatment length.

We tested whether the LOT measurement depended on the jaw aperture in dynamic procedures. We grouped the data by jaw preset (J20 or J42), by treatment unit (TOMO1 or TOMO2), and by collection type (in air or *in vivo*). For each group, we calculated the differences between planned and measured LOTs. Then, we looked for a correlation between the differences and the jaw aperture. We calculated the sliding mean and standard deviation of the differences as a function of the jaw aperture, with a window of 1 mm.

2.C.3. Leaf open time discrepancies and in air and in vivo comparison

We calculated the discrepancies between planned and measured LOTs for clinical datasets. If segments of the dynamic jaw procedures were considered as unexploitable, then they were discarded. The LOT measurement on these segments would be unreliable.

We reviewed manually the raw signal when the absolute difference between the planned and the measured LOT was greater than 15 ms. We verified that the signal was not corrupted, e.g., by excessive noise or by arcings in the LINAC.

We verified that the algorithm measured the same LOTs in air and *in vivo*. For each patient, we generated the distribution of the discrepancies measured in air and the distribution of the discrepancies measured *in vivo*. Then, we compared the two distributions.

3. RESULTS

The Richardson–Lucy algorithm executed within 20–60 s, depending on the procedure length, on a standard desktop computer. The execution time of the subsequent algorithm steps was negligible.

3.A. Leaf spread functions

Under the J42 preset, the LSFs got broader as the jaw closed (Fig. 4). Conversely, the LSFs were independent of the jaw aperture under the J20 preset.

The LSFs fitting errors were always less than 2% at the three preset jaw apertures. They were less than 3% at the narrower apertures of the J42 preset.

3.B. Comparison to Daily QA and measurement uncertainty

We calculated the differences between the LOT errors measured using our method and the errors reported by Daily QA. The mean difference was 0.0 ms and the standard deviations was 0.3 ms on both our tomotherapy units. The maximum difference was 1.7 ms. Regarding the standard deviations, the algorithm measured the LOTs with an uncertainty of 1 ms at a confidence level of 99%.

The mean standard deviation of the leaf movement advance was of 0.7 LINAC pulse, or 2.2 ms. When a leaf

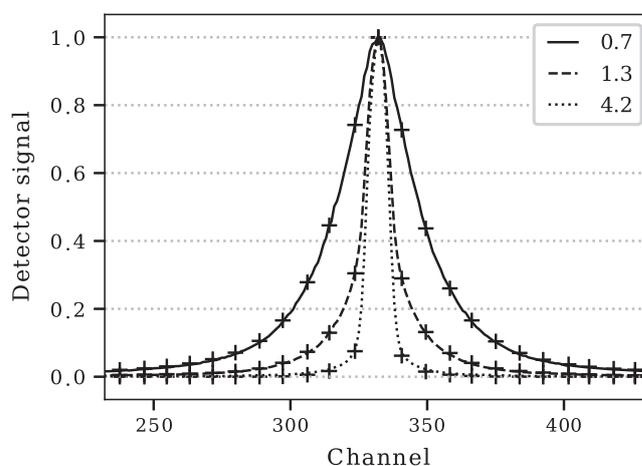


Fig. 4. Raw data (lines) and fit (cross markers) of leaf 31 spread functions at three different jaw apertures of the J42 preset measured on unit TOMO1. The fit is represented only at the neighbor leaves' channels.

TABLE II. Maximal jaw aperture (mm) under which LOT differences greater than 50 ms were measured on the in air and *in vivo* J42 procedures' data. Lines *TTT fraction* and *Dynamic jaw fraction* give the fraction of the total treatment length and of the dynamic jaw segments, respectively, in which the data were unexploitable.

	TOMO1					TOMO2					
Patient	8	20	21	22	23	9	10	11	13	14	15
In air	11.2	11.1	11.2	11.2	11.3	10.5	10.6	10.7	10.6	9.0	11.0
<i>In vivo</i>		11.8	11.0	11.0	11.3		12.4	10.1	10.6	9.7	11.2
TTT fraction (%)	6.2	14.4	9.2	5.9	7.3	7.1	5.8	5.4	13.1	3.4	8.0
Dynamic jaw fraction (%)	17.1	17.0	17.1	11.4	17.3	16.7	16.5	13.8	14.6	12.3	11.5

stays open across multiple projections, the uncertainty on the first and on the last projection LOTs is higher. The measured LOTs have then an uncertainty of 6 ms at a confidence level of 99%.

3.C. Leaf open time measurement in dynamic jaw segments

On J20 procedures, we observed no LOT differences greater than 50 ms. On J42 procedures, we observed differences greater than 50 ms up to a jaw aperture of 12.4 mm (Table II). There was no significant difference between in air and *in vivo* measurements. The unexploitable segments represented between 3.4 and 14.4% of the total treatment length (Table II).

Figure 3 illustrates the signal drop at the beginning of a J42 procedure. It illustrates that it was sometimes impossible to distinguish an open leaf signal from the noise at apertures below 13 mm. In particular, at jaw apertures between 8.2 and 9.5 mm (magnified in Fig. 3), leaf 35 opened at all projections while leaf 36 stayed closed. Yet, the signals of leaves 35 and 36 are identical at most of these projections.

Figure 5 shows the differences between planned and measured LOTs for the J20 procedures and the J42 procedures measured *in vivo* on TOMO1. It shows the sliding mean and sliding standard deviation calculated over a window of 1 mm. In J20 procedures and above 12 mm in J42 procedures, neither the mean nor the standard deviation correlated with the jaw aperture. We made the same observation for the data measured in air and on TOMO2.

3.D. Leaf open time discrepancies and in air and *in vivo* comparison

In the processing of J42 data, we discarded the projections at which the jaw aperture was below 13 mm. We had found unexploitable data segments up to 12.4 mm (Section 3.C). Figure 6 shows the distributions of the discrepancies between planned and measured LOTs. The largest discrepancies were always less than 45 ms.

We looked at the planned LOTs and signal segments corresponding to the largest discrepancies. In some instances, the measured LOT was more than 15 ms shorter than expected. Then, the planned LOT was below 30 ms. In these situations, no peaks — or only traces of them — were visible in the raw signal [e.g., projections 206, 210, and 212 in Fig. 1(a)]. Similarly, the measured LOT was more than 15 ms larger than expected in some other instances. The corresponding planned LOT was within 45 ms of the projection time in two successive projections at least. In these situations, we observed that the signal did not drop — or only partially — between the projections [e.g., projections 73 and 75 in Fig. 1(b)].

Considering any specific procedure, the in air and *in vivo* discrepancy distributions had similar shapes. The means were within 0.5 ms of each other, except for patient 1 and 5. In these instances, the difference between both means was 0.9 and 3.5 ms, respectively. The distributions were only shifted relatively to each other. They still had the same shape. In particular, the interquartile range (IQR) did not change.

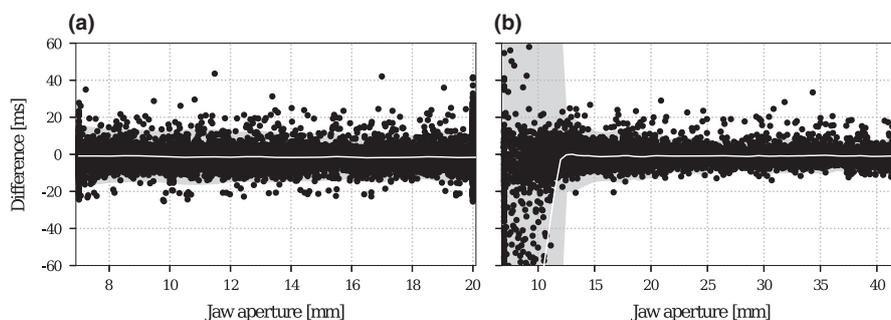


FIG. 5. Differences between planned and measured LOTs as a function of the jaw aperture. All the J20 (a) and J42 (b) *in vivo* data collected on TOMO1 were grouped together. The white line is the sliding mean of the differences for a window of 1 mm. The gray area corresponds three times the sliding standard deviation (same window).

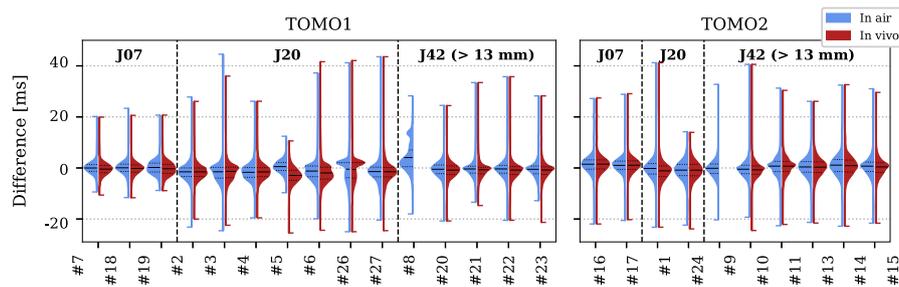


FIG. 6. Violin plots of the discrepancies between planned and measured LOTs for all the procedures measured on TOMO1 (left) and TOMO2 (right). The horizontal lines show the means, and the first and third quartiles. The labels J07, J20, and J42 correspond to the jaw settings. J42 data below a jaw aperture of 13 mm were discarded.

4. DISCUSSION

Here we present a method to measure the LOTs from the uncompressed detector raw data and its validation on two different tomotherapy units. The method aimed to extend previous works to the dynamic jaw mode. The results show that detector data are not exploitable by our algorithm below a jaw aperture of maximum 12.4 mm in dynamic J42 procedures. However, they show that the LOTs were reliably measured both in air and *in vivo* under the other jaw settings.

Our work extends the previously proposed method for in air and *in vivo* LOT measurement.¹⁵ We showed that the LSF depend on the jaw aperture under the J42 preset (Fig. 4) and incorporated these jaw-dependent functions in the algorithm. Additionally, our algorithm uses a dynamically adapting threshold to normalize the signal (Section 2.B.2). In dynamic J42 procedures, the recorded signal always drops to zero as the jaws closed (Fig. 3). The signal amplitude at 12 mm is of the same order as the signal noise at 40 mm. Using the same threshold at all jaw apertures would not be optimal. A static threshold may also be suboptimal on dynamic J20 procedures, especially in the presence of attenuation material.

The unexploitable data segments represented between 3.4% and 14.4% of the total treatment length. It would not be responsible to discard 14% of the treatment projections of a treatment plan during the DQA process. Yet, in such cases, it is still possible to measure the LOTs at all projections in air by delivering with fixed jaws. A procedure with fixed jaws can be generated from the treatment plan. With respect to the LOT measurement, this would not make any difference. Changing the jaw mode does not modify the planned LOTs and our algorithm does not verify the jaw positions anyway. The local normalization wipes the jaw effect out of the signal (Section 2.B.2). Moreover, the tolerated jaw motion uncertainty has a low dosimetric impact and Daily QA checks the jaw health regularly.

Analyzing more J42 plans could make it possible to highlight, or not, a difference in the extent of the unexploitable data segments between our two treatment units. The unexploitable data extent could be detector dependent. In addition, the segments size could be more formally expressed in terms of the signal-to-noise ratio of the detector signal.

The discrepancies, between planned and measured LOTs, greater than 15 ms were not algorithmic failures. The manual

review of the raw signal excluded this possibility. This highlights the limits of considering that the MLC is purely binary. A leaf is only considered either fully closed or fully opened, even while it is moving. This does not correspond to the reality. A leaf might stay closed or open only partially when the planned LOT is less than 30 ms [e.g., projections 206, 210, and 212 in Fig. 1(a)]. Moreover, a leaf might stay opened or close only partially when the planned LOTs are within 45 ms of the projection time in at least two successive projections [e.g., projections 73 and 75 in Fig. 1(b)].

The results indicate that, otherwise, the algorithm measured the LOTs reliably. The agreement between the LOTs measured by the algorithm and by TQA was high, indicating that the LOT measurement was reliable in air under static jaws. Then, except at the signal segments reported as unexploitable, the narrowing of the jaws did not increase the LOT measurement uncertainty. The mean and standard deviation of the differences between planned and measured LOTs did not depend on the jaw aperture (Fig. 5). In the dynamic jaw segments, the difference mean oscillated between -1.5 and 1.5 ms, which is comparable to the discrepancy means in the J07 (static jaw) procedures (Fig. 6). Finally, the beam attenuation and scattering caused by the patient did not increase the measurement uncertainty either. The algorithm measured equivalent LOTs in air and *in vivo* (Fig. 6). Therefore, as the algorithm measured the LOTs reliably in air under static jaws, independently of the jaw aperture, and independently of the patient attenuation, we assess that it measured them reliably in all situations.

The discrepancies we found between planned and measured LOTs are consistent with the findings of others.^{14–17,28,29} We could not otherwise compare the method to an external standard on clinical procedures, because there is none. Chen *et al.*¹⁵ modified the RDS software to make it record the output of the leaf position optical sensors. We did not have this opportunity because we collected the data on units in clinical use.

Our LOT measurement method can be the basis of a convenient pretreatment and *in vivo* DQA tool. A comprehensive DQA protocol would verify the dose output and the patient position. Accuray has started going in this direction with the release of Delivery Analysis (2015), a detector data processing software. Among others, Delivery Analysis provides a pretreatment dose-based QA functionality relying on

Accuray's standalone dose calculator.³³ As of today, Delivery Analysis does not measure the LOTs of the projections located in the dynamic jaw regions of the plan. Our algorithm measures the leaf opening and closing points of time before calculating the associated LOTs. It is thus compatible with Accuray's dose calculator.³³

5. CONCLUSION

We developed a method to measure the LOTs from the detector pulse-by-pulse data. The algorithm could not measure the LOTs on the first and last segments of the J42 dynamic jaw procedures at which the jaw aperture was the narrowest (from 7 to maximum 12.4 mm). It measured the LOTs otherwise reliably, with or without a patient on the treatment couch.

The method we propose is a first step toward an efficient pretreatment and *in vivo* DQA tool, and toward a more accurate knowledge of the delivered dose in tomotherapy.

ACKNOWLEDGMENTS

We thank the RTTs who willingly and carefully helped to collect the numerous data used in this study. We also thank D. Casey and C. Maurer (Accuray, Inc.) for the fruitful discussions and technical support.

CONFLICT OF INTEREST

Accuray funded this research.

^{a)}Author to whom correspondence should be addressed. Electronic mail: raphael.moeckli@chuv.ch.

REFERENCES

1. Grau C, Defourny N, Malicki J, et al. Radiotherapy equipment and departments in the european countries: final results from the ESTRO-HERO survey. *Radiother Oncol.* 2014;112:155–164.
2. Alber M, Broggi S, De Wagter C, et al. Guidelines for the verification of IMRT, Brussels, Belgium: ESTRO publishing; 2008.
3. Ezzell GA, Galvin JM, Low D, et al. Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT subcommittee of the AAPM radiation therapy committee. *Med Phys.* 2003;30:2089–2115.
4. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys.* 1998;25:656–661.
5. Zhen H, Nelms BE, Tomé WA. Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA. *Med Phys.* 2011;38:5477–5489.
6. Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V, Gabriele P. Pre-treatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram. *Med Phys.* 2012;39:7626–7634.
7. Thomas SJ, Cowley IR. A comparison of four indices for combining distance and dose differences. *Int J Radiat Oncol Biol Phys.* 2012;82:e717–e723.
8. Bresciani S, Di Dia A, Maggio A, et al. Tomotherapy treatment plan quality assurance: the impact of applied criteria on passing rate in gamma index method. *Med Phys.* 2013;40:121711.
9. Stasi M, Bresciani S, Miranti A, et al. EP-1587: sensitivity and specificity of gamma index method for tomotherapy plans. *Radiother Oncol.* 2016;119:S737.
10. Westerly DC, Soisson E, Chen Q, et al. Treatment planning to improve delivery accuracy and patient throughput in helical tomotherapy. *Int J Radiat Oncol Biol Phys.* 2009;74:1290–1297.
11. Sheng K, Jones R, Yang W, et al. 3d dose verification using tomotherapy CT detector array. *Int J Radiat Oncol Biol Phys.* 2012;82:1013–1020.
12. Handsfield LL, Jones R, Wilson DD, Siebers JV, Read PW, Chen Q. Phantomless patient-specific Tomotherapy QA via delivery performance monitoring and a secondary monte carlo dose calculation. *Med Phys.* 2014;41:101703.
13. Deshpande S, Geurts M, Vial P, Metcalfe P, Lee M, Holloway L. Clinical significance of treatment delivery errors for helical Tomotherapy nasopharyngeal plans – a dosimetric simulation study. *Physica Med.* 2017a;33:159–169.
14. Deshpande S, Xing A, Metcalfe P, Holloway L, Vial P, Geurts M. Clinical implementation of an exit detector-based dose reconstruction tool for helical tomotherapy delivery quality assurance. *Med Phys.* 2017b;44:5457–5466.
15. Chen Q, Westerly D, Fang Z, Sheng K, Chen Y. Tomotherapy MLC verification using exit detector data. *Med Phys.* 2011;39:143–152.
16. Sevillano D, Mínguez C, Sánchez A, Sánchez-Reyes A. Measurement and correction of leaf open times in helical tomotherapy. *Med Phys.* 2012;39:6972–6980.
17. Lissner S, Schubert K, Klüter S, Oetzel D, Debus J. A method for testing the performance and the accuracy of the binary MLC used in helical tomotherapy. *Zeitschrift für Medizinische Physik.* 2013;23:153–161.
18. Seibert RM, Ramsey CR, Garvey DR, Hines JW, Robison BH, Outten SS. Verification of helical tomotherapy delivery using autoassociative kernel regression. *Med Phys.* 2007;34:3249–3262.
19. Pisaturo O, Miéville F, Tercier P-A, Allal AS. An efficient procedure for tomotherapy treatment plan verification using the on-board detector. *Phys Med Biol.* 2015;60:1625–1639.
20. Pisaturo O, Miéville F, Tercier P-A, Allal AS. TransitQA — a new method for transit dosimetry of tomotherapy patients. *Med Phys.* 2017;45:438–447.
21. McNutt TR, Mackie TR, Paliwal BR. Analysis and convergence of the iterative convolution/superposition dose reconstruction technique for multiple treatment beams and tomotherapy. *Med Phys.* 1997;24:1465–1476.
22. Kapatoes JM, Olivera GH, Ruchala KJ, Smilowitz JB, Reckwerdt PJ, Mackie TR. A feasible method for clinical delivery verification and dose reconstruction in tomotherapy. *Med Phys.* 2001;28:528–542.
23. Schombourg K, Bochud FO, Mirimanoff R-O, Moeckli R. 3d dose reconstruction for narrow beams using ion chamber array measurements. *Zeitschrift für Medizinische Physik.* 2012;22:123–132.
24. Sterzing F, Uhl M, Hauswald H, et al. Dynamic jaws and dynamic couch in helical tomotherapy. *Int J Radiat Oncol Biol Phys.* 2010;76:1266–1273.
25. Schnarr E, Beneke M, Casey D, et al. Feasibility of real-time motion management with helical tomotherapy. *Med Phys.* 2018;45:1329–1337.
26. Balog J, Holmes T, Vaden R. Helical tomotherapy dynamic quality assurance. *Med Phys.* 2006;33:3939–3950.
27. Schombourg K, Bochud F, Moeckli R. Stability of the helical Tomotherapy hi-art II detector for treatment beam irradiations. *J Appl Clin Med Phys.* 2014;15:119–127.
28. Balog J, Olivera G, Kapatoes J. Clinical helical tomotherapy commissioning dosimetry. *Med Phys.* 2003;30:3097–3106.
29. Fenwick JD, Tomé WA, Jaradat HA, et al. Quality assurance of a helical tomotherapy machine. *Phys Med Biol.* 2004;49:2933–2953.
30. van dWS, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. *Computing in Science & Engineering.* 2011;13:22–30.
31. Richardson WH. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America.* 1972;62:55.
32. Lucy LB. An iterative technique for the rectification of observed distributions. *The Astronomical Journal.* 1974;79:745.
33. Chen Q, Weiguo L, Chen Y, Chen M, Henderson D, Sterpin E. Validation of GPU based Tomotherapy dose calculation engine. *Med Phys.* 2012;39:1877–1886.

Appendix C

A delivery quality assurance tool based on the actual leaf open times in tomotherapy

Mathieu Schopfer, François Bochud, Jean Bourhis, and Raphaël Moeckli

Status : submitted.

In: *Medical Physics*

A delivery quality assurance tool based on the actual leaf open times in tomotherapy

Mathieu Schopfer,^{*} François Bochud, and Raphaël Moeckli[†]

Institute of Radiation Physics, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Jean Bourhis

Radiation-oncology department, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

(Dated: February 3, 2019)

Purpose: To validate a delivery quality assurance (DQA) protocol for tomotherapy based on the measurement of the leaf open times (LOTs). In addition, to show the correlation between the mean relative LOT discrepancy and the dose deviation in the planning target volume (PTV).

Materials and methods: We used a LOT measurement algorithm presented in a previous work.

We generated TomoPhant plans with intentional random LOT discrepancies following Gaussian distributions of -6% , -4% , -2% , 2% , 4% , and 6% . We irradiated them on the Cheese Phantom with two ion chambers and collected the raw data. Using the raw data, we measured the actual LOTs and verified that the induced discrepancies were highlightable. Then we calculated the actual dose using Accuray’s standalone calculator and verified that the calculated dose was in agreement with the ion chamber measurements.

We chose randomly 60 clinical treatment plans and delivered them in air. We calculated the actual LOTs from the raw data and calculated the actual dose using Accuray’s standalone calculator. We assessed the correlation between the mean relative LOT discrepancy and the deviation between expected and actual dose in the PTV.

Results: The mean relative LOT discrepancy measured by the algorithm on the modified TomoPhant plans was within 1% of the expected discrepancy. The measured and calculated dose agreed within 1% .

On clinical plans, the means of the relative LOT discrepancies ranged between -3.0% and 1.4% . The dose deviation the PTVs ranged between -1.6% and 2.4% . The correlation between the mean relative LOT discrepancy and the dose deviation in the PTV was of 0.76 on one treatment unit and of 0.65 on another.

Conclusion: The method allowed to highlight correctly the LOT discrepancies on the TomoPhant plans. The dose subsequently calculated was accurate. The mean LOT discrepancy correlated with the dose deviation in the PTV. This makes the mean LOT discrepancy a handy indicator of the plan quality.

I. INTRODUCTION

Pretreatment patient-specific delivery quality assurance (DQA) is an essential, often mandatory step in the intensity modulated radiotherapy (IMRT) workflow [1–3]. It protects the patient against unexpected behaviours of the treatment units that may lead to inadequate dose delivery. Undesirable effects may range from poor covering of the planning target volume (PTV), hence reducing the probability of complication-free tumour control, to potentially lethal accidents.

Common pretreatment DQA protocols consists in measuring the delivered dose in a phantom with films, ion-chambers, or detector arrays. The dose quality is often assessed using the γ -index pass rate [4]. These protocols present some drawbacks. They are usually time consuming and require to handle cumbersome phantoms. The detector used may have angular dependence [5, 6]. And above all, the γ -index pass rate has been reported to be a poor predictor of the PTV dose, especially when applied to dose distributions measured by detector arrays [7–15].

Alternatives to measuring the dose in a phantom have been proposed. One approach consists in using the logged machine parameters – namely collimator position, gantry angle, dose output, and patient position – to forward calculate or reconstruct the actual dose in the patient’s planning images [16–18]. This approach has the advantage to provide more accurate knowledge of the actual dose distribution in the patient’s anatomy.

In tomotherapy, the fluence is modulated by varying the leaf open times and the jaw positions. During a treatment, gantry rotation speed, couch translation speed, and dose output are constant. Some authors have tested the accuracy and stability of these parameters. Handsfield *et al.* [19] have reported that couch positions and gantry angles, as recorded by the data acquisition system (DAS), corresponded to the planned values throughout several treatment fractions. Lee *et al.* [20] have reported a maximum jaw position error of 0.4 mm after repeating a QA test over nine months. Finally, the dose control system (DCS) is a servomechanism that stabilises the dose output. The output was reported to variate within 0.4% [21].

In contrast, others have reported that the multileaf collimator (MLC) is prone to fluctuations that may significantly impact the quality of the delivered dose [19, 22–

^{*} mathieu.schopfer@chuv.ch

[†] raphael.moeckli@chuv.ch

25]. Which is more, neither the leaves state (open, close, or moving) nor the leaves open times are recorded by the DAS. In tomotherapy, a DQA based on the machine parameters should therefore focus on the MLC verification. Other components can be checked by analysing the DAS log files and through regular machine QA [26]. Daily QA module of Tomotherapy Quality Assurance (TQA) checks daily the health of the couch, the gantry, the jaws, the LINAC (output and alignment), and the MLC [27].

In a previous article, we presented an algorithm to measure the leaf open times (LOTs) from the detector raw data [28]. Here we propose a DQA method for tomotherapy based on the measured LOTs. We demonstrate that the mean relative LOT discrepancy is a handy indicator of the plan's quality. On the Cheese Phantom, we tested the method ability to highlight LOT discrepancies and verified that the measured LOTs predicted accurately the delivered dose. We used Accuray's standalone calculator [29] to calculate the actual dose from the measured LOTs and compared it to ion chamber measurements. On clinical plans, we assessed that the mean relative LOT discrepancy correlated with the dose variation in the PTV. Again, we relied on Accuray's calculator to calculate the expected and actual dose in the PTV. In summary, we present a suitable and convenient DQA method based on the LOTs measurement.

II. MATERIALS AND METHODS

Our department possesses two TomoHDA™ units, hereafter called TOMO1 and TOMO2.

A. Tomotherapy system and standalone dose calculator

The tomotherapy binary MLC is designed to open and close quickly its leaves. The typical leaf travel time is around 10 ms. This allows for a fine modulation of the delivered fluence. Mechanical constraints may cause slight variations of the actual leaf open time. The manufacturer MLC interlock interrupts the irradiation when the delay is larger than 70 ms. At low mean planned LOT, this inaccuracy may lead to non-negligible dose discrepancy [22].

Tomotherapy units uses megavoltage computed tomography (MVCT) detector arrays for IGRT. During any procedure (imaging, treatment, or DQA), the detector collects the fluence per LINAC pulse (300 Hz). These are the detector pulse-by-pulse data or uncompressed raw data. They can be retrieved from the on-board computer (OBC) after the procedure has been delivered.

Accuray's standalone dose calculator can calculate a dose distribution independently of the commissioned treatment planning system (TPS). It needs to be provided with CT images, a tomotherapy procedure delivery plan, and a machine beam model [29]. Its software

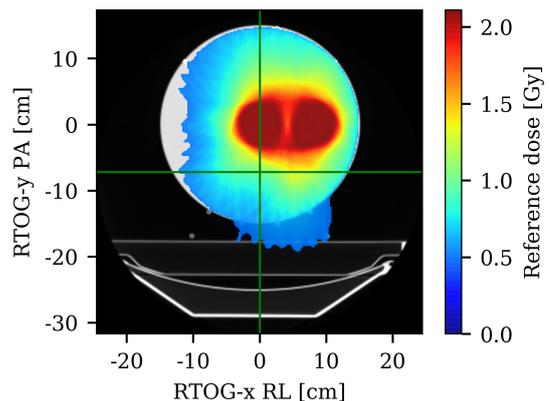


FIG. 1. Cheese Phantom CT images and reference TomoPhant plan dose (TOMO1) overlaid. The green lines correspond to the machine axes.

suite include tools to retrieve images, plans, and machine beam models from the patient archives.

B. Leaf open time measurement

In a previous work, we presented exhaustively an algorithm to measure the leaf open times from the detector raw data [28]. The algorithm performs a Richardson-Lucy deconvolution of the signal. This allows us to get a neat signal per leaf, without inter-leaf contamination. Then, it normalises the signal locally to get rid off the fluctuations due to the patient's attenuation, the jaw output, the leaves output, and the cone shape of the beam profile. Then, the period of time during which a leaf is open corresponds to a high intensity region of the normalised signal. The algorithm thus measures the FWHM of the high intensity regions.

C. Modified TomoPhant plans

To verify the ability of the leaf open time measurement algorithm to detect discrepancies and to verify the accuracy of the dose calculated from the measured LOTs with Accuray's standalone calculator [29], we modified a TomoPhant plan.

TomoPhant is a set of synthetic plans used for routine machine QA in tomotherapy. It allows to measure the delivered dose with ion chambers in the Cheese Phantom to verify its accuracy. We chose the helical plan with a dynamic jaw field of 2.5 cm. These are the commonest treatment settings in our department, because they allow to reach the highest plan quality. This plan has two off-axis cylindrical targets of 6 cm (figure 1). It is slightly modulated, the mean planned LOT being of 173 ms for projection time of 314 ms. This gave us enough mar-

gin to introduce discrepancies without running into non-deliverable LOTs. Which is more, few planned LOTs are in the non-linear regimen of the leaves, i.e. below 30 ms (where a leaf may fail to open) or within 30 ms of the projection time (where a leaf may fail to close between successive projections).

We generated random discrepancies following normal distributions of mean -6% , -4% , -2% , 2% , 4% , and 6% , and a standard deviation of 7% . 7% was the mean standard deviation between planned and actual LOTs observed in the clinical DQA (see section IID). We chose Gaussian distributions because the discrepancy distribution is a bell curve in many clinical instances [28]. No plan was generated with a mean difference of 0% . We modified only the LOTs, *i.e.* a leaf that does not open in the reference plan does not open in any of the modified plans either.

We delivered each plan twice (*i.e.* a total of 14 runs per treatment unit). We collected the detector raw data and measured the delivered dose. We used two A1SI ion chambers – one in each target – located at -0.5 cm and $+8.0$ cm along the right-left axis (RTOG frame of reference). From the detector raw data, we measured the actual LOTs.

Hereafter, “reference” and “expected” LOTs refer to original and modified TomoPhant plans, respectively. For each measurement, we calculated the differences between actual and expected LOTs and the differences between mean and standard deviation. The differences between actual and expected LOTs are caused by the intrinsic MLC inaccuracy. Then, we calculated the expected discrepancies (difference between expected and reference LOTs) and the actual discrepancies (difference between the actual and reference LOTs). Again, we calculated the differences mean and standard deviation. Finally, we calculated the actual dose from the actual LOTs and the difference between calculated and measured doses.

In addition, we calculated the correlation between the mean relative (actual) LOT discrepancy and the dose difference (between measured and reference dose).

D. Clinical pretreatment DQA

To assess the correlation between the mean relative leaf open time discrepancy and the dose deviation in the PTV, we selected randomly 60 clinical treatment plans (30 on each unit). All plans were helical, dynamic jaw deliveries with a 2.5 cm field. These are the commonest treatment settings in our department. All plans had been accepted following the DQA protocol in use in our department. This protocol consists in measuring the dose with the Octavius 729 system (PTW, Freiburg, Germany) and comparing measured and expected dose distributions using a 3D global γ -index in VeriSoft (PTW, Freiburg, Germany), suppressing dose below 10-20 % of expected dose maximum. The plan is accepted if the pass rate (relative number of dose points verifying $\gamma < 1$) is

above 90 %, for passing criteria of 3 %/3 mm. Sometimes, the dose tolerance for dose points below 0.5 Gy was increased by 10 % to 15 %.

We irradiated all plans in air using a static couch DQA procedure with the couch positioned out of the gantry bore. We collected the detector raw data and measured the actual LOTs from them. We calculated the dose distributions from the expected and actual LOTs (named hereafter expected and actual dose) in the patient’s planning images using Accuray’s standalone calculator [29]. Then, for each patient, we calculated the deviation between the expected and actual dose covering 95 % of the PTV (D_{95}). Finally, we calculated the correlation between the mean relative LOT discrepancy and the D_{95} deviation.

For each plan, we calculated the local 3D γ -index in VeriSoft, suppressing points below 10 % of the dose maximum as recommended in the AAPM TG 218 report [30]. VeriSoft does not provide an absolute dose difference threshold. We calculated the pass rates for passing criteria of 2 %/2 mm and 3 %/3 mm. We calculated the correlation between the γ -index pass rates and the absolute dose (D_{95}) deviation in the PTV.

We did not perform the clinical DQA ourselves. We used the dose measured in the Octavius phantom during the actual pretreatment DQA procedure. The γ -index calculation parameters and passing criteria used to clinically accept the plans were different from the ones used in this study. In particular, we used a local γ -index expecting a higher sensitivity [11]. Stasi *et al.* [8] reported a stronger correlation of the dose deviation with the local γ -index pass rate than with the global pass rate. Clinical DQA and in air LOT measurements relate to different runs of the same plan. They were usually performed on different days, but always within a fortnight.

III. RESULTS

A. Modified TomoPhant plans

As mentioned, “reference” and “expected” LOTs refer to original and modified TomoPhant plans, respectively. “actual” LOTs refer to the leaf open times measured from the detector raw data (section IIB).

Figure 2 shows the means and the standard deviations of the relative discrepancies between actual and expected LOTs (actual-expected). The absolute discrepancies means were 2 ms on TOMO1 and 0 ms on TOMO2 (equals in all 14 runs).

Figure 3 shows the means of the expected (expected-reference) and actual (actual-reference) relative LOT discrepancies. The difference between actual and expected values is 1 % on TOMO1 and 0 % on TOMO2 (as in figure 2).

The relative difference between the dose calculated from the actual LOTs and the dose measured with ionisation chambers was of $0.2 \pm 0.3\%$ on TOMO1 and of

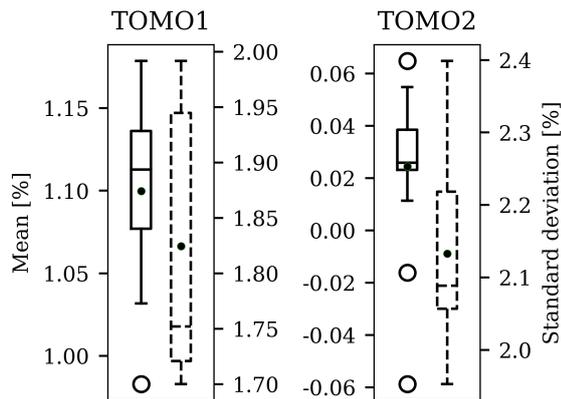


FIG. 2. Means (plain) and standard deviations (dashed) of the relative differences between actual and expected leaf open times ($N = 14$ on both units). The whiskers length equals 1.5 times the interquartile range (IQR).

TABLE I. Correlations between γ -index pass rate and absolute dose deviation in the PTV (D_{95}).

	TOMO1	TOMO2
Tolerances	Corr. (p -value)	Corr. (p -value)
2 %/2 mm	-0.23 (0.045)	0.09 (0.450)
3 %/3 mm	-0.10 (0.396)	0.14 (0.222)

0.1 ± 0.3 % on TOMO2 ($N = 28$, one standard deviation). Figure 4 shows the relative difference between calculated and measured dose.

Finally, the correlation between the mean relative actual LOT discrepancy and the relative dose difference (measured-reference) was of 1.0 on both units (figure 5). In this context, the reference dose is the planned dose of the original (unmodified) TomoPhant plan.

B. Clinical pretreatment DQA

The means of the relative leaf open time discrepancies ranged between -3.0 % and -0.8 % on TOMO1 and -1.1 % and 1.4 % on TOMO2. The deviations of the PTV dose (D_{95}) ranged between -1.6 % and 0.0 % on TOMO1 and -0.2 % and 2.4 % on TOMO2. The correlation between mean relative LOT discrepancy and dose the deviation in the PTV were of 0.76 ($p \approx 10^{-15}$) on TOMO1 and 0.65 ($p \approx 10^{-10}$) on TOMO2 (figure 6).

The correlation between γ -index pass rate for acceptance criteria of 2 %/2 mm and 3 %/3 mm and absolute dose deviation in the PTV (D_{95}) are reported in table I and figure 7.

IV. DISCUSSION

Here we validated a DQA method for tomotherapy based on the LOT measurement. We generated LOT discrepancies in TomoPhant plans and assessed that the algorithm detected them. We checked that the dose calculated from the measured LOTs was correct by comparing it to ion chamber measurements. Finally, we evaluated the impact of LOT discrepancies on clinical plans and showed how the mean relative LOT discrepancy correlates with the dose deviation in the PTV.

The actual (measured) LOTs corresponded well with the planned LOTs (figure 2). We expected a 0 % mean LOT difference, because in the TomoPhant plans most LOTs fall within the linear regime of the leaves. 1 % mean difference corresponded to 2 ms in this case. This is below the LINAC pulse rate (300 Hz). The difference between both units is notable. These results give us again a good confidence in the LOT measurement algorithm (see also [28]). The LOT differences standard deviations (figure 2) gives an estimate of the intrinsic variability of the LOTs (i.e. the real error of the MLC).

The algorithm detected correctly the MLC discrepancies. The mean LOT discrepancy measured in the modified TomoPhant plans correspond to the expected values within 1 % (figure 3).

The dose calculated from the actual LOTs was accurate. The ion chamber measurements corresponded within 1 % to the calculation (figure 4). As Accuray's standalone dose calculator was previously validated [29], an accurate dose measurement implies that the calculation parameters – the LOTs – are accurate.

The mean relative LOT discrepancy is a good indicator of the dose discrepancy (figure 6). On the TomoPhant plans, it was even a perfect indicator (figure 5). It could be perfected by weighting the discrepancies by the jaw output factor, the leaf output factor, and the beam lateral profile. But we wanted to provide an operational value, not to recalculate a fluence map.

A mean relative LOT discrepancy between -3 % and 1 % did not lead to clinically relevant dose deviations in the PTV (figure 6). Out of this range, a lack of PTV coverage or hotspots in the patient's organs at risk (OARs) are to be expected. For that, the dose calculation from the measured LOTs brings a real added value. We used Accuray's standalone calculator, but any independent dose calculation software (e.g. a Monte-Carlo based one) would do the job.

Deshpande *et al.* [25] have also reported a linear relation between the dose deviation in the PTV and the mean relative LOT difference. Though, they found steeper dose deviations than we did. In particular, at a mean LOT difference of 2 %, they reported a dose change of 5 %, whereas we reported a dose change of 1 %. It is not clear if they defined the mean relative LOT discrepancy the same way we did.

The lack of correlation between the γ -index pass rate and the dose to the PTV has already been documented

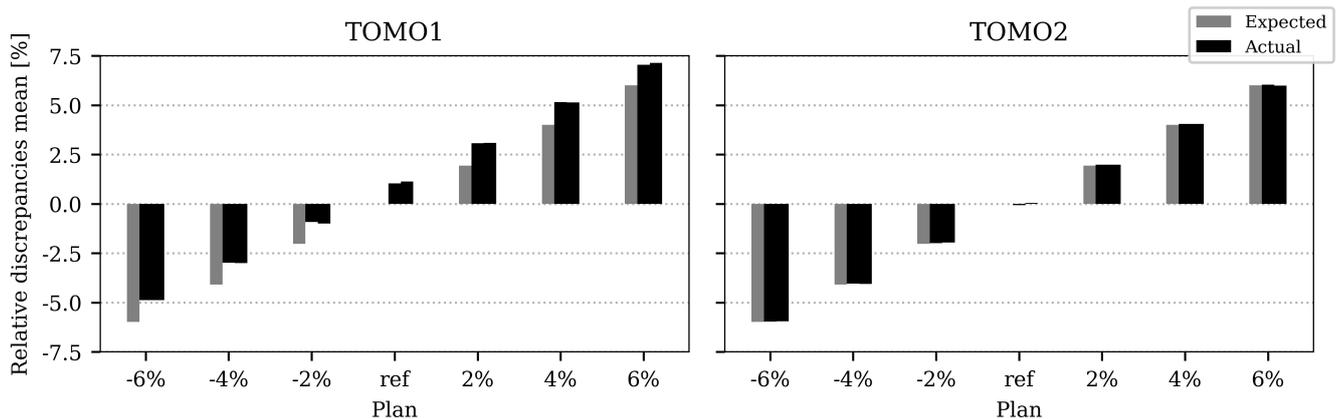


FIG. 3. Means of the expected (black) and actual (grey) relative leaf open time discrepancies per TomoPhant plan. “ref” refers to the original plan. The other values refer to the plans with modified leaf open times. They are two actual values per plan because each measurement was run twice.

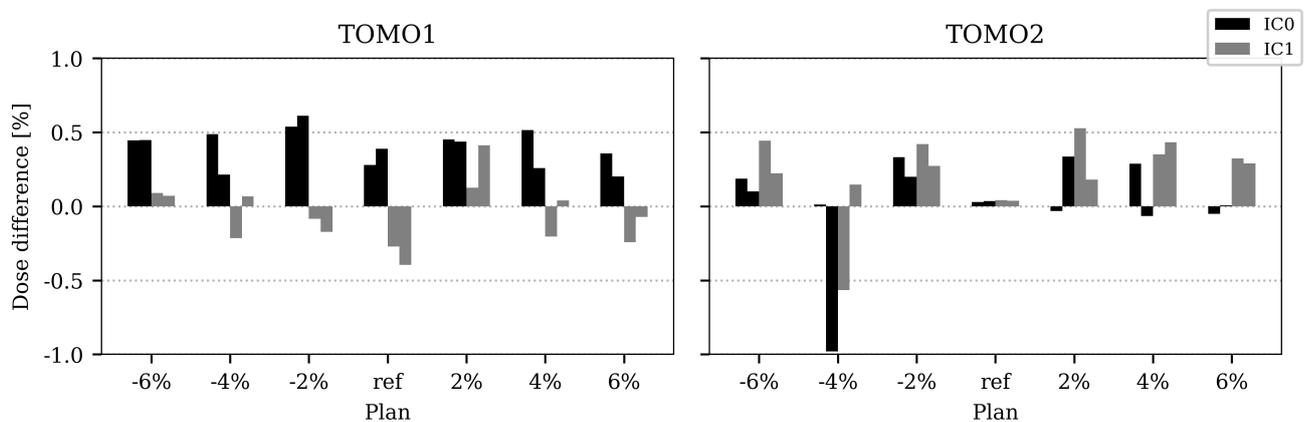


FIG. 4. Relative difference between calculated and measured dose per TomoPhant plan. The dose was calculated from the actual (measured) LOTs. “ref” refers to the original plan. The other values refer to the plans with modified leaf open times. They are two values per plan and chamber because each measurement was run twice. The ion chambers were at -0.5 cm (IC0) and 8.0 cm (IC1) along the left-right axis (RTOG frame of reference) in the Cheese Phantom.

[7–15]. This highlights the lack of uniformity in the applied passing criteria. With the criteria used in the present study, the pass rate of many plans was below 90 % (figure 7), meaning the plans would have been rejected. Yet, all plans were accepted, which was the correct decision regarding the slight dose deviations in the PTV (figure 6). The dose calculation shows that the plans were actually of good quality. This study provides a more clinically relevant DQA tool by relating directly the plan quality indicator – the mean relative LOT discrepancy – to the dose deviation in the PTV.

We previously reported that the LOT measurement algorithm works as well in air as *in vivo* [28]. Thus, the method proposed here would be also valid *in vivo*. In fact, the LOTs of the TomoPhant plans were measured while the phantom was in the beam path. Yet, we do not recommend to drop pretreatment DQA. Suboptimal

use of the planning parameters (the pitch in particular) can lead to non-negligible dose discrepancy [22, 31]. But this makes the DQA method presented here easy to commission. And it allows to monitor the MLC health on the long run. The results on the TomoPhant show that the LOT discrepancy pattern of a specific plan is reproducible. Handsfield *et al.* [19] reached the same conclusion. They measured *in vivo* the mean LOT difference at all fractions of 10 patients’ treatment.

V. CONCLUSION

This study presented a DQA method based on the LOTs measurement and the calculation of the actual dose on the patient’s planning CT. We validated the method by ion chambers measurements. We assessed the corre-

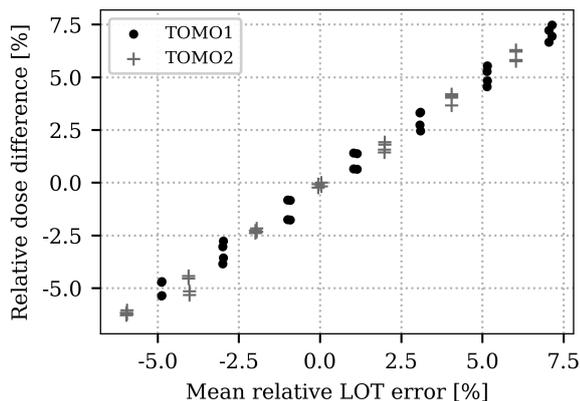


FIG. 5. Mean relative (actual) leaf open time discrepancy and relative difference between measured and reference dose of the TomoPhant plans delivered on TOMO1 (black) and TOMO2 (grey).

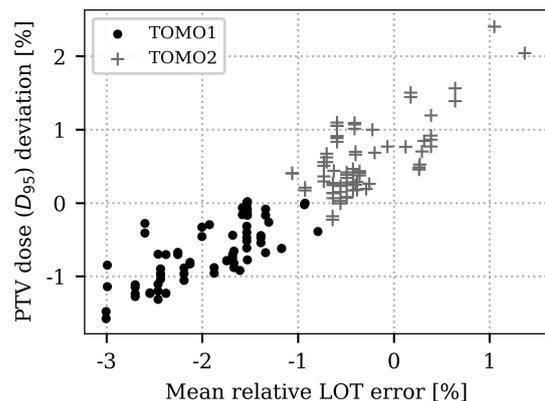


FIG. 6. Mean relative leaf open time discrepancy and deviation between expected and actual PTV dose (D_{95}) of the plans delivered on TOMO1 (black) and TOMO2 (grey).

lation of the LOT discrepancy and the dose deviation in the PTV, thus showing that the mean relative LOT discrepancy is a reliable and clinically relevant indicator of the plan quality.

ACKNOWLEDGMENTS

We wish to thank the medical physicists in our department for performing the DQA and providing the related data. We also wish to thank the RTTs who willingly helped to collect clinical detector data. Finally, we wish to thank Dylan Casey, Edward Chao, and Calvin Maurer (Accuray) for the fruitful discussions and technical support.

This research was funded by Accuray.

-
- [1] Markus Alber, Sara Broggi, Carlos De Wagter, Ines Eichwurz, Per Engström, Claudio Fiorino, Dietmar Georg, Günther Hartmann, Tommy Knöös, Antonio Leal, Hans Marijnissen, Ben Mijnheer, Marta Paiusco, Francisco Sánchez-Doblado, Rainer Schmidt, Milan Tomsej, and Hans Welleweerd, “Guidelines for the verification of IMRT,” *ESTRO* (2008).
- [2] Gary A. Ezzell, James M. Galvin, Daniel Low, Jatinder R. Palta, Isaac Rosen, Michael B. Sharpe, Ping Xia, Ying Xiao, Lei Xing, and Cedric X. Yu, “Guidance document on delivery, treatment planning, and clinical implementation of IMRT: Report of the IMRT subcommittee of the AAPM radiation therapy committee,” *Medical Physics* **30**, 2089–2115 (2003).
- [3] Alan C. Hartford, James M. Galvin, David C. Beyer, Thomas J. Eichler, Geoffrey S. Ibbott, Brian Kavanagh, Christopher J. Schultz, and Seth A. Rosenthal, “American college of radiology (ACR) and american society for radiation oncology (ASTRO) practice guideline for intensity-modulated radiation therapy (IMRT),” *American Journal of Clinical Oncology* **35**, 612–617 (2012).
- [4] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, “A technique for the quantitative evaluation of dose distributions,” *Med Phys* **25**, 656–661 (1998).
- [5] Syam Kumar, Aswathi Cheruparambil, and Anjana Parakkat Thokkayil, “Clinically evaluating directional dependence of 2d seven29 ion-chamber array with different IMRT plans,” *International Journal of Cancer Therapy and Oncology* **3**, 348 (2015).
- [6] Christopher Neilson, Michael Klein, Rob Barnett, and Slav Yartsev, “Delivery quality assurance with ArcCHECK,” *Medical Dosimetry* **38**, 77–80 (2013).
- [7] Heming Zhen, Benjamin E. Nelms, and Wolfgang A. Tomé, “Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA,” *Medical Physics* **38**, 5477–5489 (2011).
- [8] M. Stasi, S. Bresciani, A. Miranti, A. Maggio, V. Sapino, and P. Gabriele, “Pretreatment patient-specific IMRT

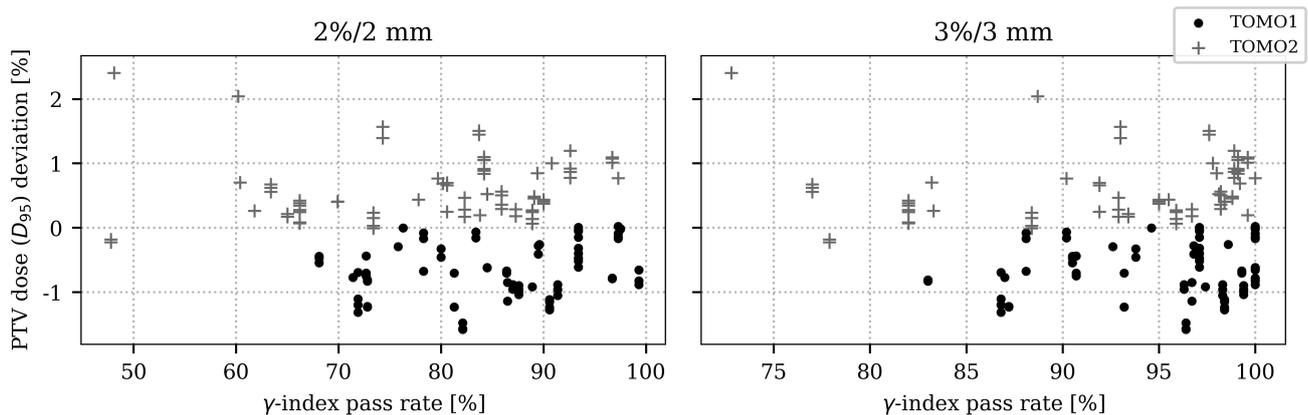


FIG. 7. γ -index pass rate and deviation between expected and actual PTV dose (D_{95}) of the plans delivered on TOMO1 (black) and TOMO2 (grey), for two different sets of passing criteria. A 3D local γ -index was calculated, suppressing low dose below 10 % of the expected dose maximum.

- quality assurance: A correlation study between gamma index and patient clinical dose volume histogram,” *Medical Physics* **39**, 7626–7634 (2012).
- [9] Simon J. Thomas and Ian R. Cowley, “A comparison of four indices for combining distance and dose differences,” *International Journal of Radiation Oncology*Biolog*Physics* **82**, e717–e723 (2012).
- [10] Sara Bresciani, Amalia Di Dia, Angelo Maggio, Claudia Cutaia, Anna Miranti, Erminia Infusino, and Michele Stasi, “Tomotherapy treatment plan quality assurance: The impact of applied criteria on passing rate in gamma index method,” *Medical Physics* **40**, 121711 (2013).
- [11] Benjamin E. Nelms, Maria F. Chan, Geneviève Jarry, Matthieu Lemire, John Lowden, Carnell Hampton, and Vladimir Feygelman, “Evaluating IMRT and VMAT dose accuracy: Practical examples of failure to detect systematic errors when applying a commonly used metric and action levels,” *Medical Physics* **40**, 111722 (2013).
- [12] Stephen F. Kry, Andrea Molineu, James R. Kerns, Austin M. Faught, Jessie Y. Huang, Kiley B. Pulliam, Jackie Tonigan, Paola Alvarez, Francesco Stingo, and David S. Followill, “Institutional patient-specific IMRT QA does not predict unacceptable plan delivery,” *International Journal of Radiation Oncology*Biolog*Physics* **90**, 1195–1201 (2014).
- [13] Elizabeth M. McKenzie, Peter A. Balter, Francesco C. Stingo, Jimmy Jones, David S. Followill, and Stephen F. Kry, “Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans,” *Medical Physics* **41**, 121702 (2014).
- [14] Strahinja Stojadinovic, Luo Ouyang, Xuejun Gu, Arnold Pompoš, Qinan Bao, and Timothy D. Solberg, “Breaking bad IMRT QA practice,” *Journal of Applied Clinical Medical Physics* **16**, 154–165 (2015).
- [15] M. Stasi, S. Bresciani, A. Miranti, M. Poli, A. Di Dia, A. Maggio, E. Delmastro, and P. Gabriele, “EP-1587: Sensitivity and specificity of gamma index method for tomotherapy plans,” *Radiotherapy and Oncology* **119**, S737 (2016).
- [16] Wojciech Osewski, Łukasz Dolla, Michał Radwan, Marta Szlag, Roman Rutkowski, Barbara Smolińska, and Krzysztof Ślosarek, “Clinical examples of 3d dose distribution reconstruction, based on the actual MLC leaves movement, for dynamic treatment techniques,” *Reports of Practical Oncology & Radiotherapy* **19**, 420–427 (2014).
- [17] Neelam Tyagi, Kai Yang, David Gersten, and Di Yan, “A real time dose monitoring and dose reconstruction tool for patient specific VMAT QA and delivery,” *Medical Physics* **39**, 7194–7204 (2012).
- [18] M. Dinesh Kumar, N. Thirumavalavan, D. Venugopal Krishna, and M. Babaiah, “QA of intensity-modulated beams using dynamic MLC log files,” *Journal of Medical Physics* **31**, 36 (2006).
- [19] Lydia L. Handsfield, Ryan Jones, David D. Wilson, Jeffery V. Siebers, Paul W. Read, and Quan Chen, “Phantomless patient-specific TomoTherapy QA via delivery performance monitoring and a secondary monte carlo dose calculation,” *Medical Physics* **41**, 101703 (2014).
- [20] Francis Kar-Ho Lee, Simon Kar-Yiu Chan, and Ricky Ming-Chun Chau, “Dosimetric verification and quality assurance of running-start-stop (RSS) delivery in tomotherapy,” *Journal of Applied Clinical Medical Physics* **16**, 23–29 (2015).
- [21] Zoë R. Moutrie, Craig M. Lancaster, and Litang Yu, “First experiences in using a dose control system on a TomoTherapy hi-art II,” *Journal of Applied Clinical Medical Physics* **16**, 277–284 (2015).
- [22] David C. Westerly, Emilie Soisson, Quan Chen, Katherine Woch, Leah Schubert, Gustavo Olivera, and Thomas R. Mackie, “Treatment planning to improve delivery accuracy and patient throughput in helical tomotherapy,” *International Journal of Radiation Oncology*Biolog*Physics* **74**, 1290–1297 (2009).
- [23] Ke Sheng, Ryan Jones, Wensha Yang, Siddharth Saraiya, Bernard Schneider, Quan Chen, Geoff Sobering, Gustavo Olivera, and Paul Read, “3d dose verification using tomotherapy CT detector array,” *International Journal of Radiation Oncology*Biolog*Physics* **82**, 1013–1020 (2012).

- [24] Shrikant Deshpande, Mark Geurts, Philip Vial, Peter Metcalfe, Mark Lee, and Lois Holloway, “Clinical significance of treatment delivery errors for helical TomoTherapy nasopharyngeal plans – a dosimetric simulation study,” *Physica Medica* **33**, 159–169 (2017).
- [25] Shrikant Deshpande, Aitang Xing, Peter Metcalfe, Lois Holloway, Philip Vial, and Mark Geurts, “Clinical implementation of an exit detector-based dose reconstruction tool for helical tomotherapy delivery quality assurance,” *Medical Physics* **44**, 5457–5466 (2017).
- [26] Katja M. Langen, Niko Papanikolaou, John Balog, Richard Crilly, David Followill, S. Murty Goddu, Walter Grant, Gustavo Olivera, Chester R. Ramsey, and Chengyu Shi, “QA for helical tomotherapy: Report of the AAPM task group 148,” *Medical Physics* **37**, 4817–4853 (2010).
- [27] John Balog, Tim Holmes, and Richard Vaden, “Helical tomotherapy dynamic quality assurance,” *Medical Physics* **33**, 3939–3950 (2006).
- [28] Mathieu Schopfer, François Bochud, Jean Bourhis, and Raphaël Moeckli, “In air and in vivo measurement of the leaf open time in tomotherapy using the on-board detector pulse-by-pulse data,” *Medical Physics* (2019).
- [29] Quan Chen, Weiguo Lu, Yu Chen, Mingli Chen, Douglas Henderson, and Edmond Sterpin, “Validation of GPU based TomoTherapy dose calculation engine,” *Medical Physics* **39**, 1877–1886 (2012).
- [30] Moyed Miften, Arthur Olch, Dimitris Mihailidis, Jean Moran, Todd Pawlicki, Andrea Molineu, Harold Li, Krishni Wijesooriya, Jie Shi, Ping Xia, Nikos Papanikolaou, and Daniel A. Low, “Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM task group no. 218,” *Medical Physics* **45**, e53–e83 (2018).
- [31] Quan Chen, David Westerly, Zhenyu Fang, Ke Sheng, and Yu Chen, “TomoTherapy MLC verification using exit detector data,” *Medical Physics* **39**, 143–152 (2011).