

PAPER

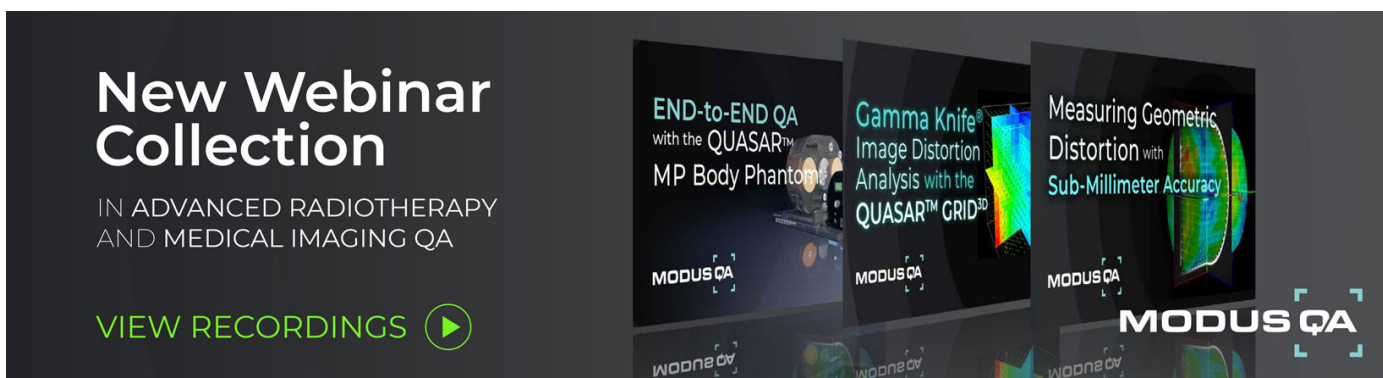
A new treatment planning approach accounting for prompt gamma range verification and interfractional anatomical changes

Recent citations


- [Roadmap: proton therapy physics and biology](#)
Harald Paganetti *et al*

To cite this article: Liheng Tian *et al* 2020 *Phys. Med. Biol.* **65** 095005

View the [article online](#) for updates and enhancements.



New Webinar Collection
IN ADVANCED RADIOTHERAPY AND MEDICAL IMAGING QA

[VIEW RECORDINGS](#) 

MODUS QA

END-to-END QA with the QUASAR™ MP Body Phantom

Gamma Knife® Image Distortion Analysis with the QUASAR™ GRID^{3D}

Measuring Geometric Distortion with Sub-Millimeter Accuracy

MODUS QA



PAPER

OPEN ACCESS


RECEIVED
17 November 2019REVISED
17 February 2020ACCEPTED FOR PUBLICATION
5 March 2020PUBLISHED
24 April 2020

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 3.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



A new treatment planning approach accounting for prompt gamma range verification and interfractional anatomical changes

Liheng Tian¹, Guillaume Landry^{1,2}, George Dedes¹, Marco Pinto¹ , Florian Kamp², Claus Belka^{2,3} and Katia Parodi^{1,4}¹ Ludwig-Maximilians-Universität München, Department of Medical Physics, Munich, Germany² Department of Radiation Oncology, University Hospital, LMU Munich, Munich, Germany³ German Cancer Consortium (DKTK), Munich, Germany⁴ Author to whom any correspondence should be addressed.E-mail: Katia.Parodi@physik.uni-muenchen.de

keywords: prompt gamma, Monte Carlo, proton therapy, treatment planning system

Abstract

Prompt gamma (PG) imaging is widely investigated for spot-by-spot *in vivo* range verification for proton therapy. Previous studies pointed out that the accuracy of prompt gamma imaging is affected by the statistics (number of protons delivered per pencil beam) of the proton beams and the conformity between prompt gamma and dose distribution (PG-dose correlation). Recently a novel approach to re-optimize conventional treatment plans by boosting a few pencil beams with good PG-dose correlation above the statistics limit for reliable PG detectability was proposed. However, up to now, only PG-dose correlation on the planning computed tomography (CT) was considered, not accounting for the fact that the robustness of the PG-dose correlation is not guaranteed in the cases of interfractional anatomical changes. In this work, this approach is further explored with respect to the robustness of the PG-dose correlation of each pencil beam in the case of interfractional anatomical changes. A research computational platform, combining Monte Carlo pre-calculated pencil beams with the analytical Matlab-based treatment planning system (TPS) CERR, is used for treatment planning. Geant4 is used for realistic simulation of the dose delivery and PG generation for all individual pencil beams in the heterogeneous patient anatomy using multiple CT images for representative patient cases (in this work, CTs of one prostate and one head and neck cancer patient are used). First, a Monte Carlo treatment plan is created using CERR. Thereby the PG emission and dose distribution for each individual spot is obtained. Second, PG-dose correlation is quantified using the originally proposed approach as well as a new indicator, which accounts for the sensitivity of individual spots to heterogeneities in the 3D dose distribution. This is accomplished by using a 2D distal surface (dose surface) derived from the 3D dose distribution for each spot. A few pencil beams are selected for each treatment field, based on their PG-dose correlation and dose surface, and then boosted in the new re-optimized treatment plan. All treatment plans are then fully re-calculated with Monte Carlo on the CT scans of the corresponding patient at three different time points. The result shows that all treatment plans are comparable in terms of dose distribution and dose averaged LET distributions. The spots recommended by our indicators maintain good PG-dose correlation in the cases of interfractional anatomical changes, thus ensuring that the proton range shift due to anatomical changes can be monitored. Compared to another proposed spots aggregation approach, our approach shows advantages in terms of the detectability and reliability of PG, especially in presence of heterogeneities.

1. Introduction

Proton therapy offers superior dosimetric properties compared to conventional photon therapy due to the favorable energy deposition peak (Bragg peak) at the end of the proton range, i.e. the beam penetration

depth in the patient. However, the proton range has uncertainties caused by factors such as the semi-empirical conversion of the x-ray computed tomography (CT) numbers into tissue stopping power ratio (relative to water), along with patient positioning and anatomical changes (Paganetti 2012, Engelsman *et al* 2013, Kraan *et al* 2013, Müller *et al* 2015). To reduce proton range uncertainties in clinical practice, prompt gamma (PG) imaging, which monitors the proton range by detecting the prompt gamma emitted by nuclear de-excitation along the beam path, was firstly proposed and investigated by Stichelbaut and Jongen (2003) and then experimentally demonstrated in Min *et al* (2006), Richter *et al* (2016), Xie *et al* (2017) for *in vivo* proton range verification.

Several studies show that the precision of the PG depends on the statistics (number of protons delivered per pencil beam) of the incident proton beam (Xie *et al* 2017, Draeger *et al* 2018). A shift retrieval precision of 2 mm can be achieved when the pencil beam has 2×10^8 protons, according to the first clinical spot-by-spot PG imaging study with a knife-edge slit camera (Xie *et al* 2017). On the other hand, as reported by a few other studies, e.g. Janssen *et al* (2014), Priegnitz *et al* (2015), Schmid *et al* (2015), Tian *et al* ((2018), the accuracy of the PG-based range retrieval could be affected by tissue heterogeneities, potentially spoiling the correlation of the PG signal to the dose delivery. In traditional treatment planning system (TPS), the factors mentioned above are not considered. As a first step to overcome this, our Monte Carlo (MC) based study (Tian *et al* 2018) proposed to re-optimize the treatment plan (TP) accounting for *in vivo* PG imaging by quantifying the conformity between the PG and dose distribution (PG-dose correlation) and boosting a few selected pencil beams (PBs) above the detectability statistic threshold.

Up to now, only PG-dose correlation on the planning CT is taken into account. The PG-dose correlation could be deteriorated due to interfractional anatomical changes, i.e. the PG-dose correlation is not guaranteed in the subsequent fractions. In this work, a new indicator quantifying the effect of heterogeneities on the 3D dose distribution of a given spot is proposed to improve the identification of PG-friendly PB selection of Tian *et al* (2018) in terms of the robustness of the PG-dose correlation in the cases of interfractional anatomical changes. The proof-of-concept is demonstrated for two representative clinical cases of one head and neck and one prostate cancer patient with multiple CT images and compared to other proposed methods such as spot aggregation (Xie *et al* 2017).

2. Materials and methods

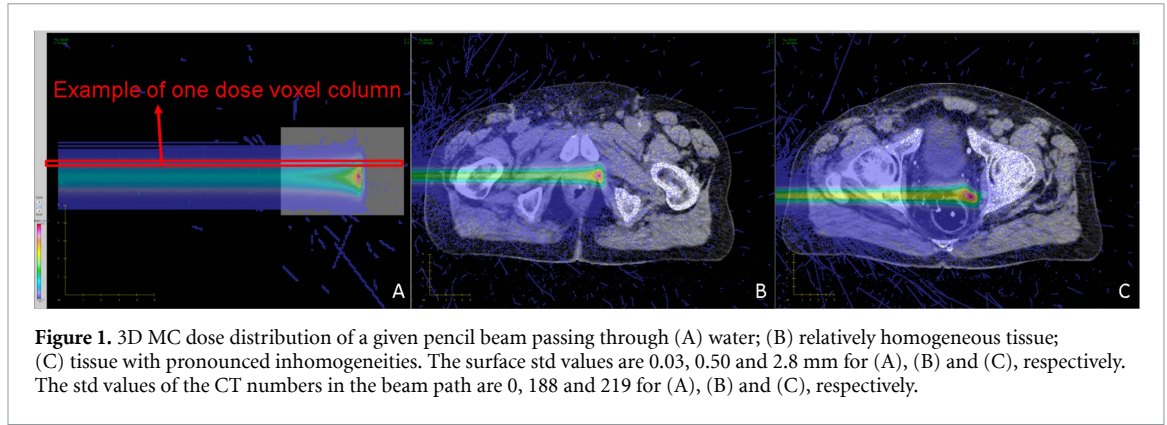
2.1. Monte Carlo treatment plan and patients data

Following (Tian *et al* 2018) MC TPs are created using the particle extension of CERR (A Computational Environment for Radiotherapy Research) (Deasy *et al* (2003), Schell and Wilkens 2010, Resch *et al* 2017). The MC dose and PG data used in this study are created by Geant4.10.02 patch-01. The Geant4 settings are the same as previous studies (Schmid *et al* 2015, Tian *et al* 2018).

CT scans taken at different time points of one head and neck and one prostate cancer patient are used for proton treatment planning. For the head and neck patient, the CT2 and CT3 were taken several weeks after CT1 during the course of the treatment, while the CT1, CT2 and CT3 of the prostate cancer patient were taken on 3 separate days in a row. The CT scanner is a Toshiba Aquilion LB (Toshiba Medical Systems, the Netherlands) and CT scans are reconstructed on a (x, y, z) grid of 1.074 mm \times 1.074 mm \times 3 mm. The CT scans at different time points for the same patient are registered using reggui (Janssens *et al* 2011). TPs for all patients are made based on the planning CT (called CT1 in the following) and then fully recalculated on the other CTs (CT2 and CT3) using Geant4 for evaluation. Only one treatment field is used for the head and neck cancer patient, and two fields are used for the prostate cancer patient. The TP created without accounting for PG emission in this paper is called initial CERR TP in the following.

2.2. Identification of PG-friendly pencil beams

The idea of using a distal 2D surface was proposed to analyze the 3D positron emission tomography (PET) data of passively scattered proton beam delivery in another study (Min *et al* 2014). In this work, a similar concept is introduced to evaluate the effect of heterogeneities on the MC dose distribution of a given PB. Such a surface is generated as follows: 1. the 3D dose distribution is rotated so that the beam direction is along the left-right (LR) axis; 2. the rotated 3D dose consists of dose voxel columns along LR axis (see example shown in figure 1(A)). The 2D dose resulting from averaging in beam direction (summation along LR) was used to select dose columns with a value higher than 50% of the maximum average dose. Thus only dose columns which are close to the center of the beam are considered; 3. the 80% fall-off position of the LR profile of each dose column is then obtained as the range of this dose column (all profiles have the same origin); 4. for the 3D dose distribution of a given PB, a 2D surface is composed of the range of each dose column (called dose surface in the following). To investigate the degree of the tissue heterogeneity in the



beam path, the CT numbers (HU) in the voxels in the same dose columns but starting from the patient skin to the 80% distal dose falloff are analyzed in terms of mean and standard deviation (std) value.

Figure 1 compares the dose distribution of a given PB delivered in different materials. The effect of heterogeneities to the dose can thus be quantified by calculating the std of the dose surface, S_{std} :

$$S_{\text{std}} = \sqrt{\frac{\sum_{i=1}^n (R_i - R_{\text{mean}})^2}{n - 1}} \quad (1)$$

$$R_{\text{mean}} = \frac{\sum_{i=1}^n R_i}{n} \quad (2)$$

where R_i is the range of the dose column i , R_{mean} is the mean value of R_i for a given PB, and n is the number of columns under consideration.

Indicators based on the previously suggested PG-dose correlation (Tian *et al* 2018) and the dose surface from CT1 are combined (figure 2). For the dose surface indicator, a S_{std} less than 1.5 mm, slightly higher than the desired PG precision of 1 mm (Tian *et al* 2018), is identified as a good surface. The threshold of 1.5 mm is set based on the distribution of the std values of the surfaces of PBs in this work. Considering the potential changes of heterogeneities, only PBs which have good surface indicators, and which are surrounded by other PBs with good surface indicators are identified as good PBs, i.e. PBs whose center do not fall within a 5 mm safety ring of any PB with a S_{std} more than 1.5 mm are considered. This threshold of 5 mm is selected the same as the spacing between the PBs used in the TP. All PBs recommended can be selected as long as the dose distribution of those PBs do not overlap (Tian *et al* 2018). The mean and std values of the CT numbers are not used when identifying good PBs, only for subsequent analysis.

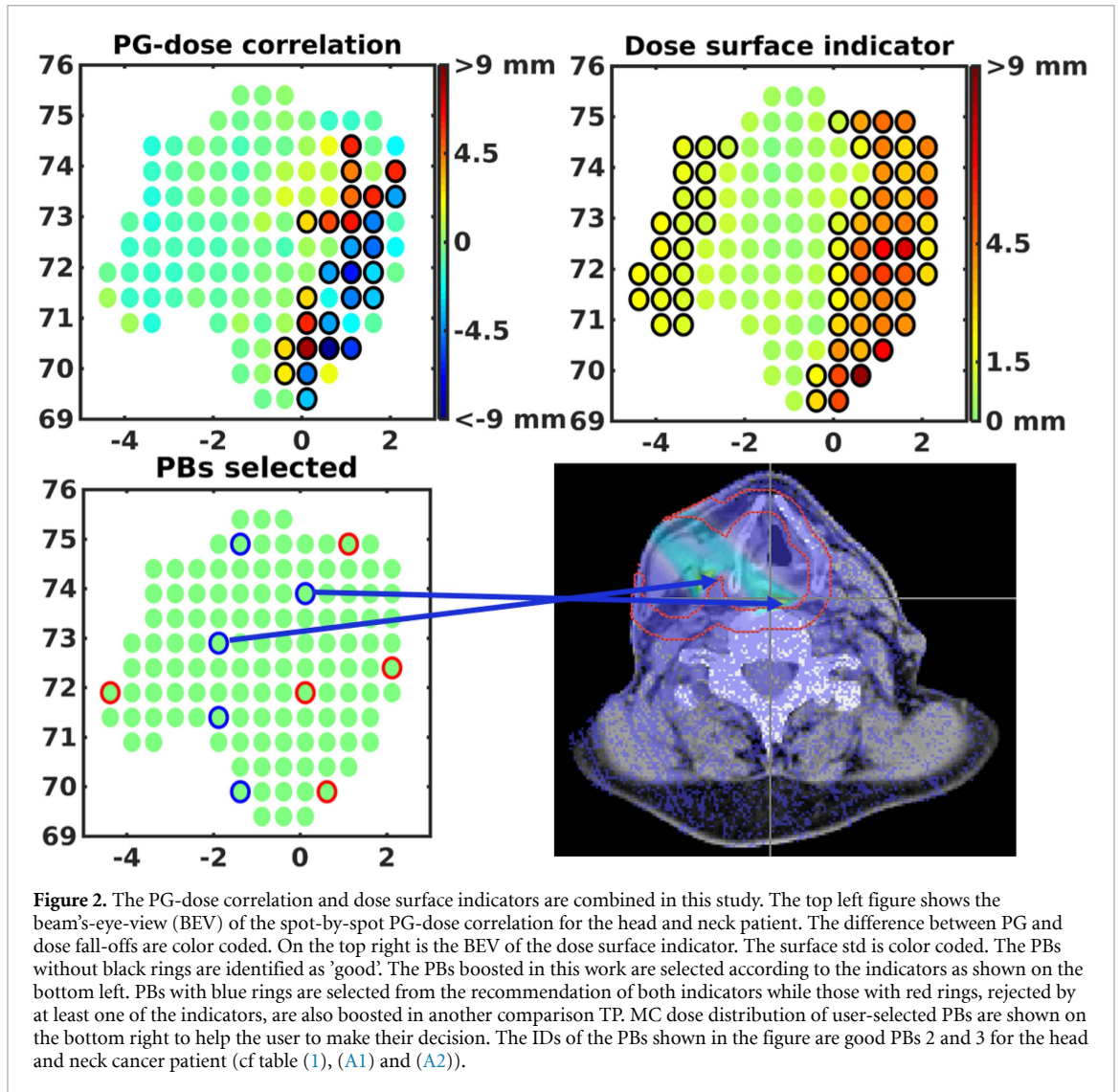
2.3. TPS re-optimization

In this work, five PBs, uniformly distributed across the entire field, are randomly selected among the PBs recommended by both indicators. These five PBs are used in a new plan called TP boosting good PBs (see blue circles in figure 2). Another five PBs per field, which are rejected by at least one of the indicators, are randomly selected in another re-optimized TP (called TP boosting counter-indicated PBs in the following) for comparison (see red circles in figure 2).

The precision of a PG profile as a function of proton statistics for the data used in this work has been studied better in Tian *et al* (2018) and 1.35×10^8 protons per PB were shown to be required to achieve a PG precision of 1 mm, which is consistent with the values reported for current prototypes of slit and Compton cameras by Xie *et al* (2017), Draeger *et al* (2018) but could be reduced in the future, if more efficient detector systems would become available. A good compromise is required above this value, taking into account the required statistics for available detector solutions and the constraints that the chosen value should not spoil the TP optimization, e.g. by introducing hot spots. In this work, new TPs are created by boosting the selected PBs above the statistics thresholds which are set to 1.8×10^8 and 2.0×10^8 protons per PB for the head and neck and the prostate cancer patient, respectively. These values are set around the highest values observed in the initial CERR plan. More details about boosting given PBs in CERR are given in our previous work (Tian *et al* 2018).

2.4. Comparison to PB aggregation approach

The alternative approach of spots aggregation proposed by Xie *et al* (2017) was tested as comparison based on the initial CERR TP. The same PBs selected above were aggregated using nearby PBs. The Gaussian kernel



was set to 7 mm as in the original proposal and PBs within 14 mm of the center PB were aggregated. Since this work focuses on the conformity between the shift of the dose and PG, it is necessary to check the contribution of the central PB to the aggregation calculated by:

$$C = \frac{N_{\text{central}}}{N_{\text{aggregation}}} \quad (3)$$

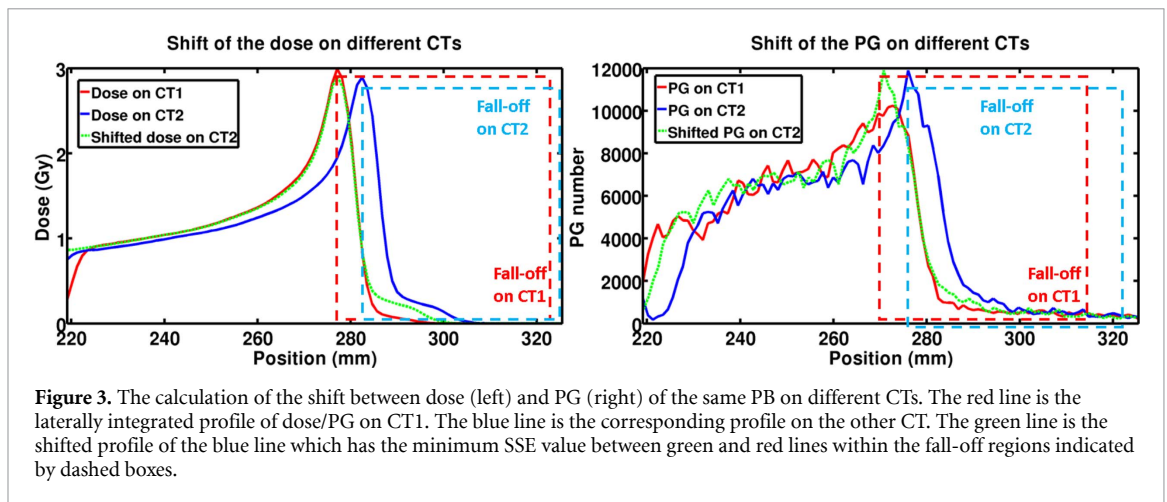
$$N_{\text{aggregation}} = \sum_{i=1}^m w_i * N_i \quad (4)$$

where C is the contribution of the given central PB, N_{central} is the statistics of this central PB, $N_{\text{aggregation}}$ is the statistics of the aggregated PBs, m is the number of nearby PBs aggregated, w_i is the Gaussian weight of PB i and N_i is its statistics in the initial CERR TP.

On the other hand, due to the tissue heterogeneities, aggregated spots may have different proton ranges despite of the same incident energy. To evaluate this effect, the range of the laterally integrated dose profile of the central PB is compared to the ranges of each spot contributing to the aggregation of said central PB. The maximum range difference of a given central PB and its nearby PBs is called range mixing (ΔRange) of this PB in the following.

2.5. MC recalculation and data analyses

All the TPs created as well as all individual PBs are fully MC recalculated on the CT1, CT2 and CT3 of the patients using Geant4. The total dose and dose averaged LET distribution of the re-optimized TPs (boosting good/counter-indicated PBs) are compared to that of the initial CERR TP.



To evaluate if a PB is reliable in terms of PG imaging, the shift of the laterally integrated dose and PG profiles of this PB on different CTs of the same patient are compared, respectively (figure 3).

For the dose profile, the fall-off region is defined from the highest point to the end of the profile. Two fall-off regions of dose profiles of the same PB on different CTs are then shifted to minimize the sum of squared errors (SSE) between each other. The shift value is regarded as the dose shift of this PB on corresponding CTs.

The fall-off region of the PG profile is identified using the approach described in Tian *et al* (2018). The PG shift of a given PB on given CTs is then calculated using the same SSE minimization approach.

A PB for which the shift of the dose and PG match, i.e. the difference between the shift of the dose and PG is less than 1 mm, is identified as a reliable PB for PG imaging.

3. Results

3.1. Total dose distribution comparison on different CTs

The total dose distribution of different TPs on different CTs for the head and neck and the prostate cancer patient are compared in figures 4 and 5, respectively. Figure 6 shows the comparison of DVHs of the target and organs at risk (OARs). As shown in the figures, the DVHs of all TPs almost overlap on the same CT scan. The dose of the initial TP and our re-optimized TPs, whether good or counter-indicated PBs are boosted, are comparable. For the prostate patient, hot spots are found on CT3 at the same position of all TPs (initial CERR TP, TP boosting good/counter-indicated PBs).

3.2. Shift of dose and PG on different CTs

Figures 7–9 show the dose surface indicator on CT1, the shift of the dose and PG along with the difference between the shift of the dose and PG on different CTs in BEV. It can be observed that all PBs, whose PG and dose shift do not match, have large surface std on CT1. The maximum difference between the shift of the dose and PG are 0.8 mm and 5.7 mm for the good and counter-indicated PBs identified by the indicators, respectively. The largest dose shifts of the good PBs observed are 5.2 (PG shift: 5.2) and 4.2 (PG shift: 4.1) mm for the head and neck and prostate patient, respectively. Detailed shift data of the selected PBs using PB aggregation and our PB boosting approach are shown in the table (A1)–(A4) in the appendix.

3.3. Comparison to PB aggregation

The selected PBs are compared in terms of statistics, range mixing and the contribution of the central PB (tables 1 and 2) as well as the shift of the dose and PG (tables (A1)–(A4)). For those good PBs, the maximum difference between the shift of the dose and PG are 0.8 and 1.3 mm for boosted PBs and aggregated spots, respectively. Our proposed PB boosting approach is comparable to the PB aggregation in terms of the conformity of the dose and PG shift of recommended PBs. On the other hand, for the counter-indicated PBs, the difference between the shift of the dose and PG can reach 5.7 and 5.8 mm for boosted PBs and aggregated PBs, respectively. Both approaches fail to predict the dose shift using the PG data in the case of monitoring the counter-indicated PBs due to large difference between the shift of the dose and PG. In the TPs created by our approach, the statistics of those selected PBs are always higher than the set statistic threshold (1.8×10^8 and 2.0×10^8 for the head and neck and the prostate cancer patient, respectively), while the PB aggregation approach does not always guarantee the desired statistics. Moreover, tissue heterogeneities cause the mixing

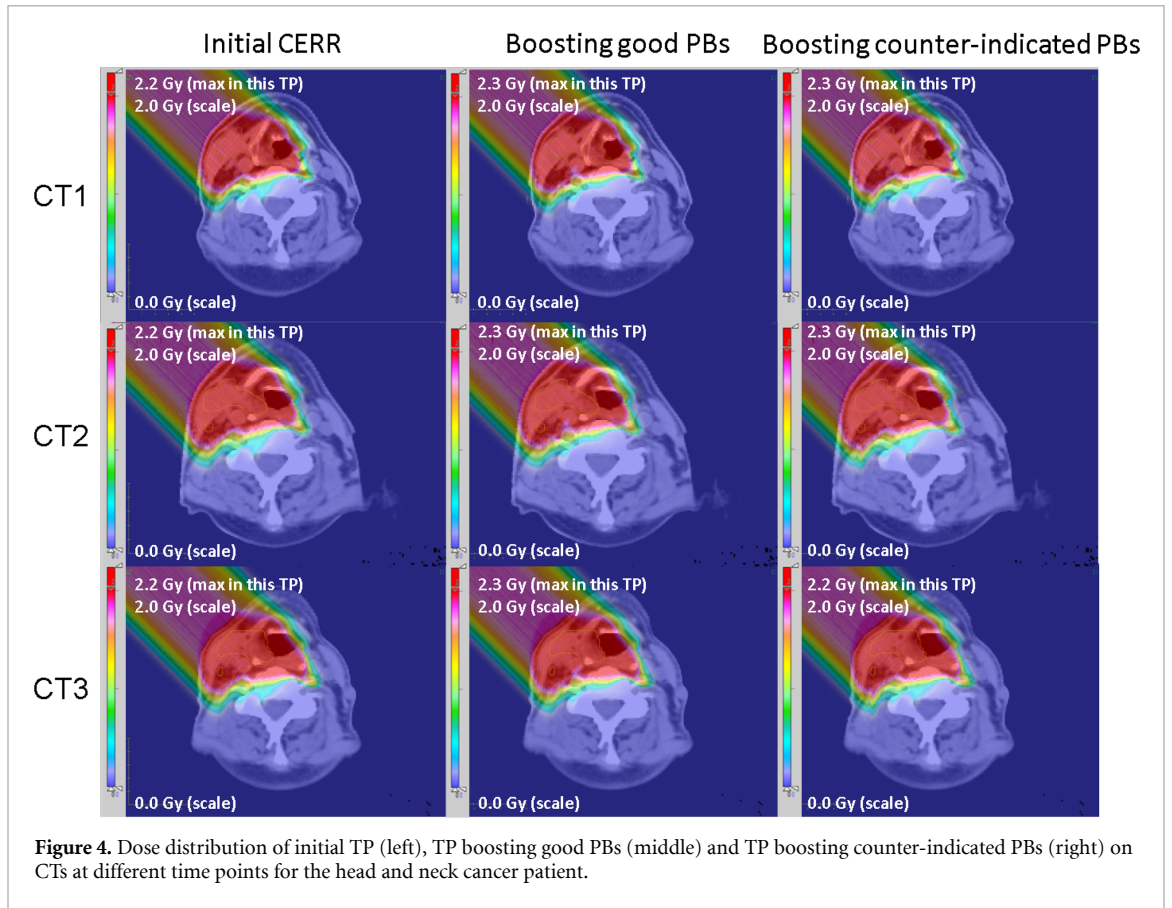


Figure 4. Dose distribution of initial TP (left), TP boosting good PBs (middle) and TP boosting counter-indicated PBs (right) on CTs at different time points for the head and neck cancer patient.

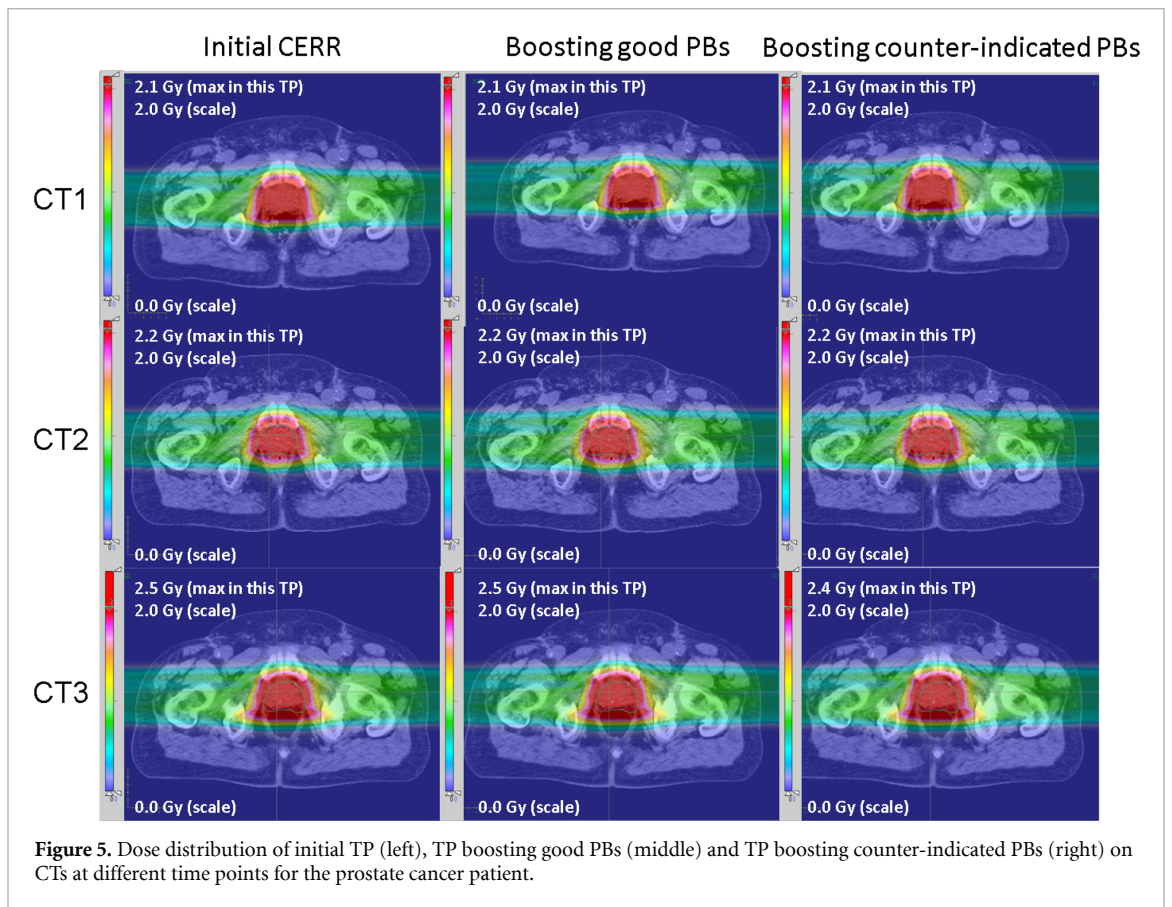
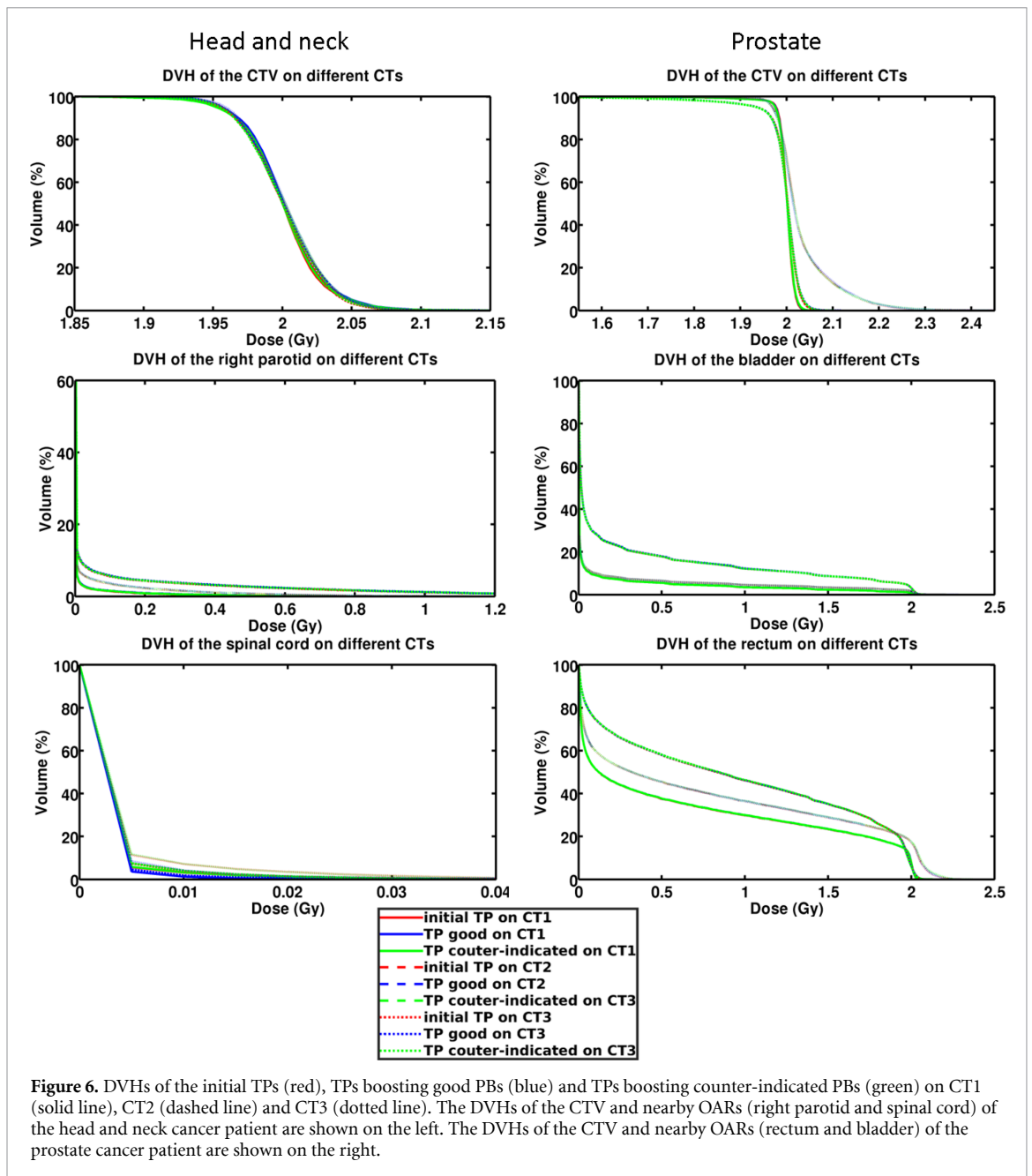


Figure 5. Dose distribution of initial TP (left), TP boosting good PBs (middle) and TP boosting counter-indicated PBs (right) on CTs at different time points for the prostate cancer patient.



of different proton ranges of aggregated PBs. The maximum observed values of range mixing are 5.3 mm and 14.6 mm when aggregating good and counter-indicated PBs, respectively. Besides, there is no guarantee that the central PB contributes dominantly to the aggregated PG profile in the initial TP due to its low statistics. For example, in the table 1, the No.4 good PB used for the head and neck cancer patient has only 1.1×10^6 protons in the initial CERR TP contributing only 0.9% to the total aggregated statistics (1.3×10^8) which still does not reach the set statistical threshold.

3.4. Dose averaged LET

The dose averaged LET distribution of TPs for different patients on different CTs are compared in figures (A1)–(A3). The results show that boosting a few PBs does not visibly change the TP in terms of dose averaged LET distribution. This is comparable to the findings of our initial work (Tian *et al* 2018).

3.5. Tissue heterogeneity in the beam path

The mean and std values of the CT numbers in the beam path of the selected PBs are reported in table A5, and the CT scans with superimposed the dose distribution of two exemplary PBs are shown in figure A4. For the head and neck cancer patient, the std values of the CT numbers in the beam path for the counter-indicated PBs (mean of the std values: 268) are higher than those of the good PBs (mean of the std

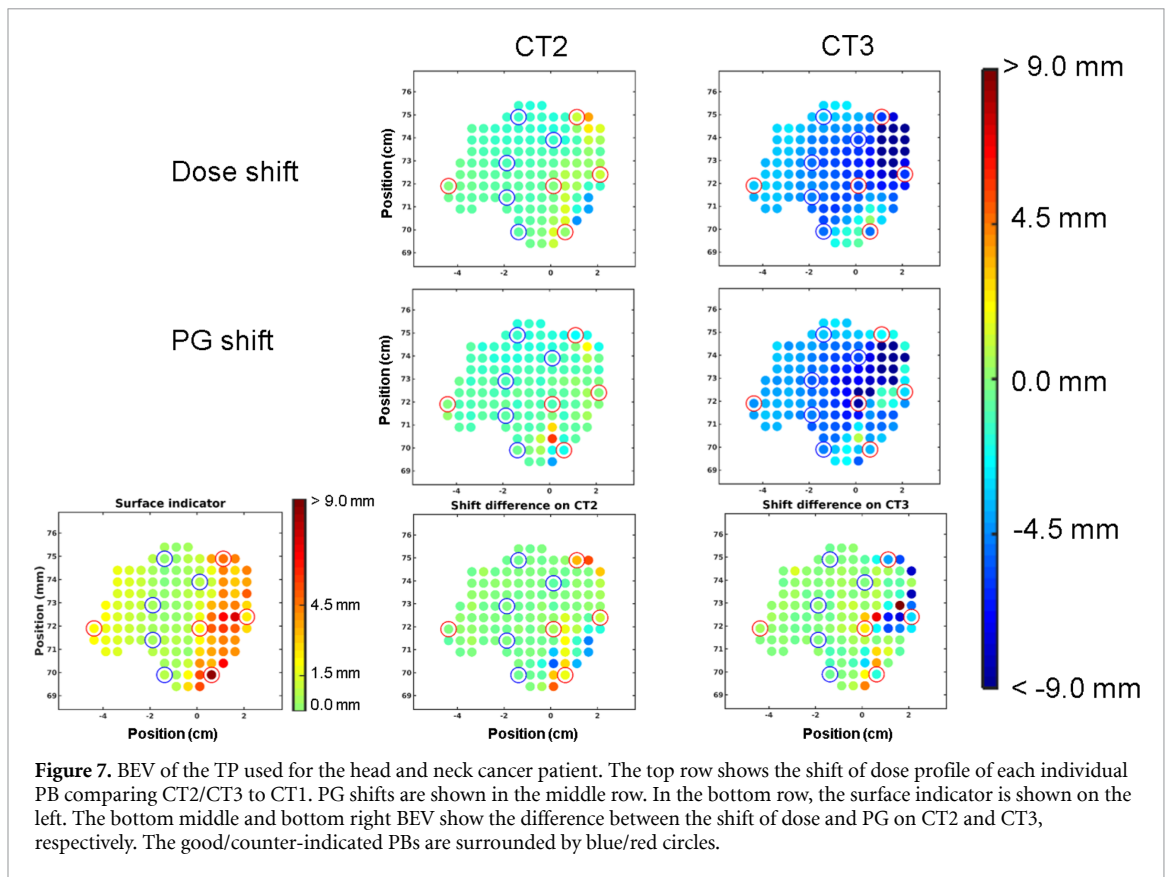


Figure 7. BEV of the TP used for the head and neck cancer patient. The top row shows the shift of dose profile of each individual PB comparing CT2/CT3 to CT1. PG shifts are shown in the middle row. In the bottom row, the surface indicator is shown on the left. The bottom middle and bottom right BEV show the difference between the shift of dose and PG on CT2 and CT3, respectively. The good/counter-indicated PBs are surrounded by blue/red circles.

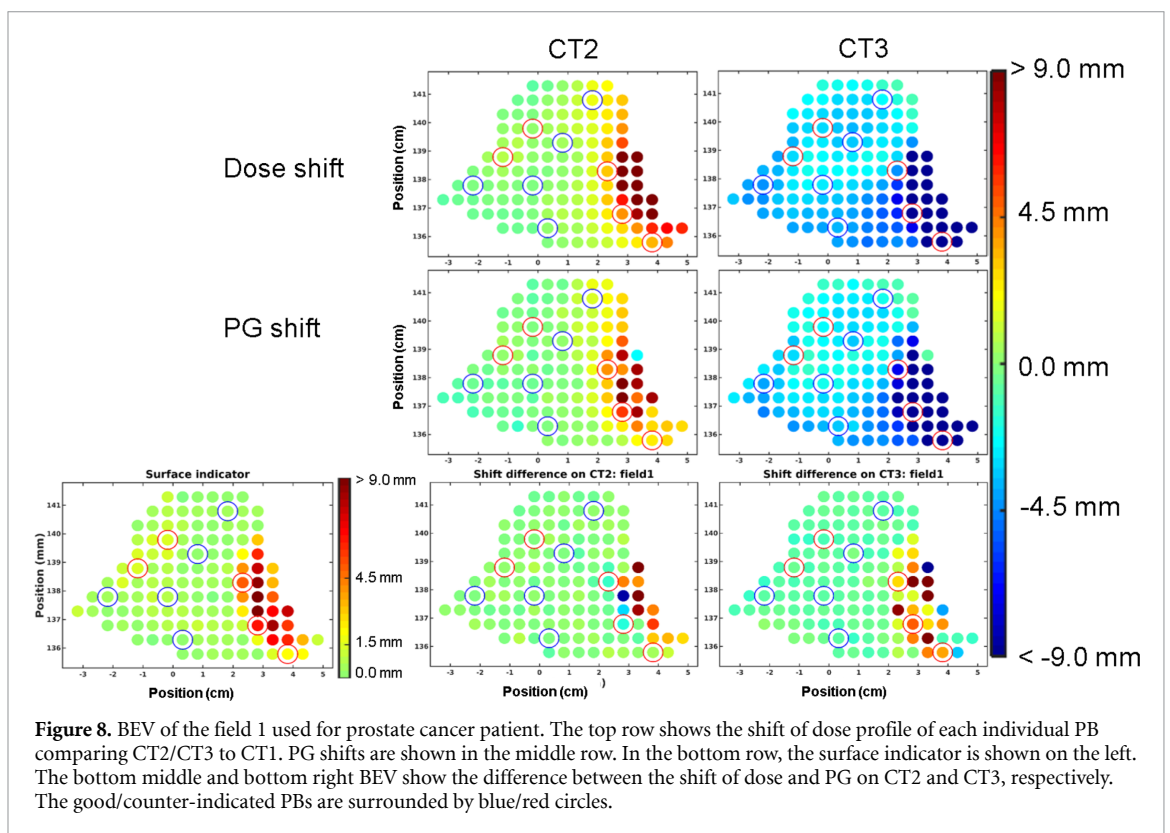


Figure 8. BEV of the field 1 used for prostate cancer patient. The top row shows the shift of dose profile of each individual PB comparing CT2/CT3 to CT1. PG shifts are shown in the middle row. In the bottom row, the surface indicator is shown on the left. The bottom middle and bottom right BEV show the difference between the shift of dose and PG on CT2 and CT3, respectively. The good/counter-indicated PBs are surrounded by blue/red circles.

values: 141), thus corresponding to anatomical regions of higher heterogeneity. For the other patient, the std values of the CT numbers, i.e. the degree of the tissue heterogeneity in the beam path of the counter-indicated PBs (mean of the std values: 194) are comparable to those of the good PBs (mean of the std values: 160). From figure A4 we see that good and counter-indicated PBs do not necessarily correspond to traversed regions of low and high tissue heterogeneity, respectively.

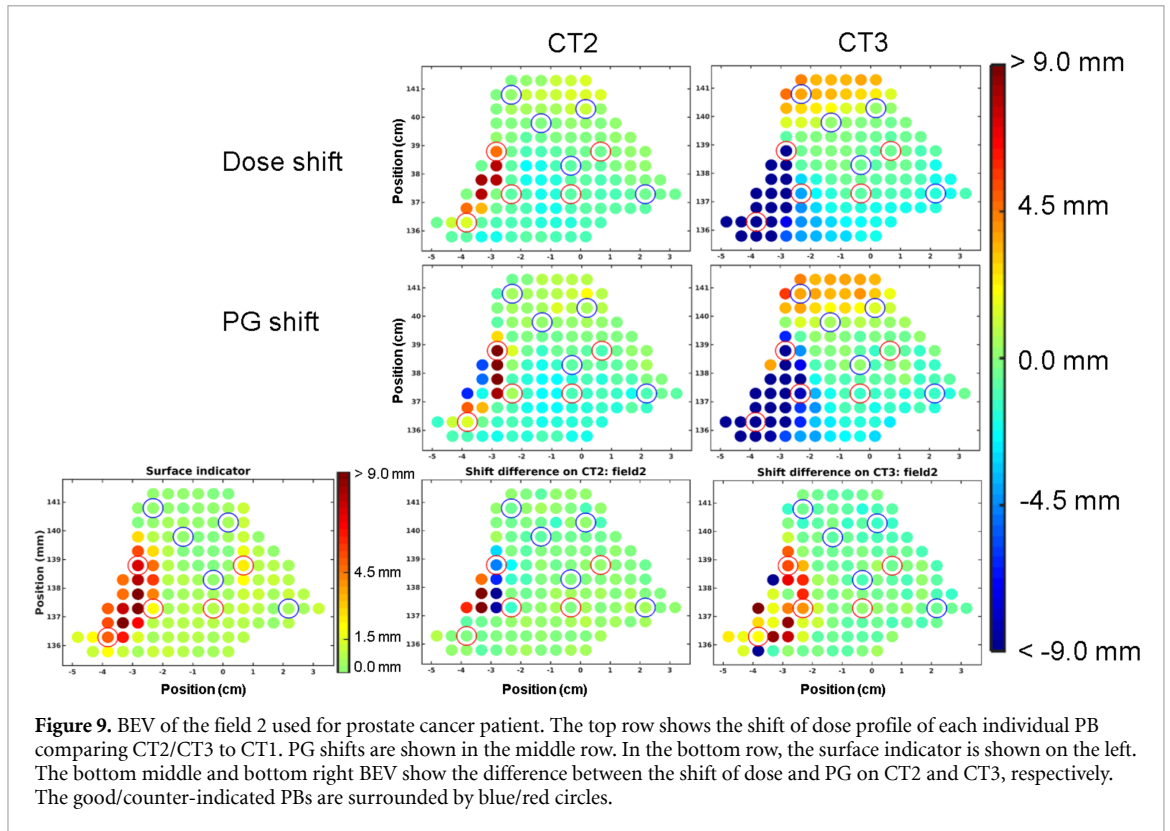


Table 1. The range mixing and the statistics of the selected good PBs using PB aggregation are shown. The range mixing, aggregated statistics, statistics of the central PB in the initial CERR TP and the contribution of the central PB are given by Δ Range, $N_{\text{aggregation}}$, N_{central} , and C, respectively.

PB id	Δ Range (mm)	$N_{\text{aggregation}}$	N_{central}	C
Prostate field1 1	2.9	3.1E+08	7.4E+07	23.8%
Prostate field1 2	2.6	1.9E+08	1.4E+07	7.3%
Prostate field1 3	3.0	1.7E+08	2.8E+07	16.4%
Prostate field1 4	3.8	4.0E+08	4.7E+07	11.8%
Prostate field1 5	4.0	1.4E+08	2.4E+06	1.7%
Prostate field2 1	1.4	1.9E+08	4.6E+07	24.2%
Prostate field2 2	1.8	7.7E+07	4.0E+06	5.2%
Prostate field2 3	4.3	2.6E+08	4.1E+07	15.8%
Prostate field2 4	4.2	6.2E+07	8.4E+05	1.4%
Prostate field2 5	5.3	1.3E+08	7.8E+06	6.0%
Head and neck 1	3.6	2.1E+08	7.7E+05	0.4%
Head and neck 2	4.1	2.4E+08	2.1E+07	8.9%
Head and neck 3	4.6	1.4E+08	8.2E+06	5.9%
Head and neck 4	3.9	1.3E+08	1.1E+06	0.9%
Head and neck 5	3.9	1.5E+08	4.1E+05	0.3%

4. Discussion

In this work, a new indicator based on the PB 3D dose is proposed to evaluate the effect of heterogeneities on the dose distribution of a given PB. In combination with the indicator built in our previous study, the new approach is tested on three CTs at different time points of one head and neck and one prostate cancer patient. PBs recommended by our approach have more robust PG-dose correlation in the case of interfractional anatomical changes. The largest difference between the shift of the PG and dose of those PBs on different CTs is 0.8 mm. This value is small although the dose/PG could shift up to 5.2 mm. On the other hand, the maximum shift difference of counter-indicated PBs is 5.7 mm. The PBs suggested by our indicators are thus better for in-vivo proton range verification. To corroborate our findings it was also verified in a few cases that the shift difference of PG and dose distributions among different CTs, assessed through laterally integrated distributions for the investigated PBs, also holds when resorting to the more detailed line-by-line evaluation. This latter analysis however requires more computationally expensive calculations for accumulating sufficient statistics, especially in the case of the PG data.

Table 2. The range mixing and the statistics of the selected counter-indicated PBs using PB aggregation are shown. The range mixing, aggregated statistics, statistics of the central PB in the initial CERR TP and the contribution of the central PB are given by Δ Range, $N_{\text{aggregation}}$, N_{central} , and C , respectively.

PB id	Δ Range (mm)	$N_{\text{aggregation}}$	N_{central}	C
Prostate field1 1	3.7	2.2E+08	3.2E+07	14.4%
Prostate field1 2	3.0	4.0E+07	2.7E+06	6.6%
Prostate field1 3	4.7	1.9E+08	4.4E+07	23.1%
Prostate field1 4	5.0	1.2E+08	9.1E+07	76.2%
Prostate field1 5	3.8	6.6E+07	5.1E+07	76.7%
Prostate field2 1	3.4	1.5E+08	5.1E+06	3.4%
Prostate field2 2	12.5	5.8E+07	9.1E+06	15.7%
Prostate field2 3	1.7	9.1E+07	4.5E+06	5.0%
Prostate field2 4	4.5	2.0E+08	4.7E+07	23.6%
Prostate field2 5	7.0	1.4E+08	3.2E+06	2.3%
Head and neck 1	14.6	1.8E+08	4.1E+05	0.2%
Head and neck 2	10.8	1.2E+08	3.1E+05	0.3%
Head and neck 3	9.3	2.3E+08	2.4E+06	1.0%
Head and neck 4	7.8	2.3E+08	1.4E+06	0.6%
Head and neck 5	9.2	1.4E+08	6.8E+05	0.5%

The re-optimized TPs (both TP boosting good and counter-indicated PBs) show similar dose and dose averaged LET distribution on the same CT. Boosting a few PBs, either good or counter-indicated, above the statistical threshold does not significantly change the quality of TP in terms of total dose and dose averaged LET distribution. All TPs for the prostate cancer patient show hot spots at the same position on CT3. These hot spots are caused by a comparably dramatic anatomical change, which is hard to predict based only on CT1. This work focuses only on the robustness of PG imaging of boosted PBs instead of the TP robustness, so the TP robustness is not discussed further here. Nevertheless, the re-optimized TPs and the initial CERR TP are still comparable in this case.

For the good PBs, the largest dose shift observed is 5.2 mm while the maximum difference between the shift of dose and PG are 0.8/1.3 mm using PG boosting/aggregation approach. For the counter-indicated PBs, the largest shift difference between the dose and individual/aggregated are 5.7/6.6 mm. The difference between the shift of the dose and PG are comparably small for good PBs using the PB boosting or PB aggregation approach, while for the counter-indicated PBs both approaches fail. As mentioned in our previous study (Tian *et al* 2018), there is no conflict between those two approaches, as PG data of individual PBs could still be aggregated regardless whether PBs are boosted or not. Nevertheless, the PBs recommended by our approach are more likely to provide reliable PG information for proton range monitoring. Monitoring a random PB may not provide correct proton range information. Besides, the statistics of the boosted PBs is guaranteed to be above the desired PG detectability threshold, while some aggregated PBs do not reach this statistical threshold. The contribution of the central PB to the monitored PG signal could be low without special TP re-optimization. For example, one of the counter-indicated PBs (head and neck 3) has only 2.4×10^6 protons, while one of its aggregated nearby PB (more than 10 mm away, with a Gaussian weight around 0.36) has more than 3.0×10^7 protons in the initial CERR TP, i.e. in this scenario the aggregated PG signal may reflect information more about that nearby PB instead of the central PB. Our approach shows advantages in terms of guaranteed statistics, lateral resolution, proton range mixing and the contribution of the central PB.

Note that PBs delivered in heterogeneities could have good dose surface and are not always rejected by the indicators. The dose surface is mainly affected by the transversal heterogeneities. Only a few good PBs presenting reliable information should be enough for the monitoring of the systematic errors, e.g. the CT-range calibration or gross body surface and anatomical changes, which causes a large part of range uncertainties. For the PBs rejected by the indicators, the shift of the PG in the future CTs is not guaranteed to match the shift of the dose, though a large PG shift is observed when the shift of the dose is large for most of the PBs and the difference between the shift of the dose and PG of those PBs may still be small. For those PBs, though a detected PG shift might not be perfectly correlated to the actual dose shift, it could still be valuable to trigger additional investigations, e.g. a new CT. The indicators only provide the information whether the PG shift of a given PB will precisely match the proton range shift or not. Our approach allows the user to select and boost the PBs of interest, e.g. the good ones to test more systematic range errors in regions of reduced transversal heterogeneities, or the counter-indicated ones which are expected to be more sensitive to range uncertainties, while not spoiling the TP.

In our approach, only PBs surrounded by good-dose-surface PBs are recommended. A margin of 5 mm is applied for potential changes of heterogeneities when identifying good PBs. As mentioned above, anatomical changes are hard to predict. Though all PBs recommended in this paper show small difference between the shift of the dose and PG on different CTs while the shift of the dose can reach 5.2 mm, our approach might still fail in the case of dramatic anatomical changes.

5. Conclusion

In this work, a previously proposed approach to select and boost a few PG-friendly PBs in the TP has been further expanded for improved identification by introducing a new indicator based on 3D PB dose distribution. This approach has been tested in the treatment planning for one head and neck and one prostate cancer patient under consideration of interfractional anatomical changes. TPs have been simulated on CTs at three different time points of each patient. The results show that the initial CERR TPs and our re-optimized TPs are comparable in terms of dose and dose averaged LET distribution for all CTs and patients. The recommended PBs show advantages for PG based proton range verification in terms of dose shift fidelity. Besides, compared to spot aggregation, our approach shows advantages in terms of counting statistics, lateral resolution and proton range mixing.

Acknowledgments

This work is funded by EU-MSCA GA n. 675265 (OMA). Financial support from the DFG Excellence Cluster 'Munich Center for Advanced Photonics' is also gratefully acknowledged.

Appendix A.

Table A1. Comparison of the shift of the dose and PG of the good PBs on CT2 and CT1 for both considered patient cases. The shift of the dose, the aggregated PG, and the PG of the boosted central PB as well as shift difference between the dose and aggregated PG, PG of the boosted central PB are given by ΔDose , $\Delta\text{PG}_{\text{aggregation}}$, $\Delta\text{PG}_{\text{boosted}}$, $\Delta\text{Shift}_{\text{aggregation}}$ and $\Delta\text{Shift}_{\text{boosted}}$, respectively. The units are mm. The maximum difference between the shift of the dose and PG are 0.6 mm and -0.6 mm for PB aggregation and PB boosting, respectively.

PB id	ΔDose	$\Delta\text{PG}_{\text{aggregation}}$	$\Delta\text{PG}_{\text{boosted}}$	$\Delta\text{Shift}_{\text{aggregation}}$	$\Delta\text{Shift}_{\text{boosted}}$
Prostate field1 1	2.1	1.7	1.7	0.4	0.4
Prostate field1 2	0.8	0.7	0.4	0.1	0.4
Prostate field1 3	0.1	-0.5	-0.2	0.6	0.3
Prostate field1 4	0.1	-0.4	-0.4	0.5	0.5
Prostate field1 5	0.3	0.3	0.2	0.0	0.1
Prostate field2 1	1.0	0.6	0.8	0.4	0.2
Prostate field2 2	0.4	0.5	0.6	-0.1	-0.2
Prostate field2 3	-0.5	-0.4	-0.4	-0.1	-0.1
Prostate field2 4	-0.3	-0.6	-0.9	0.3	0.6
Prostate field2 5	1.4	1.3	1.0	0.1	0.4
Head and neck 1	-1.2	-1.1	-1.2	-0.1	0.0
Head and neck 2	-1.3	-1	-0.7	-0.3	-0.6
Head and neck 3	-0.8	-0.7	-0.9	-0.1	0.1
Head and neck 4	-0.5	-0.4	-0.5	-0.1	0.0
Head and neck 5	-0.5	-0.5	-0.2	0.0	-0.3

Table A2. Comparison of the shift of the dose and PG of the good PBs on CT3 and CT1 for both considered patient cases. The shift of the dose, the aggregated PG, and the PG of the boosted central PB as well as shift difference between the dose and aggregated PG, PG of the boosted central PB are given by $\Delta Dose$, $\Delta PG_{aggregation}$, $\Delta PG_{boosted}$, $\Delta Shift_{aggregation}$ and $\Delta Shift_{boosted}$, respectively. The units are mm. The maximum difference between the shift of the dose and PG are -1.3 mm and -0.8 mm for PB aggregation and PB boosting, respectively.

PB id	$\Delta Dose$	$\Delta PG_{aggregation}$	$\Delta PG_{boosted}$	$\Delta Shift_{aggregation}$	$\Delta Shift_{boosted}$
Prostate field1 1	-2.1	-1.7	-1.6	-0.4	-0.5
Prostate field1 2	-3.0	-2.3	-3.0	-0.7	0.0
Prostate field1 3	-4.0	-3.6	-3.5	-0.4	-0.5
Prostate field1 4	-2.4	-2.4	-1.9	0.0	-0.5
Prostate field1 5	-3.5	-2.8	-3.0	-0.7	-0.5
Prostate field2 1	4.2	4.0	4.1	0.2	0.1
Prostate field2 2	1.0	0.7	1.2	0.3	-0.2
Prostate field2 3	-0.3	-0.2	0.5	-0.1	-0.8
Prostate field2 4	-1.5	-0.5	-1.2	-1.0	-0.3
Prostate field2 5	1.1	2.4	1.8	-1.3	-0.7
Head and neck 1	-3.0	-3.0	-3.2	0.0	0.2
Head and neck 2	-5.2	-4.8	-5.2	-0.4	0.0
Head and neck 3	-4.6	-4.6	-4.9	0.0	0.3
Head and neck 4	-4.2	-4.4	-4.3	0.2	0.1
Head and neck 5	-4.6	-4.4	-4.3	-0.2	-0.3

Table A3. Comparison of the shift of the dose and PG of the counter-indicated PBs on CT2 and CT1 for both considered patient cases. The shift of the dose, the aggregated PG, and the PG of the boosted central PB as well as shift difference between the dose and aggregated PG, PG of the boosted central PB are given by $\Delta Dose$, $\Delta PG_{aggregation}$, $\Delta PG_{boosted}$, $\Delta Shift_{aggregation}$ and $\Delta Shift_{boosted}$, respectively. The units are mm. The maximum difference between the shift of the dose and PG are 4.2 mm and -4.0 mm for PB aggregation and PB boosting, respectively.

PB id	$\Delta Dose$	$\Delta PG_{aggregation}$	$\Delta PG_{boosted}$	$\Delta Shift_{aggregation}$	$\Delta Shift_{boosted}$
Prostate field1 1	0.5	0.2	0.3	0.3	0.2
Prostate field1 2	1.0	0.4	0.4	0.6	0.6
Prostate field1 3	3.2	3.6	4.3	-0.4	-1.1
Prostate field1 4	4.4	5.1	5.8	-0.7	-1.4
Prostate field1 5	3.6	3.0	2.7	0.6	0.9
Prostate field2 1	-0.2	-0.2	-0.6	0.0	0.4
Prostate field2 2	5.1	2.5	9.1	2.6	-4.0
Prostate field2 3	-0.8	1.4	0.6	-2.2	-1.4
Prostate field2 4	-0.8	-1.4	-0.8	0.6	0.0
Prostate field2 5	2.1	0.5	1.7	1.6	0.4
Head and neck 1	0.3	-1.2	-1.6	1.5	1.9
Head and neck 2	1.6	-2.6	-2.1	4.2	3.7
Head and neck 3	0.0	-1.0	-0.9	1.0	0.9
Head and neck 4	0.6	0.2	0.0	0.4	0.6
Head and neck 5	1.4	0.0	0.1	1.4	1.3

Table A4. Comparison of the shift of the dose and PG of the counter-indicated PBs on CT3 and CT1 for both considered patient cases. The shift of the dose, the aggregated PG, and the PG of the boosted central PB as well as shift difference between the dose and aggregated PG, PG of the boosted central PB are given by $\Delta Dose$, $\Delta PG_{aggregation}$, $\Delta PG_{boosted}$, $\Delta Shift_{aggregation}$ and $\Delta Shift_{boosted}$, respectively. The units are mm. The maximum difference between the shift of the dose and PG are 6.6 mm and 5.7 mm for PB aggregation and PB boosting, respectively.

PB id	$\Delta Dose$	$\Delta PG_{aggregation}$	$\Delta PG_{boosted}$	$\Delta Shift_{aggregation}$	$\Delta Shift_{boosted}$
Prostate field1 1	-2.8	-2.3	-1.9	-0.5	-0.9
Prostate field1 2	-2.7	-1.9	-2.7	-0.8	0.0
Prostate field1 3	-3.9	-4.7	-6.9	0.8	3.0
Prostate field1 4	-15.6	-20.8	-20.8	5.2	5.2
Prostate field1 5	-19.3	-25.1	-23.4	5.8	4.1
Prostate field2 1	-0.4	-0.1	-0.2	-0.3	-0.2
Prostate field2 2	-8.6	-3.9	-14.3	-4.7	5.7
Prostate field2 3	-3.8	-10.4	-8.0	6.6	4.2
Prostate field2 4	-0.8	-1.1	-1.0	0.3	0.2
Prostate field2 5	-27.2	-26.5	-29.4	-0.7	2.2
Head and neck 1	-4.6	-3.5	-3.3	-1.1	-1.3
Head and neck 2	-5.8	-2.1	-2.0	-3.7	-3.8
Head and neck 3	-5.3	-7.2	-8.1	1.9	2.8
Head and neck 4	-2.6	-4.8	-3.9	2.2	1.3
Head and neck 5	-5.6	-3.5	-2.8	-2.1	-2.8

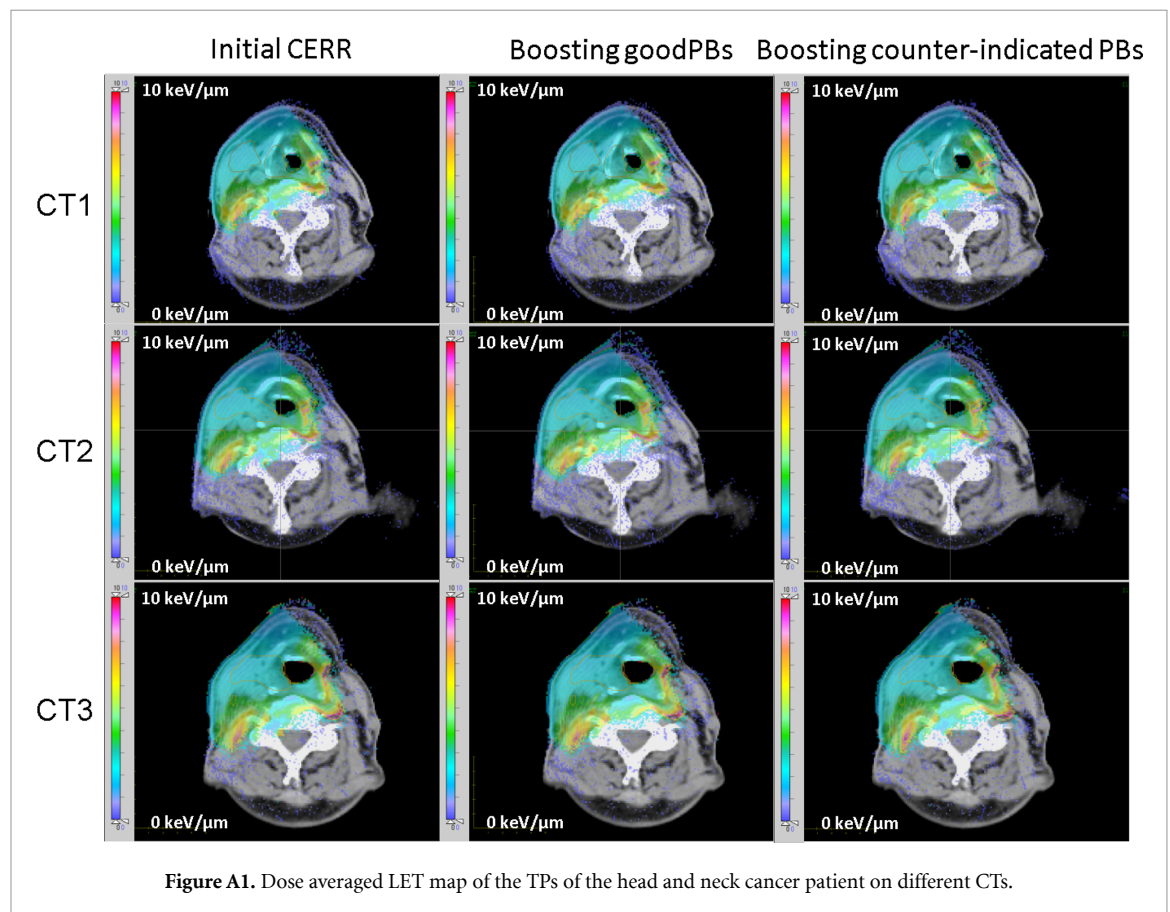


Figure A1. Dose averaged LET map of the TPs of the head and neck cancer patient on different CTs.

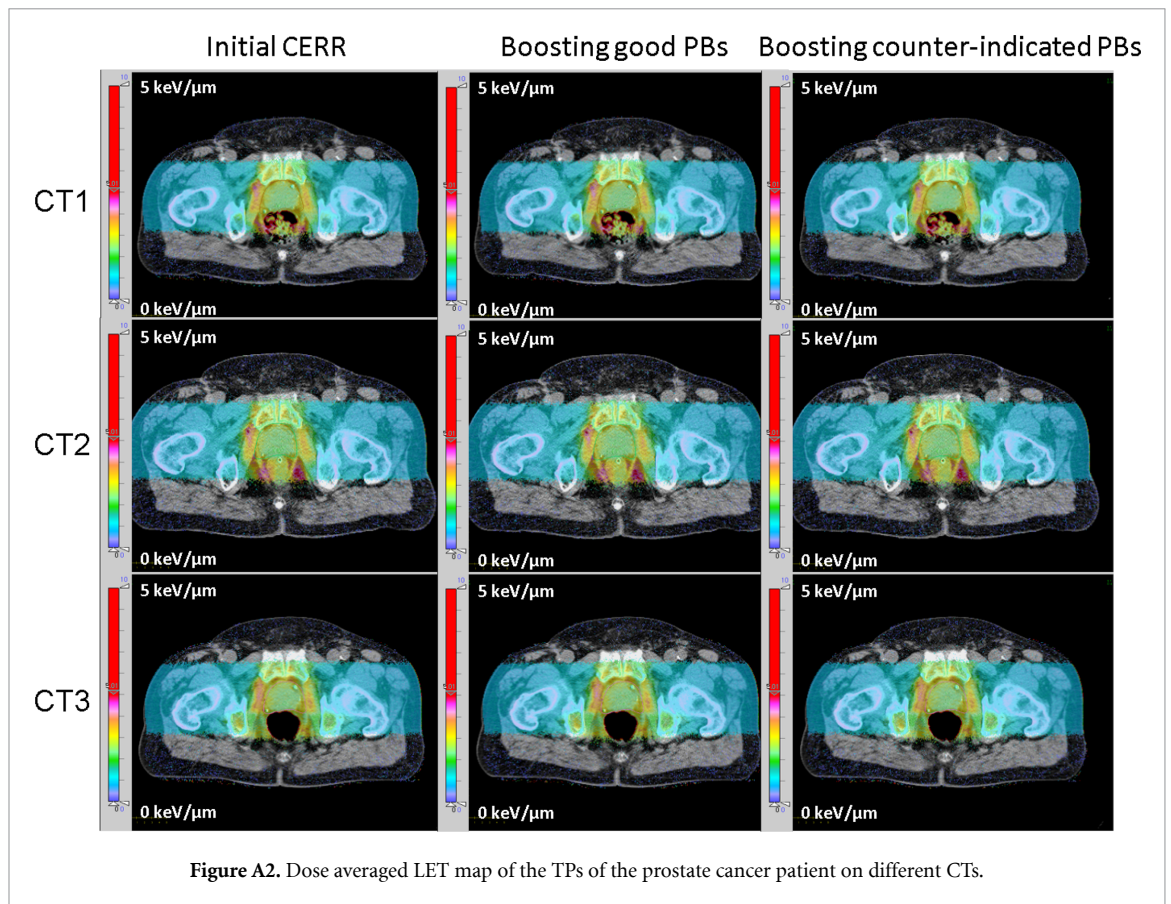
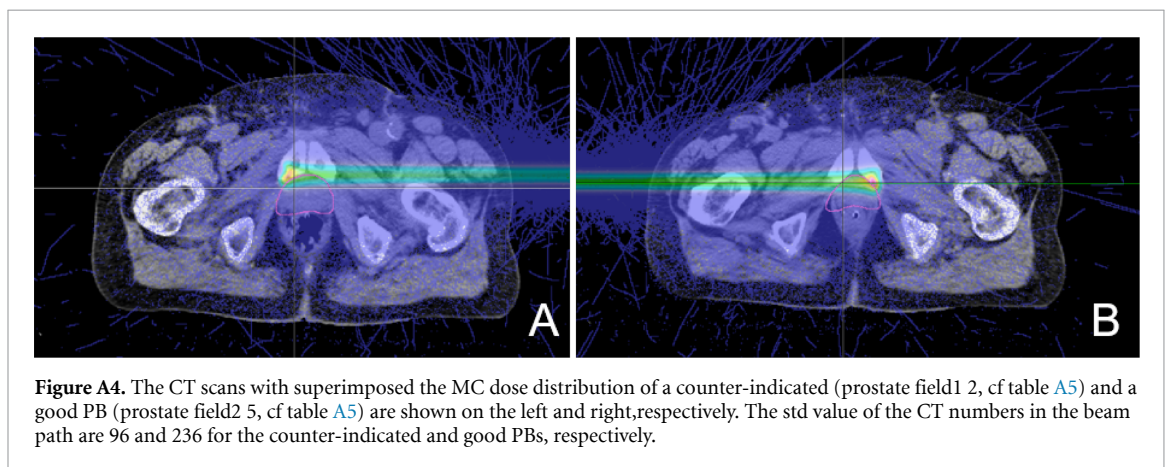
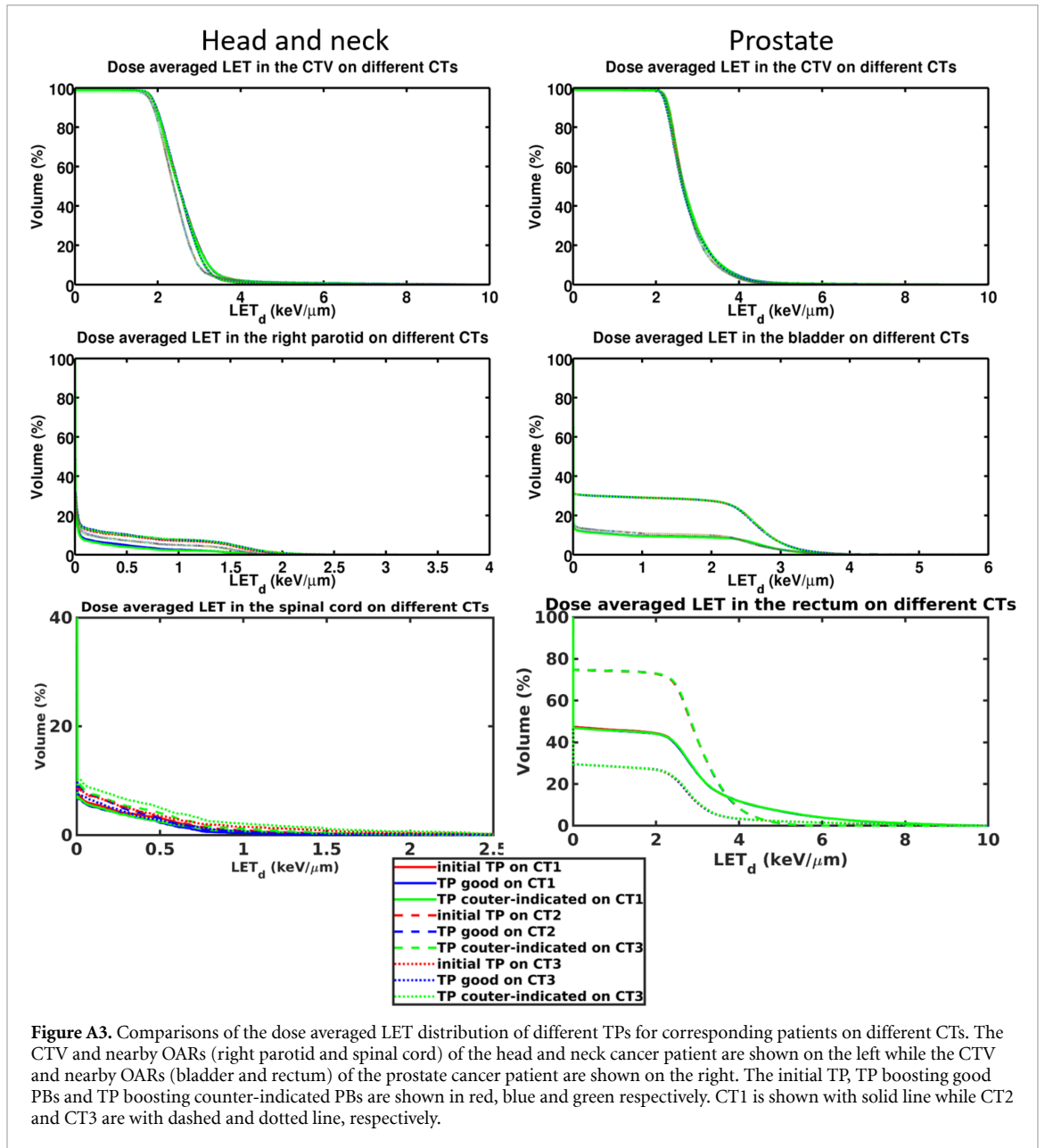


Table A5. Tissue heterogeneities in the beam path for the good and counter-indicated PBs (mean value \pm std value).

PB id (good)	Heterogeneities (HU)	PB id (counter-indicated)	Heterogeneities (HU)
Prostate field1 1	50 ± 192	Prostate field1 1	38 ± 177
Prostate field1 2	40 ± 186	Prostate field1 2	18 ± 96
Prostate field1 3	29 ± 123	Prostate field1 3	48 ± 243
Prostate field1 4	43 ± 137	Prostate field1 4	-3 ± 225
Prostate field1 5	47 ± 130	Prostate field1 5	17 ± 197
Prostate field2 1	25 ± 134	Prostate field2 1	29 ± 154
Prostate field2 2	39 ± 181	Prostate field2 2	-26 ± 291
Prostate field2 3	43 ± 151	Prostate field2 3	50 ± 199
Prostate field2 4	23 ± 129	Prostate field2 4	40 ± 134
Prostate field2 5	71 ± 237	Prostate field2 5	-26 ± 219
Head and neck 1	5 ± 140	Head and neck 1	-185 ± 432
Head and neck 2	28 ± 156	Head and neck 2	-123 ± 326
Head and neck 3	0 ± 136	Head and neck 3	-26 ± 204
Head and neck 4	-34 ± 147	Head and neck 4	-47 ± 169
Head and neck 5	-32 ± 126	Head and neck 5	-40 ± 211



ORCID iD

Marco Pinto  <https://orcid.org/0000-0001-6835-2561>

References

- Deasy J O, Blanco A I and Clark V H 2003 CERR: A computational environment for radiotherapy research *Med. Phys.* **30** 979–85
- Draeger E, Mackin D, Peterson S, Chen H, Avery S, Beddar S and Polf J C 2018 3d prompt gamma imaging for proton beam range verification *Phys. Med. Biol.* **63** 533–6
- Engelsman M, Schwarz M and Dong L 2013 Physics controversies in proton therapy *Semin. Radiat. Oncol.* **23** 88–96
- Janssen F M, Landry G, Cambraia Lopes P, Dedes G, Smeets J, Schaart D R, Parodi K and Verhaegen F 2014 Factors influencing the accuracy of beam range estimation in proton therapy using prompt gamma emission *Phys. Med. Biol.* **59** 4427–41
- Janssens G, Jacques L, Orban de Xivry J, Geets X and Macq B 2011 Diffeomorphic registration of images with variable contrast enhancement *Int. J. Biomed. Imaging* **2011** 891585
- Kraan A C, van de Water S, Teguh D N, Al-Mangani A, Madden T, Kooy H M, Heijmen B J and Hoogeman M S 2013 Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range and setup errors *Int. J. Radiat. Oncol. Biol. Phys.* **87** 888–96
- Min C H, Kim C H, Youn M Y and Kim J W 2006 Prompt gamma measurements for locating the dose falloff region in the proton therapy *Appl. Phys. Lett.* **89** 183517
- Min C H, Zhu X, Grogg K, El Fakhri G, Winey B and Paganetti H 2014 A recommendation on how to analyze in-room PET for in vivo proton range verification using a distal PET surface method *Technol. Cancer Res. Treat.* **14** 320–5
- Müller B S, Duma M N, Kampfer S, Nill S, Oelfke U, Geinitz H and Wilkens J J 2015 Impact of interfractional changes in head and neck cancer patients on the delivered dose in intensity modulated radiotherapy with protons and photons *Phys. Med.* **31** 266–72
- Paganetti H 2012 Range uncertainties in proton therapy and the role of Monte Carlo simulations *Phys. Med. Biol.* **57** R99–117
- Priegnitz M, Helmbrecht S, Janssens G, Perali I, Smeets J, Vander Stappen F, Sterpin E and Fiedler F 2015 Measurement of prompt gamma profiles in inhomogeneous targets with a knife-edge slit camera during proton irradiation *Phys. Med. Biol.* **60** 4849–71
- Resch A F, Landry G, Kamp F, Cabal G, Belka C, Wilkens J J, Parodi K and Dedes G 2017 Quantification of the uncertainties of a biological model and their impact on variable rbe proton treatment plan optimization *Phys. Med.* **36** 91–102
- Richter C et al 2016 First clinical application of a prompt gamma based in vivo proton range verification system *Radiother. Oncol.* **118** 232–7
- Schell S and Wilkens J J 2010 Advanced treatment planning methods for efficient radiation therapy with laser accelerated proton and ion beams *Med. Phys.* **37** 5330–40
- Schmid S, Landry G, Thieke C, Verhaegen F, Ganswindt U, Belka C, Parodi K and Dedes G 2015 Monte Carlo study on the sensitivity of prompt gamma imaging to proton range variations due to interfractional changes in prostate cancer patients *Phys. Med. Biol.* **60** 9329–47
- Stichelbaut F and Jongen Y 2003 Verification of the proton beam position in the patient by the detection of prompt gamma-rays emission *39th Meeting of the Particle Therapy Co-Operative Group* pp 26–29
- Tian L, Landry G, Dedes G, Kamp F, Pinto M, Niepel K, Belka C and Parodi K 2018 Toward a new treatment planning approach accounting for in vivo proton range verification *Phys. Med. Biol.* **63** 215025
- Xie Y H et al 2017 Prompt gamma imaging for in vivo range verification of pencil beam scanning proton therapy *Int. J. Radiat. Oncol. Biol. Phys.* **99** 035019