Magphan Phantoms for MR for radiation therapy and quantitative imaging applications.

Smári

Sub-voxel geometric distortion measurements
and critical image quality metrics engineered for
ease of use in the clinical workflow.

The Phantom Laboratory manufactures high-precision phantoms
coupled with Smári image analysis service and innovative custom
solutions for the medical imaging and radiation therapy fields.

[Click to see our latest phantoms and schedule a demo
of our Smári image analysis service.](#)

openPR — A computational tool for CT conversion assessment with proton radiography

Sylvain Deffet^{a)}

Institute of Information and Communication Technologies, Université catholique de Louvain, Louvain-La-Neuve 1348, Belgium

Marie Cohilis and Kevin Souris

Center of Molecular Imaging, Radiotherapy and Oncology, Institut de Recherche Expérimentale et Clinique (IREC), Université catholique de Louvain, Louvain-La-Neuve 1348, Belgium

Koen Salvo

Department of Oncology, Katholieke Universiteit Leuven, Leuven 3000, Belgium

Tom Depuydt

Department of Oncology, Katholieke Universiteit Leuven, Leuven 3000, Belgium

Department of Radiation Oncology, University Hospitals Leuven, Leuven 3000, Belgium

Edmond Sterpin

Center of Molecular Imaging, Radiotherapy and Oncology, Institut de Recherche Expérimentale et Clinique (IREC), Université catholique de Louvain, Louvain-La-Neuve 1348, Belgium

Department of Oncology, Laboratory of Experimental Radiotherapy, Katholieke Universiteit Leuven, Leuven 3000, Belgium

Benoît Macq

Institute of Information and Communication Technologies, Université catholique de Louvain, Louvain-La-Neuve 1348, Belgium

(Received 8 June 2020; revised 20 August 2020; accepted for publication 15 October 2020; published xx xxxx xxxx)

Purpose: One of the main sources of uncertainty in proton therapy is the conversion of the Hounsfield Units of the planning CT to (relative) proton stopping powers. Proton radiography provides range error maps but these can be affected by other sources of errors as well as the CT conversion (e.g., residual misalignment). To better understand and quantify range uncertainty, it is desirable to measure the individual contributions and particularly those associated to the CT conversion.

Methods: A workflow is proposed to carry out an assessment of the CT conversion solely on the basis of proton radiographs of real tissues measured with a multilayer ionization chamber (MLIC). The workflow consists of a series of four stages: (a) CT and proton radiography acquisitions, (b) CT and proton radiography registration in postprocessing, (c) sample-specific validation of the semi-empirical model both used in the registration and to estimate the water equivalent path length (WEPL), and (d) WEPL error estimation. The workflow was applied to a pig head as part of the validation of the CT calibration of the proton therapy center PARTICLE at UZ Leuven, Belgium.

Results: The CT conversion-related uncertainty computed based on the well-established safety margin rule of 1.2 mm + 2.4% were overestimated by 71% on the pig head. However, the range uncertainty was very much underestimated where cavities were encountered by the protons. Excluding areas with cavities, the overestimation of the uncertainty was 500%. A correlation was found between these localized errors and HUs between −1000 and −950, suggesting that the underestimation was not a consequence of an inaccurate conversion but was probably rather due to the resolution of the CT leading to material mixing at interfaces. To reduce these errors, the CT calibration curve was adapted by increasing the HU interval corresponding to the air up to −950.

Conclusion: The application of the workflow as part of the validation of the CT conversion to RSPs showed an overall overestimation of the expected uncertainty. Moreover, the largest WEPL errors were found to be related to the presence of cavities which nevertheless are associated with low WEPL values. This suggests that the use of this workflow on patients or in a generalized study on different types of animal tissues could shed sufficient light on how the contributions to the CT conversion-related uncertainty add up to potentially reduce up to several millimeters the uncertainty estimations taken into account in treatment planning. All the algorithms required to perform the workflow were implemented in the computational tool named openPR which is part of openREGGUI, an open-source image processing platform for adaptive proton therapy. © 2020 American Association of Physicists in Medicine [https://doi.org/10.1002/mp.14571]

Key words: CT calibration, multilayer ionization chamber, particle imaging, proton radiography, range uncertainty

1. INTRODUCTION

Range uncertainty in proton therapy is still a topical problem despite more than one decade of research worldwide aimed at developing means of *in vivo* range verification. In 2012, Paganetti listed in a seminal paper the various sources of uncertainty among which the conversion of the planning computed tomography (CT) image to relative proton stopping powers (RSPs) was referenced as one of the most important.¹

Among the potential solutions to the problem of range uncertainty, proton imaging has benefited from an active body of research.^{2–5} Proton CT, in particular, would directly provide the three-dimensional (3D) map of RSPs, relegating to the past the inaccurate conversion of the planning CT. Nonetheless, the implementation of proton CT in the clinic is challenged by physical and engineering considerations. For example, the nature of the anatomical sites that can be imaged by proton CT is constrained by the maximum energy available. In practice, only head and neck cases could potentially benefit from this technology in most existing facilities. Another approach to reduce range uncertainty is based on recent studies which show that a correction of the planning CT conversion using simple proton radiography could be satisfactory enough in most of the cases.^{6–9} Different proton imaging systems have been developed during the past decade and can be classified into two categories: list-mode and integrated mode.⁵ List-mode proton imaging offers the greatest potential in terms of precision since it provides individual measurements for each of the protons crossing the object. Integrated mode proton imaging, in contrast, integrates the contribution of every proton during the irradiation time of a pencil beam shot. This kind of acquisition is much less complex both in terms of hardware and data throughput which makes it particularly easy to integrate into the clinic.^{10,11} Nonetheless, this comes at the cost of signal degradations resulting from the presence of lateral inhomogeneities along the path of the beam and, to a lesser extent, from multiple Coulomb scattering (MCS).¹² However, deconvolution techniques can be used to increase the spatial resolution and obtain a satisfactory image to assess the range prediction.¹³

In a recent publication, Meijers et al. used a proton radiography technique relying on a multilayer ionization chamber (MLIC) to validate their site-specific HU-RSP conversion, with measurements performed on animal samples.¹⁴ Their results showed that the classic margin rule ($1\text{ mm} + 2.4\%$ when using Monte Carlo.¹) was too conservative with regard to the errors that they observed. To obtain a range error map for each animal sample that was imaged, they implemented a workflow originally proposed by Mumot et al.¹⁵ Namely, the CT scans of the animal samples were imported into the treatment planning system (TPS) and the individual integral depth dose profiles (IDDs) corresponding to the beamlets were simulated. The simulated IDDs were then compared to those measured by the MLIC. Range error maps were eventually obtained by computing for each beamlet, the shift that minimized the squared error between the simulated and the measured IDDs.

The range error maps obtained in the Meijers et al. workflow provide information directly related to the importance of uncertainties in proton therapy. Nonetheless, the workflow could be improved in two ways. First, it does not allow to isolate the range error coming from the conversion of the CT solely. This is nevertheless all the more important today with the development of robust optimization methods, which require an estimate of each individual contribution to the overall range uncertainty. The differences in range observed by Meijers et al. could in fact come from four main sources: CT resolution and artifacts, CT conversion, modeling of the IDDs by the TPS (range mixing and range straggling modeling) and residual setup errors. For example, they observed that increased range errors were well correlated with intersections between materials of clearly different densities. Previous work had already warned about the extreme sensitivity of range error estimation with respect to small residual misalignment and approximations in IDD modeling in the presence of lateral heterogeneities.^{10–13} Second, the workflow relies on simulations carried out by the TPS to estimate the range error. This can be useful to take into account in the range uncertainty assessment, the uncertainty coming from the modeling carried out by the TPS. However, to study the uncertainty associated to the conversion of the CT solely, it is desirable to have a range error estimation method which is independent of any TPS.

In order to overcome this dependence of range error maps on TPS dose calculation, Farace et al. had proposed a simple analytical model validated by Monte Carlo simulations and real measurements.¹⁰ This was the foundation stone of two subsequent developments which provided a solution to the aforementioned limitations. First, a CT-proton radiography registration algorithm was developed to correct the residual alignment errors, in postprocessing.¹⁶ Second, a deconvolution algorithm was developed in order to estimate a WEPL map without resorting to an intermediate stage of CT-based simulations.¹³ The range error could then be calculated no longer as the offset between the predicted IDDs and those measured by the MLIC but as the difference between the actual WEPL and that estimated using the planning CT.

It is of course desirable to integrate these solutions into the workflow proposed by Meijers et al. Therefore, we propose an open-source tool, named openPR, to achieve an end-to-end validation of the calibration curve with limited impacts of residual misalignment and inaccuracies in IDD modeling. However, adding two steps to the workflow to perform the registration and the WEPL estimation with the deconvolution algorithm cannot be done without taking into account that both methods depend on an analytical model which has been validated only on an anthropomorphic phantom. This is actually a source of additional uncertainty. The validity of this model could indeed vary depending on the structures crossed by the protons and on the beam characteristics. To eliminate this source of uncertainty, we propose to add one more step which consists in a Monte Carlo validation specific to the imaged object.

The whole improved workflow was applied in real conditions as part of the validation of the CT calibration of the proton therapy center PARTICLE at UZ Leuven, Belgium.

2. MATERIALS AND METHODS

The workflow detailed in this section aims to assess the difference between the WEPL estimated using the proton radiography data and the WEPL predicted on the basis of the CT with the same HU-RSP conversion as the TPS. In the first instance, we describe the types of data needed as input and the techniques used to acquire them. Then, we detail the entire workflow presented in Fig. 1.

This workflow was applied as part of the validation of the CT calibration of PARTICLE. The animal sample used for this purpose was a pig head. Its maximum thickness was just small enough to be crossed by the protons with the energy of 200 MeV which was chosen for the radiography. The pig head included tissues leading to HUs covering the whole range of the calibration curve (except for the parts corresponding to lung and metal) as well as complex structures such as cavities.

2.A. Proton radiography and CT acquisitions

A single-energy CT scan of the pig head was acquired at 120 kV with the in-room SOMATOM Definition Edge CT

FIG 1. Schematics of the workflow implemented in openPR to assess computed tomography conversion using proton radiography.

scanner (Siemens, Germany), which is the same model as the planning CT scanner. The reconstruction was performed with a 512×512 matrix, $512 \times 512 \text{ mm}^2$ field of view, 1 mm slice thickness, Qr36s\3 reconstruction kernel (according to the clinically applicable head scan protocol), iMAR artifact correction for dental filling and beam hardening correction.

The proton radiograph was acquired with the Giraffe MLIC (IBA, Belgium). This device is composed of 180 chambers, each one having a water equivalent thickness of 1.8 mm. It therefore covers the range of protons having an energy up to 230 MeV. However, the nominal energy used in our application was only 200 MeV because the last channels of the Giraffe available at PARTICLE were not calibrated. The MLIC was placed at the proton exit side of the pig head, as shown in Fig. 2. The ProteusOne delivery system (IBA, Belgium) was operated to deliver a map of spots uniformly spaced by 5 mm according to the method established by Farace et al.¹⁷ for a ProteusPlus (IBA, Belgium), which we had to slightly adapt to make it compatible with the ProteusOne. Indeed, with a ProteusPlus, the beamlets are sent according to a scheme known in advance. This makes the proton radiograph easy to obtain by collecting the IDD's measured by the Giraffe for each beamlet sent at a specified location. However, the beam produced by the S2C2 synchrocyclotron of the ProteusOne system is pulsed, and the dose delivered by each individual pulse cannot be accurately predicted. Therefore, the blind golfer algorithm is used to repeatedly deliver a fraction of the planned dose for each spot, measuring every time the delivered dose and calculating the remaining dose to be delivered. This intrinsic re-scanning delivery would make the analysis of the proton radiography acquisitions more complex. As it is not crucial to know the absolute dose actually delivered in our imaging application, the blind golfer delivery mode was disabled during our proton radiography acquisition.

The pig head was roughly positioned at isocenter using the laser system after the CT scan and before the proton

FIG 2. Experimental setup. The multilayer ionization chamber was placed on a lift table and aligned with the laser system.

radiography. Accurate alignment was not required since residual misalignment can be corrected in postprocessing using a CT-proton radiography registration algorithm.¹⁶ A single direction of acquisition was considered with the gantry positioned at 90°, which gave a radiograph along the lateral direction. Despite a large entrance diameter of 12 cm, the Giraffe's field of view is actually smaller due to internal scattering and beam divergence. In a previous publication, the effective field of view was determined to be only $45 \times 45 \text{ mm}^2$ for a similar experimental setup.¹⁰ Therefore, a whole radiography required several acquisitions between each of which the object was translated. These translations were performed by moving the couch a distance equal to the field of view. The total acquisition time was approximately 10 min, which is twice as fast as the time mentioned by Farace et al. for a similar radiograph.¹⁰ This reduction in the acquisition time was achieved thanks to the update of the interface provided the manufacturer of the proton delivery system.

2.B. Computed tomography calibration

In the RayStation (RaySearch Laboratories, Sweden) TPS used at PARTICLE, the input needed for CT conversion is a table linking HUs to mass densities. Internally, the TPS interpolates the mass density corresponding to each HU of the CT scan and converts it into one of the materials of its database using a nearest neighbor approach. The stopping power of each voxel is finally computed by using both its estimated mass density and its associated elemental composition.

In order to get the required HU-to-density table, Gammex inserts (Gammex Inc, Middleton, USA) were scanned in each one of the three planning CT scanners of PARTICLE, including the in-room one. Average HU values were computed over various protocols and phantoms (a Gammex phantom and two homemade phantoms of a thorax and a head). The differences between measured CT numbers of each scanner were deemed negligible, and only one calibration curve was calculated for all three scanners. To this end, the recommended stoichiometric calibration was used.^{18, 19}

2.C. Computed tomography conversion assessment

2.C.1. Water equivalent path length error map

In the workflow proposed by Meijers et al., the WEPL error is estimated by calculating the shift to be applied to the simulated IDD to minimize the quadratic error between the simulated and the measured IDD.¹⁴ This method is, however, not optimal because it does not use the fact that a measured IDD contains information on the WEPL not only at the spot position but also in its vicinity. In a recent publication,¹³ we developed a deconvolution algorithm which takes advantage of this phenomenon called range mixing, to estimate the WEPL map with a higher resolution than that of the proton radiograph. In other words, this method estimates not only

the WEPL at the center of the beamlets but also between them. This can be achieved because each IDD is a weighted sum of pristine Bragg curves, the ranges of which are directly related to the WEPLs located within the cross section of the overlapping beamlets. However, the WEPL estimation problem is particularly ill-posed because the cross-sectional profile of the beamlets is Gaussian and their overlaps are limited. To solve it, an assumption on the sparsity of the WEPL map was used as a regularization. In addition, a variant was proposed to optionally take advantage of the prior information of the planning CT, in a robust manner to further improve the result of the deconvolution. This particular variant was used in this paper. In the present case, the WEPL map was estimated with a resolution equivalent to that of the CT, namely $1 \times 1 \text{ mm}^2$, on the basis of proton radiography data acquired with a 5 mm spot spacing.

In order to establish the error map between the estimated WEPL and that which would be calculated by the TPS, the two-step conversion method of our TPS, namely RayStation 8.0 (RaySearch, Sweden), was implemented in openPR. To reduce the delay between the CT acquisition and the proton radiography, the alignment between the CT and the proton radiograph was not performed using the in-room x-ray alignment system but rather after the acquisition, in postprocessing, with a specific registration algorithm.¹⁶ Moreover, we observed that this was more accurate than a conventional kV-kV alignment, as already suggested by the literature.¹⁶

2.C.2. Sample-specific validation with Monte Carlo simulations

Although the tools mentioned above have been evaluated in their respective publications,^{10,13,16} they have limitations related to the imaged object. All these methods are based on a model according to which the proton beamlets propagate in straight lines, parallel to each other, with negligible scattering. These methods assume that a measured IDD, named IDD_m consists of a convolution of a Gaussian kernel G having a standard deviation equal to the spot size with shifted versions of a pristine Bragg curve called IDD_{ref} which corresponds to the IDD that would be measured without the phantom through the beam path:

$$IDD_m(x, y, z) \approx G(x, y) * IDD_{ref}(z + WEPL(x, y)) \quad (1)$$

where z refers to the depth axis of the IDD and (x, y) are the spatial coordinates in the plane perpendicular to the beam.

In this paper, IDD_{ref} was obtained by measuring an IDD under the same condition as for a proton radiography but without any phantom. Therefore, Eq. (1) is referred to as a semi-empirical model throughout this paper.

This approximation was validated on several phantoms, including an anthropomorphic head phantom, against both proton radiography measurements and Monte Carlo simulations performed by the TPS.¹⁰ However, it cannot be excluded that it may fail for some anatomical structures. In particular, the pig head imaged at PARTICLE exceeds both in terms of

structural complexity and thickness the phantoms considered in the past.

In order to carry out an accurate analysis of the WEPL error, it is necessary to validate that the semi-empirical model expressed in Eq. (1) and used by the deconvolution and the registration algorithms is correct. To do this, we propose to systematically perform an object-specific validation of the analytical model by Monte Carlo simulations. Indeed, these are nowadays often considered as the gold standard in terms of simulations accuracy, especially in complex geometries such as the presently used pig head. We thus achieved a validation of our semi-empirical method independently of the TPS with a fast Monte Carlo simulation tool named MCsquare²⁰ which is part of the openREGGUI platform on which openPR is based.

First, MCsquare was commissioned for PARTICLE. The obtained beam model was evaluated using the in-water validation measurements previously obtained during the validation of the TPS beam data library. Then, the HU-to-SP(R) conversion of the TPS was adapted so as to get the same SP in MCsquare, which uses a different material assignment scheme.

To assess the accuracy of the analytical model, the MLIC was modeled by a water box added in the CT image, at the exit side of the patient. In order to mimic an acquisition with the Giraffe and in particular to take into account its limited field of view, a treatment plan with a specific isocenter was automatically created by openPR for each position at which the object was placed. On the other hand, it was not necessary to create a specific CT per phantom position, in which the MLIC modeled by a water box would be centered at the isocenter. Indeed, the field of view of the MLIC was previously determined as the entrance region ensuring in particular that few protons escape from the detector by beam divergence or by internal scattering. There was therefore no difference in providing the Monte Carlo tool with a CT containing a much larger detector than the actual MLIC to measure IDD_s regardless of the isocenter of the treatment plan. Finally, the dose maps were exported separately for each beamlet so that the IDD_s could later be reconstructed.

Since the aim of Monte Carlo simulations is to evaluate the accuracy of the analytical model and the deconvolution method also depends on this model, it cannot be used to estimate the WEPL in such context. Instead, the range error, named *RE*, was estimated by computing the displacement that minimizes the quadratic error between two dose deposition profiles, named *IDD*₁ and *IDD*₂:

$$RE = \left| \operatorname{argmin}_s \sum_z (IDD_1(z+s) - IDD_2(z))^2 \right| \quad (2)$$

2.C.3. Range uncertainty analysis

In his seminal paper on range uncertainties in proton therapy, Paganetti evaluated the range uncertainty to 1.2 mm + 2.4% with Monte Carlo dose calculation and to 1.2 mm + 4.6% with less accurate dose calculation

algorithms, such as the pencil beam superposition method.¹ The various contributions to the range uncertainty listed by Paganetti are now broadly accepted as a guide for beam-specific PTV margins, or to set error parameters in robust optimization. Therefore, it is highly relevant to set up a methodology to validate those uncertainties using proton radiography. However, this cannot be done immediately by taking the percentage of errors lower than the margin of 1.2 mm + 2.4% calculated by Paganetti.

First, we note that Paganetti expressed his estimate of the proton range uncertainty as a 1.5 standard deviation interval. This means that the amount of range errors confined in this interval in fact depend on the probabilistic distribution of the error. For a normal distribution, this approximately corresponds to a 87% confidence interval.

A second element to consider to assess the margin defined by Paganetti is that it takes into account patient setup errors and uncertainties in the modeling of range degradation. However, those should be considered as negligible in the present study thanks to the accurate registration in postprocessing and to the validation of the semi-empirical model against Monte Carlo. Estimating the remaining range uncertainty the same way as Paganetti did, we obtain a 87% interval given by 0.6 mm + 2.4%. From there, we can assess the relevance of the margins used so far by comparing the percentage of range errors lower than 0.6 mm + 2.4% to the expected value, namely 87%.

2.D. openPR

All the algorithms involved in the workflow described above were implemented in openPR which was developed as a plugin in openREGGUI, an image processing open-source platform for adaptive proton therapy. The code is mainly written in Matlab (MathWorks, USA) and is available on the openREGGUI main repository (openreggui.org). The user can interact with the data and call the methods either via a graphical interface or through the scripting functionality available in openREGGUI. The features offered by the graphical interface also cover data import, data and metrics visualization, region of interest drawing, etc. In addition, several options were made available for the various methods. For example, a registration can be performed either with one proton radiograph or with two orthogonal proton radiographs. Parameters such as the spot spacing and the number of spots can also be adapted by the user.

3. RESULTS

3.A. Proton radiography acquisition

The proton radiograph of the pig head was acquired with a spot spacing of 5 mm, which means that an IDD was measured for each beamlet sent every 5 mm. The deconvolution algorithm was run after co-registration with the CT to estimate the WEPL map presented in Fig. 3. It was computed with a spatial resolution of 1 × 1 mm² so that it could be

compared in this section with that estimated from the CT converted to RSPs.

3.B. Sample-specific validation with Monte Carlo simulations

By applying CT-proton radiography registration in post-processing, we could reduce the impact of setup errors and therefore remove them from the several possible contributions to the range error. Consequently, the remaining major sources of errors were, in accordance with the table established by Paganetti,¹ an inaccurate conversion of HUs to RSPs, an insufficient resolution of the image acquired by the CT or CT artifacts. This assumes that the semi-empirical model used by most of the algorithms implemented in openPR was accurate. Although this model was validated both experimentally and using Monte Carlo simulations in a previous publication,¹⁰ the absence of a scattering model could be prejudicial in the current case where many cavities and heterogeneous structures are present. This example demonstrates the need for the systematic validation step based on Monte Carlo simulations that we added to the workflow.

The range error depicted in Fig. 4 shows overall an excellent agreement between Monte Carlo and semi-empirical simulations except in the upper left corner. Some Bragg curves selected at arbitrary locations are also shown. They correspond to the centers of areas A, B, and C of Fig. 5, which will be discussed in the next section. An IDD representative of the area where the error is greater was also selected at the location named E. An explanation for these errors is that the semi-empirical method is based on a constant spot size, typically defined at isocenter, and does not account for scattering properties. Bone tissues, or air gaps between tissues, for example, can thus lead to important discrepancies between Monte Carlo and semi-empirical simulations.

FIG 3. Water equivalent path length map of the pig head resulting from the deconvolution of the proton radiograph.

The map of range differences between semi-empirical simulations and Monte Carlo simulations was used to draw a region of interest in which 90% of the error is below 0.5 mm, as shown on Fig. 4.

3.C. Range uncertainty analysis

Figure 5 shows the absolute difference between the WEPL predicted using the calibration curve of our TPS and the WEPL obtained via a sparse deconvolution of the proton radiography data, after registration with the CT. The error appears to greatly vary depending on the location. The biggest discrepancies are in the ears. However, they are outside the region of interest drawn in the previous section. In addition, these organs may have moved between CT and proton radiography acquisitions, which is why we will not analyze them. Other important errors take place in cavities, and in a particularly remarkable way in the sinus cavities.

Several areas have been drawn on Fig. 5. They correspond to distinct levels of errors but also to different underlying tissues. Area A has the lowest error and does not contain any cavity. Area C, in contrast, presents the highest error and corresponds to the sinus cavities. Area B has moderate errors and its underlying tissues are made out of cavities, bones, and soft tissues. Area D is the region of interest determined following the sample-specific validation of the analytical model by Monte Carlo simulations.

Table I lists in its first column, the percentage of WEPL errors (PWE) which are lower than the range uncertainty of $0.6 \text{ mm} + 2.4\%$ calculated according to Paganetti's definition. Except for area C, the margin was clearly overestimated with more than 91.1% of the errors lower than the margin. Region C which corresponds to the sinus cavities constitutes a very problematic particular case, which also appears clearly on the error map shown in Fig. 5. The validation previously carried out with Monte Carlo simulations makes it possible to rule out any modeling error. We believe that the cause of this discrepancy is to be sought at the level of the anatomical structures present within these cavities. These have many transitions between thin cartilages and hollow regions where CT resolution may be problematic. In particular, these cavities present HUs with values lying between -1000 and -950. These are intermediate values between air and pulmonary tissues. Consequently, it was decided to extend the range of HUs corresponding to air, that is. having an RSP of 0, up to $\text{HU} = -950$. This strategy is suitable for head and neck cases for which this curve was established. However, it should be re-evaluated with suitable phantom measurements if low density tissues were to be traversed, for example, using a four-dimensional (4D) lung phantom.²¹

The second column of Table I lists the results obtained after the CT conversion is adapted. It did not only have a positive impact in the areas with cavities where 10.6% of additional pixels were found in the margin but also in all the other regions with a general improvement of a few percent.

In their paper, Meijers et al. calculated the ratios of the range error and the uncertainty margin. They found out that

FIG 4. Range difference between semi-empirical simulations and Monte Carlo simulations. 90% of the values inside the region of interest are lower than 0.5 mm. The plots labeled with the letters A, B, C, and E refer to the points on the left image labeled with the same letter.

FIG 5. Absolute error between the water equivalent path length (WEPL) predicted using the calibration curve of our treatment planning system and the WEPL obtained via a sparse deconvolution of the proton radiography data.

1.5 standard deviation of the values did not exceed 75%. This actually means that the margin was overestimated by at least 33.3% for approximately 87% of the spots (assuming a normal distribution which seems valid with regard to their

TABLE I. Percentage of WEPL errors (PWE) confined in a 0.6 mm + 2.4% margin in areas A, B, C, and D of Fig. 5 before and after that the calibration curve is adapted.

Area	PWE before correction	PWE after correction
A	99.8%	100.0%
B	91.1%	99.3%
C	35.9%	46.5%
D	95.4%	97.9%

figure). For the sake of comparison with Meijers' results, we did a similar experiment. However, we considered a 90% confidence interval rather than a 1.5 standard deviation interval, because the number of errors contained in the 1.5 standard deviation interval depends on the shape of the error distribution. It would be 87% for a normal distribution but it is different for non-perfectly Gaussian distribution. The percentage of WEPL errors lower than the margin is plotted on Fig. 6 for various margin sizes ranging from 0.6 mm + 0% to 0.6 mm + 5%. In the region of interest D which includes area A, B, and C, this value is 90% for a margin of 0.6 mm + 1.4% which means that the part of the initial margin of 0.6 mm + 2.4% which depends on dose computation was

actually overestimated by 71% for 90% of the pixels, or equivalently could be decreased by 42%, which is similar to Meijers conclusions.¹⁴ In the region A which does not have any cavity, the margin corresponding to a 90% confidence interval is 0.6 mm + 0.4%. This means that the part of the initial margin of 0.6 mm + 2.4% which depends on dose computation was overestimated by 500% for 90% of the pixels, or equivalently could be decreased by 83%. In the case where cavities were present, calculating the underestimation does not seem adequate. We observe, for example, that a margin much greater than 0.6 mm + 5% would be necessary in area A while a margin of only 0.6 mm + 1.5% would be sufficient in area B which nevertheless has a higher WEPL but much smaller cavities.

Finally, Fig. 6(b) shows the pixels which are confined in a margin of 0.6 mm + 2.4% on the one hand and 0.6 mm + 1.4% on the other hand.

4. DISCUSSION

The use of treatment margins, or more recently robust optimization methods, is necessary to ensure the robustness of proton therapy plans against treatment errors. A major source of uncertainty is the conversion of CT images to RSPs.²² Proton computed tomography would make it possible to obtain a map of RSPs directly and is naturally the ideal modality, but its clinical use in the short term remains an unresolved challenge. In contrast, proton radiography was recently used to validate the conversion scheme of the proton therapy facility of University Medical Center Groningen (Groningen, The Netherlands).¹⁴ The workflow proposed by the authors made it possible to reduce the range uncertainty, not only by quantifying the error made by the TPS on the RSP computation but also by allowing the HU-RSP conversion to be accordingly adapted.

The workflow detailed in the paper of Meijers et al. is based on the comparison of measurements acquired with an MLIC with simulations generated by the TPS on the basis of a CT of the imaged object.¹⁴ The output is a range error map intended to assess the CT conversion. Nonetheless, the range difference can come from several sources: CT resolution and artifacts, CT conversion, modeling of the IDD by the TPS (range mixing and range straggling modeling) and residual setup errors. To assess solely the conversion of the CT, it was therefore desirable to adapt this workflow to make it on the one hand independent of the TPS dose computation, and on the other hand robust to residual setup errors. In the present paper, we proposed to include a previously published CT-proton radiography registration method for correcting residual misalignment and a deconvolution method for estimating a high-resolution WEPL map from the low spatial resolution proton radiography data, without relying on a TPS. In addition, we showed the importance of adding to the workflow a stage of Monte Carlo simulations in order to validate the use of the tools relying on a semi-empirical model. We have shown that certain particular anatomical configurations such as the ears of the pig head studied throughout this paper,

could cause a significant divergence of the beam which is not modeled by analytical simulations. It was thus necessary to draw a region of interest based on this validation.

Our workflow was applied to a pig head in the context of the validation of the CT conversion scheme at PARTICLE. According to the various contributions to the range uncertainty listed by Paganetti,¹ the overall uncertainty was estimated to be 0.6 mm + 2.4%. Large variations in the WEPL were observed to be related to the presence of cavities. Where no cavity was traversed, the WEPL error was way below the expected value. In contrast, only 35.9% of the WEPL errors were within the margin at the exit of the sinus cavities, using the initial calibration curve obtained using the stoichiometric method. The agreement between Monte Carlo simulations and those relying on the semi-empirical used by openPR confirmed that these discrepancies were not related to the modeling of the dose deposition. The values of the HUs responsible for the errors were between -1000 and -950. These do not correspond to biological tissues but rather to average values between these and the surrounding air. It therefore seems that the error was not really a consequence of an inaccurate conversion but rather of a CT resolution issue or of CT artifacts. We therefore proposed a correction procedure consisting in forcing to 0 the RSP of all HUs between -1000 and -950. No other correlation between the error and specific tissues was found.

After correction of the calibration curve, we obtained values in line with the results presented by Meijers et al. regarding the overestimation of the margins, except for the sinus cavities. The overestimation of the margin was 71% for 90% of the region of interest. A margin of 0.6 mm + 1.4% would actually have been enough to guarantee 90% of the WEPL to be confined in the margin. Moreover, the use of proton radiography showed that the range uncertainty can vary considerably from one area to another. For the pig head considered in this study, the classical rules were underestimated in the cavities while they were overestimated by 500% everywhere else. The underestimation in the pig cavities shows a potential limitation of the theoretical uncertainty model used in this paper since it seems that the main source of uncertainties in this case was related to only a small part of the conversion curve to RSPs. Nevertheless, this should not have a major impact in the clinic where such cavities are typically avoided. Moreover, the content of the pig's cavities might have slightly changed between the acquisition of the CT and the proton radiography since we unfortunately do not have a CT acquired after the proton radiography to rule out this possibility.

Those results seem to confirm that the validation of CT conversions on real biological tissues could potentially improve up to several millimeters the way that the uncertainties are accounted for in treatment planning. A generalized study on different types of animal tissues and across proton therapy centers would make it possible to determine and better understand the existence of correlations between WEPL errors and underlying tissues. In addition, a specific use per patient should be possible thanks to the low dose associated

FIG 6. (a) Relationship between the part of the margin dependent on the dose computation and the percentage of water equivalent path length (WEPL) errors confined in this margin, for a pig head. (b) WEPL errors confined in the initial margin of $0.6 \text{ mm} + 2.4\%$ on the one hand and in the margin of $0.6 \text{ mm} + 1.4\%$ corresponding to a 90% confidence interval, on the other hand.

with the proton radiography. Estimations performed with a treatment planning system for a similar experimental setup showed that the dose delivered with such a method is less than 1 cGyE which is comparable to the one delivered by megavoltage portal imaging.¹⁰ The acquisition time is mainly determined by the amount of couch movements required to image the whole area. With the delivery system used in this study, approximately 20 seconds were necessary for each acquisition of $45 \times 45 \text{ mm}^2$ equivalent to the field of view of the detector. The total acquisition time could be further reduced by focusing on a region of interest or by using a detector with a larger field of view.¹²

Each step of the workflow presented in this paper was experimentally validated on an animal and also in the reference publications, on a synthetic head phantom. This anatomical site is particularly well suited to proton radiography because it can be entirely crossed by protons with moderate energy, whatever the direction of acquisition. The main acquisition parameters were a spot size between 3 mm (typical size for a 210 MeV beam emitted by the ProteusPlus system used in the referenced papers) and 3.7 mm (typical size for a 200 MeV beam emitted by the ProteusOne system used in this paper), a machine output between 0.01 and 0.1 MU, an energy between 200 and 210 MeV, an air gap between the

detector and the object of 30 cm maximum and a spot spacing of 5 mm. Each step of the workflow would of course have to be revalidated if other parameters or phantoms were to be used.

All the tools necessary to carry out the workflow presented in this paper are available on the openREGGUI platform under an open-source license. In addition, the use cases of openPR can be extended to any specific need of the user thanks to the scripting tool offered by openREGGUI.

5. CONCLUSIONS

In this paper, we developed a workflow to achieve an end-to-end validation of the CT conversion to RSPs in proton therapy using range probe measurements. The proposed method has two advantages over those already published: it mitigates residual misalignment in postprocessing as well as potential modeling errors in the dose deposition profiles. The workflow was applied to a pig head as part of the validation of the CT calibration of the proton therapy center PARTICLE at UZ Leuven, Belgium. The results showed that the CT conversion-related uncertainty computed based on the usual Paganetti's rule were overestimated by 71%. Excluding areas with cavities, the overestimation was 500% which showed

that a major contribution to the uncertainty was a localized problem related to a specific type of tissue. This shows that the validation of CT conversion in a generalized study on different types of animal tissues could potentially reduce the range uncertainties generally accounted for in treatment planning by several millimeters. All the tools necessary to carry out the proposed workflow were made available open-source.

ACKNOWLEDGMENTS

The authors thank the IBA site engineers at PARTICLE for their technical support. Sylvain Deffet was supported in the confines of this study by the Fonds européen de développement régional (FEDER) and the Walloon region in the framework of the operational program Wallonie-2020.EU. Marie Cohilis is supported by the Télévie Grant from the Belgian Fonds National pour la Recherche Scientifique F.R.S-FNRS (Grant No. 7450517F). Kevin Souris is funded by the Walloon region (MECATECH/BIOWIN, grant number 8090).

^{a)}Author to whom correspondence should be addressed. Electronic mail: sylvain.deffet@uclouvain.be.

REFERENCES

1. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol* 2012;57:R99.
2. Knopf A, Lomax A. In vivo proton range verification: a review. *Phys Med Biol*. 2013;58:R131.
3. Poludniowski G, Allinson NM, Evans PM. Proton radiography and tomography with application to proton therapy. *Br J Radiol*. 2015;88:20150134.
4. Parodi K, Polf JC. In vivo range verification in particle therapy. *Med Phys*. 2018;45:e1036–e1050.
5. Johnson R. Review of medical radiography and tomography with proton beams. *Rep Prog Phys*. 2018;81:016701.
6. Schneider U, Pedroni E. Proton radiography as a tool for quality control in proton therapy. *Med Phys*. 1995;22:353–363.
7. Doolan PJ, Testa M, Sharp G, Bentefour EH, Royle G, Lu H-M. Patient-specific stopping power calibration for proton therapy planning based on single-detector proton radiography. *Phys Med Biol*. 2015;60:1901.
8. Collins-Fekete C-A, Brousmiche S, Hansen DC, Beaulieu L, Seco J. Pre-treatment patient-specific stopping power by combining list-mode proton radiography and x-ray CT. *Phys Med Biol*. 2017;62:6836.
9. Krah N, Patera V, Rit S, Schiavi A, Rinaldi I. Regularised patient-specific stopping power calibration for proton therapy planning based on proton radiographic images. *Phys Med Biol*. 2019;64:065008.
10. Farace P, Righetto R, Deffet S, Meijers A, Stappen FV. Technical note: a direct ray-tracing method to compute integral depth dose in pencil beam proton radiography with a multilayer ionization chamber. *Med Phys*. 2016;43:6405–6412.
11. Krah N, Marzi LD, Patriarca A, Pittá G, Rinaldi I. Proton radiography with a commercial range telescope detector using dedicated post processing methods. *Phys Med Biol*. 2018;63:205016.
12. Deffet S. Proton radiography to reduce range uncertainty in proton therapy. Université catholique de Louvain. 2018.
13. Deffet S, Farace P, Macq B. Sparse deconvolution of proton radiography data to estimate water equivalent thickness maps. *Med Phys*. 2020;47:509–517.
14. Meijers A, Free J, Wagenaar D, et al. Validation of the proton range accuracy and optimization of CT calibration curves utilizing range probing. *Phys Med Biol*. 2020;65:03NT02.
15. Mumot M, Algranati C, Hartmann M, Schippers JM, Hug E, Lomax AJ. Proton range verification using a range probe: definition of concept and initial analysis. *Phys Med Biol*. 2010;55:4771.
16. Deffet S, Macq B, Righetto R, Stappen FV, Farace P. Registration of pencil beam proton radiography data with X-ray CT. *Med Phys*. 2017;44:5393–5401.
17. Farace P, Righetto R, Meijers A. Pencil beam proton radiography using a multilayer ionization chamber. *Phys Med Biol*. 2016;61:4078.
18. Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol*. 1996;41:111.
19. Ainsley CG, Yeager CM. Practical considerations in the calibration of CT scanners for proton therapy. *J Appl Clin Med Phys*. 2014;15:202–220.
20. Souris K, Lee J, Sterpin E. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. *Med Phys*. 2016;43:1700–1712.
21. Meijers A, Seller OC, Free J, et al. Assessment of range uncertainty in lung-like tissue using a porcine lung phantom and proton radiography. *Phys Med Biol*. 2020;65:155014.
22. Bär E, Lalonde A, Royle G, Lu H-M, Bouchard H. The potential of dual-energy CT to reduce proton beam range uncertainties. *Med Phys*. 2017;44:2332–2344.