The effect of surgical titanium rods on proton therapy delivered for cervical bone tumors: experimental validation using an anthropomorphic phantom

2014 Phys. Med. Biol. 59 7181

(http://iopscience.iop.org/0031-9155/59/23/7181)

IOP Publishing | Institute of Physics and Engineering in Medicine Physics in Medicine & Biology

Phys. Med. Biol. **59** (2014) 7181–7194 doi:10.1088/0031-9155/59/23/7181

The effect of surgical titanium rods on proton therapy delivered for cervical bone tumors: experimental validation using an anthropomorphic phantom

Isabelle Dietlicher[1], Margherita Casiraghi[1], Carmen Ares[1], Alessandra Bolsi[1], Damien C Weber[1,2], Antony J Lomax[1,3] and Francesca Albertini[1]

[1] Center for Proton Therapy, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland
[2] Department of Radio-Oncology, Bern University Hospital, University of Bern, Switzerland
[3] Department of Physics, Swiss Institute of Technology (ETH), Zurich, Switzerland

E-mail: francesca.albertini@psi.ch

Received 13 May 2014, revised 24 August 2014
Accepted for publication 16 September 2014
Published 4 November 2014

Abstract

To investigate the effect of metal implants in proton radiotherapy, dose distributions of different, clinically relevant treatment plans have been measured in an anthropomorphic phantom and compared to treatment planning predictions. The anthropomorphic phantom, which is sliced into four segments in the cranio-caudal direction, is composed of tissue equivalent materials and contains a titanium implant in a vertebral body in the cervical region. GafChromic® films were laid between the different segments to measure the 2D delivered dose. Three different four-field plans have then been applied: a Single-Field-Uniform-Dose (SFUD) plan, both with and without artifact correction implemented, and an Intensity-Modulated-Proton-Therapy (IMPT) plan with the artifacts corrected. For corrections, the artifacts were manually outlined and the Hounsfield Units manually set to an average value for soft tissue. Results show a surprisingly good agreement between prescribed and delivered dose distributions when artifacts have been corrected, with > 97% and 98% of points fulfilling the gamma criterion of 3%/3 mm for both SFUD and the IMPT plans, respectively. In contrast, without artifact corrections, up to 18% of measured points fail the gamma criterion of 3%/3 mm for the SFUD plan. These measurements indicate that correcting manually for the reconstruction artifacts resulting from metal implants substantially improves the accuracy of the calculated dose distribution.

Keywords: proton radiotherapy, metal implants, phantom, IMPT

(Some figures may appear in colour only in the online journal)

1. Introduction

Many patients treated with proton therapy (PT) along the spinal axis are referred after a partial or complete removal of one or more vertebral body. The spine is then stabilized with the insertion of anterior (cage and anterior plate) or posterior (posterior rod and pedicular screw) implants, usually made of titanium. In a recent paper and following on from the paper of (Rutz *et al* 2007 Staab *et al* 2011) studied the follow-up of 40 patients treated extra-cranially for chordoma and chondrosarcoma between 1999 and 2006 at PSI with spot scanned proton radiotherapy, noting that 12 of the 13 local failures observed in this patient population occurred in patients with stabilizing titanium implants and that the presence of implants was a significant risk factor ($p < 0.01$) for local failure.

There are several potential clinical (e.g. tumor more aggressive or in a more advanced stage) reasons to explain the observed correlation between the presence of titanium implants and local failure. Nevertheless, the presence of such metal implants can cause difficulties for dose calculations, especially in proton therapy. For instance, their sharp density interfaces with the surrounding tissues can degrade the delivered target dose homogeneity. In addition however, such implants, being of high density, can create severe reconstruction artifacts in the planning CT. As accurate proton dose calculations are based on CT values (Hounsfield Units, HU) which are then calibrated to relative proton stopping powers (Mustafa and Jackson 1983 Schneider *et al* 1996) this implies that CT data corrupted by metal artifacts could introduce substantial errors in the evaluation of proton range and therefore the resultant calculated dose distribution.

This problem has been investigated by a number of authors, either through simulations, through re-calculations on artifact free CT data sets or experimentally in simple geometric phantoms. For instance, Kim *et al* 2006 compared the dose distributions obtained from plans calculated on CTs contaminated by metal artifacts resulting from dental implants in the head and neck region to those obtained by the same plan re-calculated on CT's where an artifact-reduction algorithm had been applied. They found that, when compared to the nominal dose distribution, the corrected and uncorrected dose distributions displayed an average hot dose of 1 Gy in the parotids and an average cold dose spot of 2 Gy within the target volume. On the other hand, Wei *et al* 2006 investigated the perturbation on the dose distribution produced in a patient with a bilateral hip prosthesis for a proton plan as well as for photon and electron plans. They found that for all the three irradiation modalities and particularly for PT, if treatment planning is based on CT images contaminated by artifacts, then the target will be severely under-dosed. Similarly, Lim *et al* 2009 and Cheung *et al* 2010 investigated the dose perturbation effects caused by different fiducial markers used for prostate localization, whilst Newhauser *et al* 2007 found that tantalum markers, used for the localization of the eye in uveal melanoma treatments, produce significant perturbations in the absorbed dose distribution and may cause significant dose shadows beyond the clips. Furthermore, both Albertini *et al* 2006 and Newhauser *et al* 2008 have investigated the possibility of using megavoltage CT (MVCT) to create artifact-free images in the presence of metal implants, with these images being then used to re-calculate the delivered dose.

In this paper, we wish to expand on this previous work and experimentally investigate the effect of metal implants on the delivered dose distribution in spot scanned PT, but using

a sophisticated anthropomorphic phantom in order to reproduce the conditions of an *in-vivo* measurement as closely as possible. As such, we have directly measured the relative doses delivered to an anthropomorphic phantom containing a metal implant in the neck region. The phantom used corresponds in size and anatomic characteristics to a real patient and is therefore as close as possible to a true '*in-vivo*' measurement as is currently possible. The full-plan dose distributions in the anthropomorphic phantom have been measured and compared to the calculated dose distributions to check the accuracy of the dose calculation. Furthermore, the validity of the artifact correction procedure currently used in our clinic, which consists of outlining the artifacts as anatomical structures and manually assigning an average soft tissue HU value to these regions, has also been investigated.

2. Material and methods

2.1. The anthropomorphic phantom

For all measurements, a sophisticated anthropomorphic phantom has been used to emulate *in-vivo* conditions as accurately as possible. For such measurements, the Alderson® phantom (Radiology Support Devices, Long Beach, CA, USA) is usually used. This phantom has however two main disadvantages for proton verifications. First, it is divided into sections cut orthogonally to the cranial–caudal axis of the phantom, which leads to air gaps between the slices that perturb the CT image and make accurate measurements with co-planar proton beams extremely difficult. Second, it is composed of just two materials; one for bone and one for all non-bony tissues (see e.g. www.rsdphantoms.com/rt_art.htm). Because density heterogeneities affect the range of the protons significantly, it is therefore important to have a more realistic composition. As such, with the help of CIRS (Norfolk, VA, USA), we designed a new anthropomorphic phantom specifically for PT verifications and measurements (Albertini *et al* 2011) that eliminates these two aforementioned shortcomings.

Size- and anatomical structure-wise, our CIRS anthropomorphic phantom corresponds to an average adult human head, with many anthropomorphic structures, such as brain, bone, larynx, trachea, sinuses, nasal cavities and teeth being modelled in detail and in all cases, a variety of different tissue among equivalent materials have been used (ATOM®max resins, figure 1). As part of its design, the phantom also contains a titanium rod fixed with two screws at a vertebra in the neck and a metal filling (sintered tungsten ball with a diameter of 5 mm) in one tooth. The phantom is halved in the cranio-caudal direction along the middle line and one half is further sliced into three segments, with the first and second cuts being 2 cm and 4 cm respectively away from the mid-line of the phantom. The cuts through the phantom allow us to position up to three GafChromic® films at three different anatomical positions (figure 1). In addition, the titanium rod has been implanted immediately before the second cut, meaning that the film situated in position 2 is in contact with the implant (figure 1(*c*)).

2.2. Treatment planning

In order to make the irradiations as realistic as possible and to be able to ensure consistent positioning among irradiations, the anthropomorphic phantom has been immobilized in such a way that it is possible to reproduce its position before each irradiation. As such, an individualized thermoplastic mask and vacuum pillow have been used, similar to those used clinically for treatments at our center. With these immobilization devices, a planning CT was acquired using a 2 mm slice separation and our standard acquisition protocol for planning CTs in this

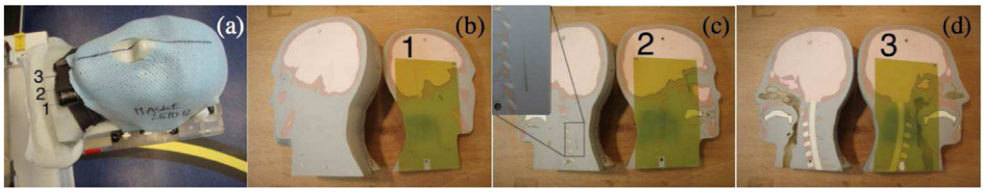

Figure 1. (*a*) The anthropomorphic phantom is shown together with the immobilization device used. The sliced segments are named in the figure from 1 to 3. (*b–d*) illustrate the three segments of the anthropomorphic phantom. Anatomical details and the Gafchromic® films used are visible. In figure (*c*) in the zoomed window the metal rod is visible posterior to the spinal cord in the neck area.

anatomical region (120 KV, standard reconstruction filter). For the dose calculation, the HU values of this CT were then converted to relative proton stopping powers using the stoichiometric calibration curve introduced by Schneider *et al* (1996). It should be noted however, that as the density of titanium is somewhat higher than normal tissues or bone, its true HU value (if measureable) would be considerably higher than the maximum value supported by our CT scanner (3095), which, when using our unaltered stoichiometric calibration curve, would give a relative SP of 2.513, well below the actual, measured relative SP of surgical titanium of 3.250. In order to take this into account, we use clinically a HU-SP calibration curve where the proton SP assigned to the maximum HU value is artificially set to the relative SP of titanium. Such a calibration curve has also been used in this work. Furthermore, as the materials of the phantom are non-biological and our HU-SP calibration curve has been optimized for biological tissues (Schneider *et al* 1996), the actual proton SP's of all materials in the phantom have been directly measured and compared to those calculated from the CT and calibration curve. These show an agreement within 1.5% for the soft tissues and of about 2% for bony structures (Albertini *et al* 2011).

On the planning CT of the phantom, a planning target volume (PTV) of 105 cm^3 has been defined to simulate a cervical spine chordoma, such that the metal rod in the neck of the phantom is embedded within the target volume, together with the spinal cord (figure 2). In addition and as accurately as possible, the extent of all reconstruction artifacts in the soft tissue regions of the CT has also been delineated, being careful to exclude bony structures and the implant itself. This resulted in a cumulative delineated volume of 45 cm^3. For all voxels within this delineated region, the HU units in a copy of the planning CT have then been set to an average HU value for soft tissue, to create an 'artifact corrected' planning CT. This is the procedure used routinely at out clinic for correcting such reconstruction artifacts in planning CTs (Rutz and Lomax 2005).

Two 4-field plans (figure 2) have been planned on the 'artifact corrected' CT, using the PSI treatment planning system (Lomax *et al* 2004). The first was a Single Field Uniform Dose (SFUD) plan and the second an Intensity Modulated (IMPT) plan, with constraints on the spinal cord. While for an SFUD plan every field has a homogeneous dose distribution within the target, the individual fields of an IMPT plan are highly inhomogeneous (Lomax 1999).

Additionally, to evaluate the necessity of correcting for the metal reconstruction artefacts, the more homogeneous plan, i.e. the SFUD plan, has been re-planned on the original, uncorrected CT data set. The IMPT plan was not re-planned on the un-corrected CT data sets as the same effect is expected. All plans have been calculated using the ray casting pencil beam algorithm described by Schaffner *et al* 1999 on a dose calculation grid of 5 × 5 × 5 mm, as used clinically at our institute. Pencil beam size is on average 8 mm sigma at the Bragg peak

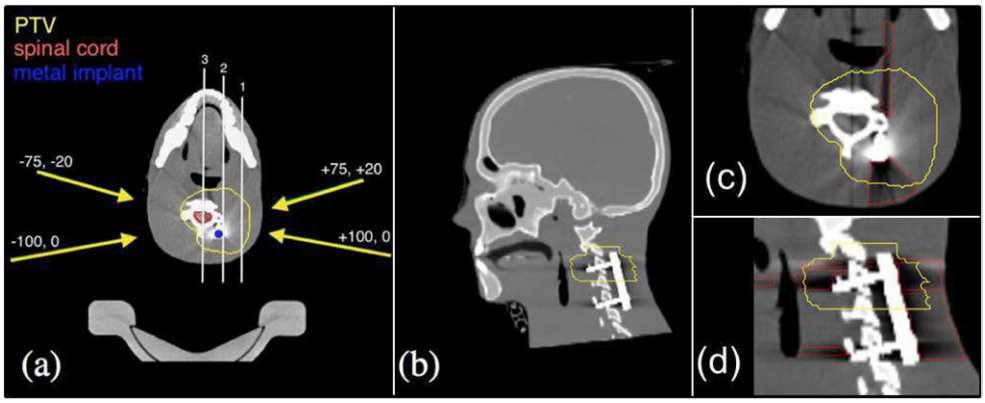

Figure 2. (*a*) Axial view of the 'artifact corrected' CT scan with the beam directions indicated by the yellow arrows (namely, (±75º gantry; ±20º table) and (±100º gantry; 0º table)). The three measurement positions are indicated as white lines along the cranio-caudal direction. The yellow contour shows the PTV, the red one the spinal cord and the blue spot indicates the metal rod. (*b*) Sagittal view of the CT scan with the PTV indicated in yellow. The metal rod and the artefacts are visible in the neck area. (*c*)–(*d*) zoomed views of the axial and the sagittal views of the un-corrected CT, respectively. The degree of the artefacts can be appreciated around the rods and the screw. In red is highlighted the volume defined as 'reconstruction artifacts' and corrected in the planning CT.

(Lomax 2001) and pencil beams are separated by 5 × 5 mm orthogonal to the field direction and 4.6 mm in water equivalent range along the beam direction (the latter value being determined by the water equivalent thickness of the range shifters used to modulate the range on our Gantry 1, see e.g. Pedroni *et al* 1995). Finally, for all plans, the beam angles used were quasi-coplanar and symmetrical from both lateral directions, forming a 'star arrangement' as is our clinical practice for such cases. Figure 2 shows the beam directions, together with the planes of the three measurement positions.

2.3. GafChromic® film measurements

For the measurements, GafChromic® EBT3 films have been used. The use of GafChromic® films for validation of proton dose distributions has been previously reported by different authors (Vatnisky 1997a, Vatnisky *et al* 1997b, Mumot *et al* 2009, Martisikova and Jäkel 2010, Arjomandy *et al* 2010, Zhao and Das 2010, Albertini *et al* 2011, Park *et al* 2011, Reinhardt *et al* 2012, Sorriaux *et al* 2013, Fiorini *et al* 2014). The films were laid between the three different segments of the anthropomorphic phantom (separated by 20 mm) in order to measure the 2D dose distributions in these three planes (figure 1 and figure 2). The films were cut to fit to the phantom and selected anatomical structures were marked to allow a correlation between the measured dose distribution and the phantom. The irradiated films were scanned with a commercial A3 size flatbed scanner (Epson, Expression 10000 XL). The images were acquired in 48 bit color (RGB) mode with a resolution of 72 dpi and saved as uncompressed TIFF format. As suggested by the manufacturer, the films were placed at the center of the scanner bed, in landscape orientation, in order to minimize the error caused by the scanner readout un-uniformity and film orientation, known as the main sources of experimental error (Sorriaux *et al* 2013). Moreover, each image was acquired 3 times and the average image was

used for the analysis, to additionally reduce the residual readout un-uniformity. Images of all the films were acquired before irradiation and subtracted to the images acquired after the irradiation in order to correct for the unexposed film un-uniformity. The films were scanned 48 h after the irradiation in order to ensure a darkening variation less than 0.01% (Sorriaux *et al* 2013). The red color channel was used for the analysis of the digital images. All films used in this work were from the same batch (A03051205). The dose response curve of the batch was obtained from a calibration procedure. Precisely, a film was placed in a PMMA phantom and irradiated with eight uniform square fields of 3 cm side delivering a dose ranging from 0.2 to 4 Gy in the middle of a Spread-Out-Bragg Peak (SOBP) of 10 cm depth. The relation between the film darkening and the dose measured with an ionization chamber (Exradin T2 0.5 c.c.) was obtained. This calibration curve was used to linearize the film response.

Film dosimetry was performed for each of the three plans and in order to further investigate the spatial dependency between the metal, its artifacts and the dose distribution, for each individual field as well.

2.4. Data analysis

The accuracy of the dose calculation in the presence of metal has been investigated by comparing the 2D dose distributions, measured with the films, to the corresponding 2D dose distributions extracted from the treatment planning system. This comparison was done using the IDAS (**I**MRT **D**ose **A**nalysis **S**ystem) software, developed at the Department of Radiation Oncology, University of Ulm, for the dosimetric verification of IMRT plans (Salk *et al* 2003). The measured and calculated dose distributions were firstly cropped to fit the measured data. This process guarantees to analyze all the measured points with a dose greater than zero, while minimizing the points with a dose equal to zero. Afterward these were aligned manually and normalized to a common point selected within a uniform part of the calculated dose distribution. From these data, iso-dose overlays, dose differences and gamma evaluations (Low *et al* 1998) can be performed. For the gamma evaluation, tolerance limits of 3% of the prescribed dose and 3 mm in distance-to-agreement have been used (Zhu *et al* 2011). Additionally and to directly correlate the distortion of the dose distribution with the geometrical localization of the metal, the measured dose distributions have been aligned to the planning CT. This has been achieved by optimizing the difference between the marks drawn on the GafChromic® films (see above) and the anatomy of the CT. Furthermore, dose differences between the measured and calculated doses, together with the gamma evaluation, have been calculated with an in-house developed MATLAB® program.

3. Results

3.1. Artifact corrected plans

3.1.1. SFUD plan. Figure 3 shows the comparison of the calculated and the measured dose distributions of the SFUD plan (see 2.2), calculated on the artifact corrected CT, for all three measurement planes. The doses are represented as color maps as defined by the color bar on the right. As described in section 2.4, the two dose distributions have been normalized to a common point selected in a relatively uniform region of the calculated dose distribution. In the third column of the figure, the iso-dose contours of the calculated and measured overlays are shown using the 10%, 30%, 50%, 70% and 90% iso-dose levels, with the solid and dashed lines representing the calculated and measured doses respectively. Finally, in the last column, the point-to-point gamma index (3%/3 mm) is represented, again in the form of a color map.

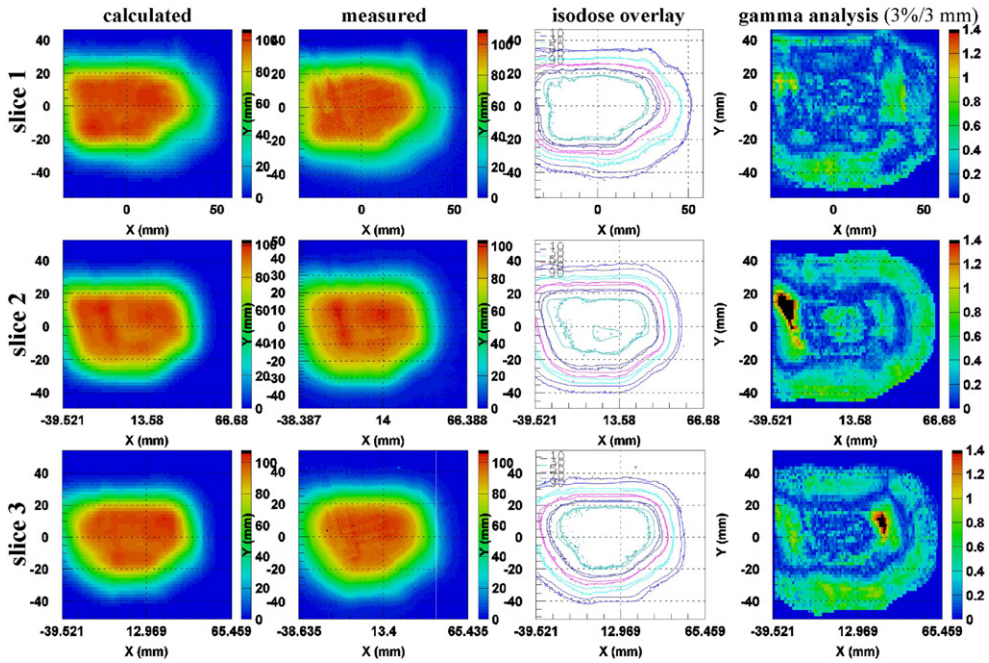

Figure 3. Calculated and measured dose distributions as 2D color maps for the SFUD plan with artefact correction for the three different slices of the anthropomorphic phantom. Isodose contours are overlaid with solid lines for the calculated doses and the dashed ones for the measured. The last column shows the gamma index maps. Tolerance limits of 3%/3 mm were used for the gamma evaluation (see text for details).

Table 1. Number of points satisfying the gamma tolerance limits of 3% in dose and 3 mm in distance.

	artefact corrected		artefact uncorrected
	SFUD	IMPT	SFUD
slice 1	100%	98%	94%
slice 2	97%	98%	89%
slice 3	99%	99%	97%

For all three slices, more than 97% of measured points have a gamma value < 1 (see table 1), indicating a good agreement between the delivered and prescribed dose at all measurement planes However, while for the slice at the distal edge of the target volume (slice 1), the agreement is nearly perfect (nearly 100% of the measured points fulfill the gamma criterion) for the other two slices, there are regions where the gamma evaluation fails, with obvious under-dosage being observed. In particular, for slice 2 (closest to the titanium rod) the highest gamma failure rate of about 3% of all measured points was observed. Figure 4 shows the dose difference and the gamma evaluation for all three slices superimposed on to the relevant sagittal slice of the planning CT. From these, it is clear that the under-dosage on slice 2 is directly next to the metal, with local under-dosages of up to −17.9% being found. Furthermore, a total area of 9.4 cm^2 was found to have dose differences larger than 5%. Maximum under-dosages on slices 1 and 3 are somewhat less (−14.3% and −13.4% respectively) as are the absolute

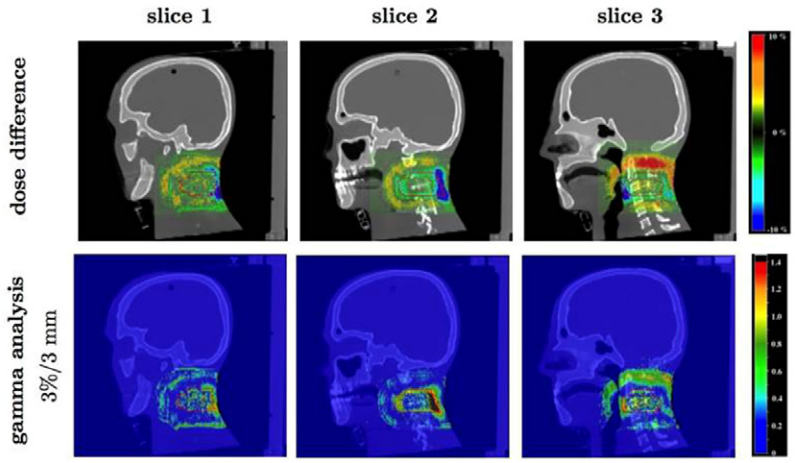

Figure 4. Dose difference and gamma analysis (3%/3 mm) (calculated with a Matlab® program) superimposed on the CT data for the calculated and measured dose distribution of the SFUD plan for all three slices of the anthropomorphic phantom. The images are inverted along the craio-caudal direction in comparison to figure 3.

areas with dose differences >5% (4.6 cm^2 and 4.5 cm^2 respectively). Similarly, the gamma criterion fails in a total area of 1 cm^2 in slice 1 and in about 5 cm^2 and 2.6 cm^2 in slices 2 and 3 respectively. However, only about 0.2 cm^2, 5 cm^2 and 1 cm^2 of the contoured PTV resulted with a difference greater than 5% on slice 1, slice 2 and slice 3, respectively.

Figure 5 shows the calculated and the measured doses, as well as the gamma evaluation, for each individual field compared to the whole plan. Generally, for each slice, the gamma evaluation of the two fields with positive gantry angles are worse than for the other two fields and the regions where the gamma evaluation fails, are mostly situated on the left side. This corresponds to the area in proximity to the surface of the neck of the phantom, as is visible in figure 4. The target boundary coincides here with the quite complex neck profile. Additionally highly weighted pencil beams traversing metal and artefacts deposit the dose there. All together contribute in deteriorating the dose distribution in this area.

Furthermore, on the measured dose distributions, clear projections of the metal implant are visible for some fields. In particular, for slice 1, the projections are visible for beams with a negative gantry angle (field 1 and 3) and for slice 3 for fields with a positive gantry angle (field 2 and 4). These projections are also visible, but not so distinctly, on the calculated distributions and are also faintly visible for the measurements of the whole plan.

3.1.2. IMPT plan. Figure 6 shows comparisons of calculated and measured dose distributions for the IMPT plan. These results are comparable to the results of the SFUD plan, with more than 98% of the measured points satisfying the gamma criterion of 3%/3 mm. Similarly to the SFUD plan, slices 2 and 3 exhibit the largest differences and on the slice closest to the patient surface (slice 1) and on the slice in the middle of the phantom (slice 3) faint projections of the metal are visible.

Similar results as for the SFUD plan were found also for the individual fields of the IMPT plan (not shown here). The metal implant is projected on the measured dose distributions of those fields crossing through it and the gamma analysis fails between 5% and 15% of the measured points depending on the field incidence.

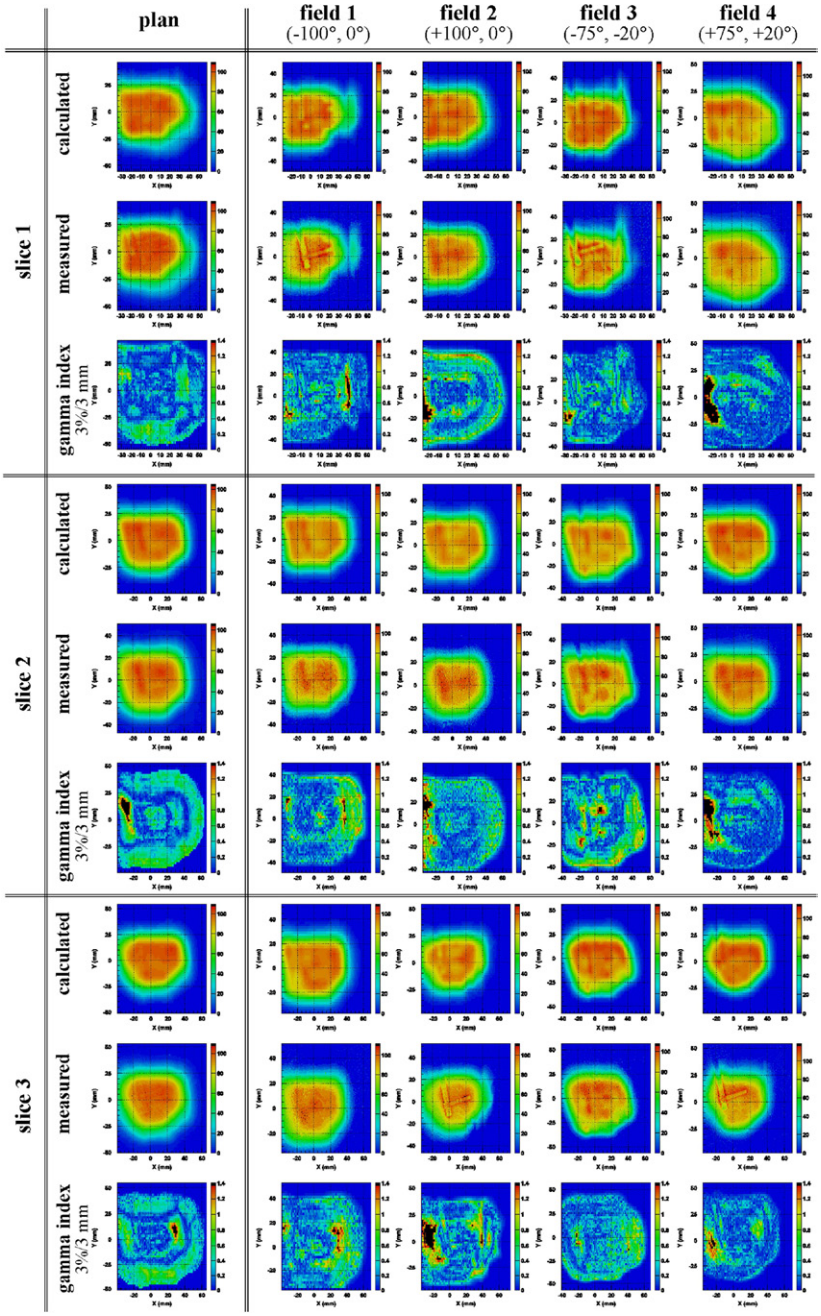

Figure 5. Calculated and measured dose distributions as 2D colour maps and the gamma index map for the SFUD plan and all individual fields for the three different slices of the anthropomorphic phantom.

3.2. Uncorrected plan

As mentioned in section 2.2, all metal artifacts on the planning CT used for the plans in section 3.1 were corrected with our standard procedure. In order to evaluate the effect of this

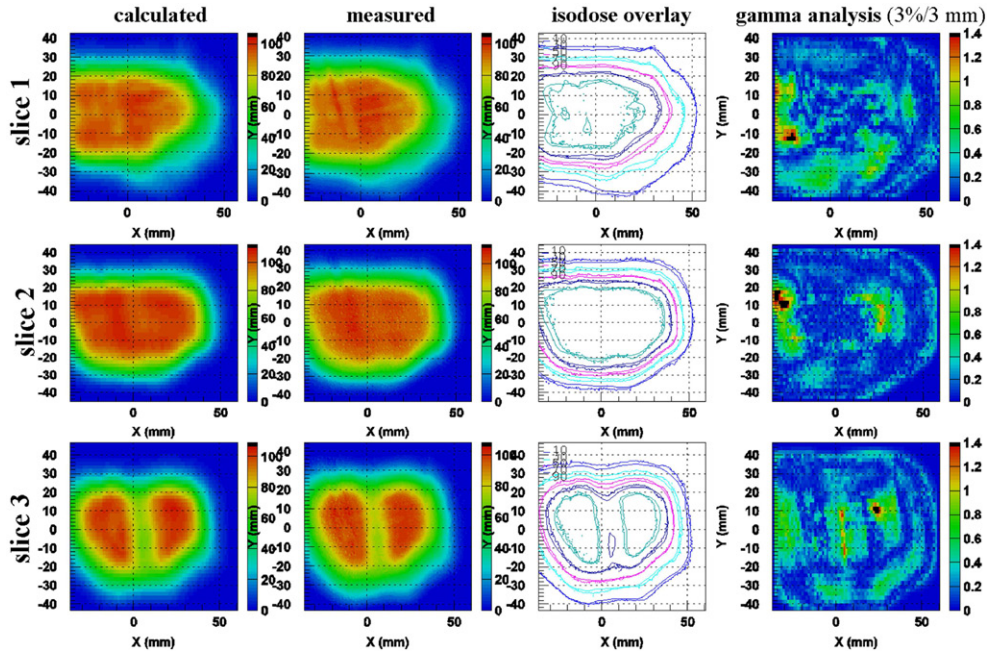

Figure 6. Calculated and measured dose distributions as 2D colour maps for the IMPT plan with artefact correction for the three different slices of the anthropomorphic phantom. Isodose contours are overlaid with solid lines for the calculated doses and the dashed ones for the measured. The last column shows the gamma index maps. Tolerance limits of 3%/3 mm were used for the gamma evaluation.

correction, an additional SFUD plan has been calculated on the original planning CT, i.e. without any artifact correction, and the resulting dose delivered to the phantom was also measured. Figure 7 shows the results of this SFUD plan. For each slice, the normalization point was chosen in the same region as for the SFUD plan with artifact correction.

For the SFUD plan without artifact correction, the gamma criterion fails for about 6% of the measured points for slice 1, 11% for slice 2 and 3% for slice 3 and is significantly higher than the results for the artifact corrected CT. Main areas of poor agreement between the measurements and calculated dose distribution are located in the same positions as for the corrected CT plans (i.e. directly behind the metal implant). In addition, the projections of the metal rod, visible on slice 1 and 3, are more distinct than for the artefact corrected plans.

4. Discussion

In this paper, we have assessed the impact of metal implants and related reconstruction artifacts, on the delivered dose for realistic plans delivered to a sophisticated anthropomorphic phantom.

We have shown a surprisingly good agreement between the measured and calculated doses for both IMPT and SFUD plans calculated in the artifact corrected CT, with more than 97% of the measured points satisfying the gamma tolerance limits of 3%/3 mm (see table 1 and figures 3 and 6) in both cases. This is in agreement with the results achieved earlier by our group for an IMPT plan calculated in a region not affected by metal artefacts (Albertini *et al*

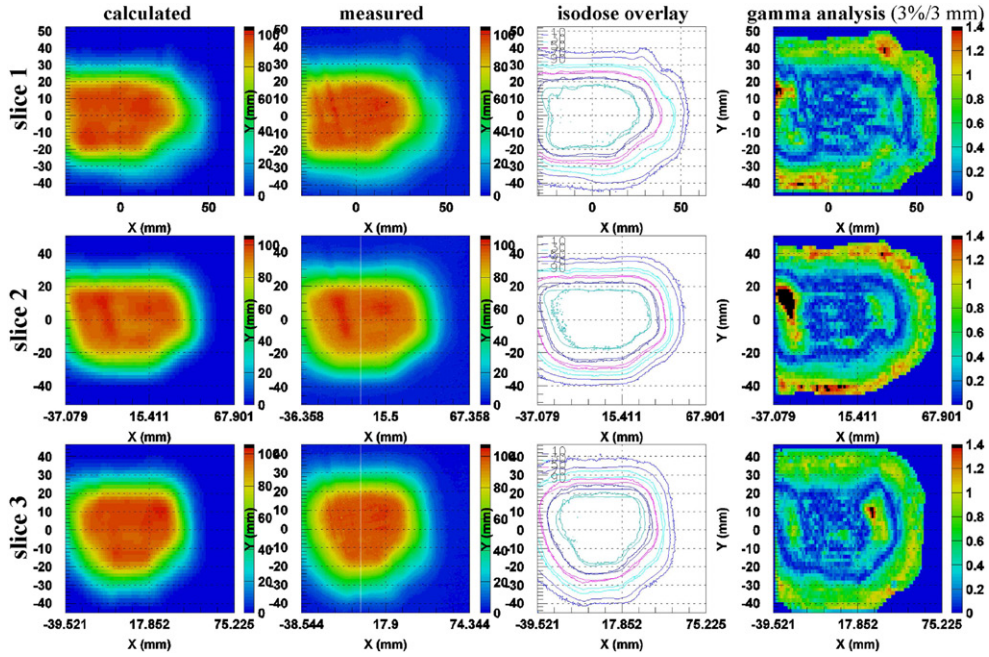

Figure 7. Calculated and measured dose distributions as 2D color maps for the SFUD plan without artefact correction for the three different slices of the anthropomorphic phantom. Isodose contours are overlaid with solid lines for the calculated doses and the dashed ones for the measured. The last column shows the gamma index maps. Tolerance limits of 3%/3 mm were used for the gamma evaluation.

2011). For this case more than 99% of measured points were satisfying the gamma tolerance limits of 3% /3 mm.

Nevertheless, for the actual measurements, some differences are visible when analyzing the plans in more detail. Local differences up to −17.9% have been measured behind the metal for the SFUD plan and for slice 2 (i.e. the slice containing the metal rods), differences greater than 5% have been detected in an area larger than 9 cm^2, whereas the gamma analysis failed in an area of up to 5 cm^2. Furthermore, the rate of failure of the gamma criterion for the individual fields is always worse than for the whole plan (in detail in figure 5). Although this would be expected, these results nevertheless reinforce the importance of using multiple fields, especially when treating patients with metal implants.

In addition and as pointed out in other publications (e.g. Albertini *et al* 2006 and Newhauser *et al* 2008) it has been quantitatively shown here that correcting for artifacts can improve the accuracy of the delivered dose distribution significantly in the presence of metal artifacts. For instance, figures 3 and 7 show that when the metal artifacts are not corrected, there is a deterioration in the agreement between measured and calculated dose (more points failing the gamma tolerances). In contrast, when artifacts are corrected, the same SFUD plan shows a much better agreement to measurements (figure 3), with more than 97% of the measured points satisfying the 3%/3 mm tolerance limits of the gamma analysis. However, as some local differences are still present, there is still a clear need for improved strategies for correcting for artifacts.

On the other hand, a more accurate dose calculation algorithm could improve the agreement between the measurement and the calculation. The ray-casting algorithm clinically

implemented at the Paul Scherrer Institute (Scheib and Pedroni 1992, Schaffner *et al* 1999), as any analytical dose calculation, is extremely fast but reveals its limit especially for pencil beams crossing regions of density heterogeneities (such as air-bone interfaces, metal etc). Therefore, the measurement has been also compared to the in-house developed Monte Carlo Code (Tourovsky *et al* 2005). As expected, the disagreement behind the metal is reduced. In particular, the metal projections, caused by the back-scatter of the metal, measured behind the metal (see figures 5) can be better reproduced by the Monte–Carlo simulation.

Interestingly, the results of this paper suggest that the dosimetric impact of implanted material in chordoma patients is clinically questionable. This is consistent with the recent finding of Verburg and Seco 2013. In their paper, the authors investigated the effect of titanium implants on the clinical dose distribution for chordoma patients treated with passive scattering, by comparing the delivered treatment plans with the ones recalculated with a Monte Carlo simulation employing a metal artifact reduction algorithm. Identically to our study, they observe that chordoma plans for passively scattered proton plans are quite robust against errors introduced by metal implants. As such, one could argue that the observed detrimental effect of metal implants on patient outcome, as reported by our group (Rutz *et al* 2007, Staab *et al* 2011) and other authors (Delaney *et al* 2009), may not be due to the dose distortion induced by the metal artifacts itself but may be due to the local aggressiveness of the tumor. As such, the presence of metal implants may be a proxy to the biological behavior of the tumor or the intrinsic unresectability of this rare bone tumor.

However, it should be noted that the presence of heterogeneity (as e.g. the metal, air-bone interfaces…) along the beam path increases the sensitivity of the plan to uncertainties (Lomax 2008). For that reason, to reduce the set-up uncertainties, it is recommended to carefully immobilize and position each patients based on daily images. Additionally, it is worth accounting for the increase range uncertainty either via the robust optimization (see e.g. Unkelbach *et al* 2009, Fredrikson *et al* 2011) or by increasing the PTV margins.

Finally, it should be noted that GafChromic® films irradiated with particle beams are affected by quenching in the Bragg peak, which leads to a under-response of the films in this region (see e.g. Mumot *et al* 2009, Reinhardt *et al* 2012, Fiorini *et al* 2014). A quenching effect up to 12 % was measured by Mumot *et al* 2009 in the Bragg peak area for an unmodulated beam of 175 MeV. However, the film under-response was only 4% in a spread-out-Bragg peak (SOBP) field, due to the mix of proton energies contributing dose in the SOBP. An even lower under-response of only 2.1% was recently measured by Fiorini *et al* 2014 in the case of a SOBP field for EBT3 films. Similarly for this work, as a full dose distribution achieved by a combination of thousands individually weighted Bragg peaks, incident from different fields, have been analysed, we believe that the quenching effect is to a large part reduced and smoothed out, at least for the analysis of the full plan distributions. Nevertheless, for an accurate determination of the dose at the distal edge of the target a quenching correction factor should be applied. However, as this has not been done in this work, we would expect more discrepancies between calculated and measured dose in the distal ends of the individual field measurements. Indeed, as seen in figure 2, slice 1 (for negative gantry angles) and slice 3 (for positive gantry angles) are close to the distal-end of the target volume and therefore could be affected by such an effect. However, the dose differences here detected for the individual fields can be related also to undershoot caused by range uncertainties. In any case, the dose discrepancy is less evident for the measurements of the full plans, as both effects (quenching and range uncertainties) are reduced with the use of co-planar, laterally opposed field arrangements.

5. Conclusions

In this paper, we have investigated the accuracy of dose calculations in the presence of metal by measuring the dose delivered to an anthropomorphic phantom with a titanium implant in the neck area. Clinically relevant SFUD and IMPT treatment plans have been delivered to a volume simulating a cervical spine chordoma case. Our results indicate that when artifacts are corrected in the planning CT, the accuracy of the dose calculation, even in the presence of metal, is surprisingly good. This implies that, at least from a dosimetric point of view, provided that the beam directions are carefully selected and any reconstruction artifacts are adequately corrected, patients with metal implants could be treated clinically with SFUD and IMPT composite plans.

Acknowledgment

Authors are thankful to Stefano Lorentini, Gabriel Meier and to MD Dr Sabina Vennarini for the useful discussions regarding the radiochromic films analysis, for the support in performing the gamma analysis in the patient anatomy and for the initial definition of the target volume respectively.

References

Albertini F, Bolsi A, Ares C, Broggi S, Cattaneo G M and Lomax T 2006 Advantage of using a MVCT for proton planning *PTCOG 44 Zurich and PSI*

Albertini F, Casiraghi M, Lorentini S, Rombi B and Lomax A J 2011 Experimental verification of IMPT treatment plans in an anthropomorphic phantom in the presence of delivery uncertainties *Phys. Med. Biol.* **56** 4415–31

Arjomandy B, Tailor R, Anand A, Sahoo N, Gillin M, Prado K and Vicic M 2010 Energy dependence and dose response of GafChromic® EBT2 film over a wide range of photon, electron and proton beam energies *Med. Phys.* **37** 1942–7

Cheung J, Kadchadker R J, Zhu X R, Lee A K and Newhauser W D 2010 Dose perturbations and image artifacts caused by carbon-coated ceramic and stainless steel fiducials used in proton therapy for prostate cancer *Phys. Med. Biol.* **55** 7135–47

Delaney T F *et al* 2009 Phase II study of high-dose photon/proton radiotherapy in the management of spine *Int. J. Radiat. Oncol. Biol. Phys* **74** 732–9

Fiorini F, Kirby D, Thompson J, Green S, Parker D J, Jones B and Hill M A 2014 Under-response correction for EBT3 films in the presence of proton spread out Bragg peaks *Phys. Med.* **30** 454–61

Fredriksson A, Forsgren A and Hardemark B 2011 Minimax optimization for handling range and setup uncertainties in proton therapy *Med. Phys.* **38** 1672–84

Kim Y, Tomé W A, Bal M, Mc Nutt T R and Spies L 2006 The impact of dental metal artifacts on head and neck IMRT dose distributions *Radiother. Oncol.* **79** 198–202

Lim Y K 2009 Microscopic gold particle-based fiducial markers for proton therapy of prostate cancer *Int. J. Radiat. Oncol. Biol. Phys.* **74** 1609–16

Lomax A 1999 Intensity modulation methods for proton radiotherapy *Phys. Med. Biol.* **44** 185–205

Lomax A J et al 2007 Intensity modulated proton therapy: a clinical example *Med. Phys.* **28** 317–24

Lomax A J *et al* 2004 Treatment planning and verification of proton therapy using spot scanning: initial experiences *Med. Phys.* **31** 3150–7

Lomax A J 2008b Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2. The potential effects of inter-fraction and interfiled motion *Phys. Med. Biol.* **53** 1043–56

Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61

Martisikova M and Jäkel O 2010 Dosimetric properties of Gafchromic EBT films in monoenergetic medical ion beams *Phys. Med. Biol.* **55**:3741–51

Mumot M, Mytsin G V, Molokanov A G and Malicki J 2009 The comparison of dose measured by radiochromic films and semiconductor detector in a 75 MeV proton beam *Phys. Med.* **25** 105–10

Mustafa A and Jackson D F 1983 The relation between x-ray CT numbers and charged particle stopping powers and its significance for radiotherapy treatment planning *Phys. Med. Biol.* **28** 169–76

Newhauser W D, Giebeler A, Langen K M, Mirkovic D and Mohan R 2008 Can megavoltage computed tomography reduce proton range uncertainties in treatment plans for patients with large metal implants? *Phys. Med. Biol.* **53** 2327–44

Newhauser W D, Koch N C, Fontenot J D, Rosenthal S J, Gombos D S, Fitzek M M and Mohan R 2007 Dosimetric impact of tantalum markers used in the treatment of uveal melanoma with proton beam therapy *Phys. Med. Biol.* **52** 3979–90

Park S A, Kwak J W, Yoon M G, Shin D H, Lee S B, Choc K H, Kang S K, Kim K J, Bae H S and Park S Y 2011 Dose verification of proton beam therapy using the Gafchromic EBT film *Radiat. Meas.* **46** 717–21

Petroni E *et al* 1995 The 200 MeV proton therapy project at PSI: conceptual design and practical realisation *Med. Phys.* **22** 37–53

Reinhardt S, Hillbrand M, Wilkens J J and Assmann W 2012 Comparison of Gafchromic EBT2 and EBT3 films for clinical photon and proton beams *Med. Phys.* **39** 5257–62

Rutz H P and Lomax A 2005 Donut shaped high-dose configuration for proton beam radiation therapy *Strahlenther. Onkol.* **181** 49–53

Rutz H P, Weber D C, Sugahara S, Timmermann B, Lomax A J, Bolsi A, Pedroni E, Coray A, Jermann M and Goitein G 2007 Extracranial chordoma: outcome in patients treated with function-preserving surgery followed by spot-scanning proton beam irradiation *Int. J. Radiat. Oncol. Biol. Phys.* **67** 512–20

Salk J, Kosta M, Blank P and Röttinger E M 2003 IDAS—an extensible framework for IMRT verification *Radiother. Oncol.* **68** (Suppl. 1) 104

Schaffner B, Pedroni E and Lomax A 1999 Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation *Phys. Med. Biol.* **44** 27–41

Scheib S and Pedroni E 1992 Dose calculation and optimization for 3D conformal voxel scanning *Radiat. Environ. Biophys.* **31** 251–526

Schneider U, Pedroni E and Lomax A 1996 The calibration of CT Hounsfield units for radiotherapy treatment planning *Phys. Med. Biol.* **41** 111–24

Sorriaux J, Kacperek A, Rossomme S, Lee J A Bertrand D, Vynckier S and Sterpin E 2013 Evaluation of Gafchromic EBT3 films characteristics in therapy photon, electron and proton beams *Phys. Med.* **29** 599–606

Staab A, Rutz H P, Ares C, Timmermann B, Schneider R, Bolsi A, Albertini F, Lomax A, Goitein G and Hug E 2011 Spot-scanning-based proton therapy for extracranial chordoma *Int. J. Radiat. Oncol. Biol. Phys.* **81** e489–96

Tourovsky A, Lomax A J, Schneider U and Pedroni E 2005 Monte Carlo dose calculations for spot scanned proton therapy *Phys. Med. Biol.* **50** 971–81

Unkelbach J, Bortfeld T, Martin B C and Soukup M 2009 Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning *Med. Phys.* **36** 149–63

Vatnitsky S 1997a Radiochromic film dosimetry for clinical proton beams *Appl. Radiat. Isot.* **48** 643–51

Vatnitsky S M, Schulte R W M, Galindo R, Meinass H J and Miller D W 1997b Radiochromic film dosimetry for verification of dose distributions delivered with proton-beam radiosurgery *Phys. Med. Biol.* **42** 1887–98

Verburg J M and Seco J 2013 Dosimetric accuracy of proton therapy for chordoma patients with titanium implants *Med. Phys.* **40** 071727

Wei J, Sandison G A, His W C, Ringor M and Lu X 2006 Dosimetric impact of a CT metal artifact suppression algorithm for proton, electron and photon therapies *Phys. Med. Biol.* **51** 5183–97

Zhao L and Das I J 2010 GafChromic® EBT film dosimetry in proton beams *Phys. Med. Biol.* **55** N291–301

Zhu X R *et al* 2011 Patient-specific quality assurance for prostate cancer patients receiving spot scanning proton therapy using single-field uniform dose *Int. J. Radiat. Oncol. Biol. Phys.* **81** 552–9