Received: 6 March 2021 | Revised: 30 June 2021 | Accepted: 11 October 2021

DOI: 10.1002/mp.15336

RESEARCH ARTICLE

MEDICAL PHYSICS

Localization of anatomical changes in patients during proton therapy with in-beam PET monitoring: A voxel-based morphometry approach exploiting Monte Carlo simulations

Aafke Christine Kraan[1] | **Andrea Berti**[1,2] | **Alessandra Retico**[1] | **Guido Baroni**[3,4] | **Giuseppe Battistoni**[5] | **Nicola Belcari**[1,2] | **Piergiorgio Cerello**[6] | **Mario Ciocca**[3] | **Micol De Simoni**[7,8] | **Damiano Del Sarto**[1,2] | **Marco Donetti**[3] | **Yunsheng Dong**[5,9] | **Alessia Embriaco**[10] | **Veronica Ferrero**[6] | **Elisa Fiorina**[3,6] | **Marta Fischetti**[8,11] | **Gaia Franciosini**[7,8] | **Giuseppe Giraudo**[6] | **Francesco Laruina**[1,2] | **Davide Maestri**[3] | **Marco Magi**[11] | **Giuseppe Magro**[3] | **Carlo Mancini Terracciano**[7,8] | **Michela Marafini**[8,12] | **Ilaria Mattei**[5] | **Enrico Mazzoni**[5] | **Paolo Mereu**[6] | **Riccardo Mirabelli**[7,8,12] | **Alfredo Mirandola**[3] | **Matteo Morrocchi**[1,2] | **Silvia Muraro**[5] | **Alessandra Patera**[6] | **Vincenzo Patera**[8,11,12] | **Francesco Pennazio**[6] | **Angelo Rivetti**[6] | **Manuel Dionisio Da Rocha Rolo**[6] | **Valeria Rosso**[1,2] | **Alessio Sarti**[8,11,12] | **Angelo Schiavi**[8,11] | **Adalberto Sciubba**[11,12,13] | **Elena Solfaroli Camillocci**[7,8] | **Giancarlo Sportelli**[1,2] | **Sara Tampellini**[3] | **Marco Toppi**[11,13] | **Giacomo Traini**[8,12] | **Serena Marta Valle**[5] | **Francesca Valvo**[3] | **Barbara Vischioni**[3] | **Viviana Vitolo**[3] | **Richard Wheadon**[6] | **Maria Giuseppina Bisogni**[1,2]

[1] Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Pisa, Italy

[2] Dipartimento di Fisica, Università di Pisa, Pisa, Italy

[3] Centro Nazionale di Adroterapia Oncologica, Pavia, Italy

[4] Politecnico di Milano, Milano, Italy

[5] Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy

[6] Istituto Nazionale di Fisica Nucleare, Sezione di Torino, Torino, Italy

[7] Dipartimento di Fisica, Sapienza Università di Roma, Roma, Italy

[8] Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Roma, Italy

[9] Dipartimento di Fisica, Università di Milano, Milano, Italy

[10] Istituto Nazionale di Fisica Nucleare, Sezione di Pavia, Pavia, Italy

[11] Dipartimento di Scienze di Base e Applicate per l'Ingegneria, Sapienza Università di Roma, Roma, Italy

[12] Museo Storico della Fisica e Centro Studi e Ricerche "E. Fermi", Roma, Italy

[13] Istituto Nazionale di Fisica Nucleare, Sezione dei Laboratori di Frascati, Frascati, RM, Italy

Correspondence
Alessandra Retico, INFN, Sezione di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy.
Email: alessandra.retico@pi.infn.it

Funding information
Italian Ministry of Education, Grant/Award Number: PRIN MIUR 2010P98A75 INSIDE; Italian Institute of Nuclear Physics, Grant/Award Numbers: RDH, INFN-RT2 PETRA 172800; Historical Museum of Physics and the Enrico Fermi Study and Research Center and the; Tuscany Government, Grant/Award Numbers: POR FSE 2014–2020, INFN-RT2 PETRA 172800; CNAO Foundation, Grant/Award Number: INSIDE2

Abstract

Purpose: In-beam positron emission tomography (PET) is one of the modalities that can be used for in vivo noninvasive treatment monitoring in proton therapy. Although PET monitoring has been frequently applied for this purpose, there is still no straightforward method to translate the information obtained from the PET images into easy-to-interpret information for clinical personnel. The purpose of this work is to propose a statistical method for analyzing in-beam PET monitoring images that can be used to locate, quantify, and visualize regions with possible morphological changes occurring over the course of treatment.

Methods: We selected a patient treated for squamous cell carcinoma (SCC) with proton therapy, to perform multiple Monte Carlo (MC) simulations of the expected PET signal at the start of treatment, and to study how the PET signal may change along the treatment course due to morphological changes. We performed voxel-wise two-tailed statistical tests of the simulated PET images, resembling the voxel-based morphometry (VBM) method commonly used in neuroimaging data analysis, to locate regions with significant morphological changes and to quantify the change.

Results: The VBM resembling method has been successfully applied to the simulated in-beam PET images, despite the fact that such images suffer from image artifacts and limited statistics. Three dimensional probability maps were obtained, that allowed to identify interfractional morphological changes and to visualize them superimposed on the computed tomography (CT) scan. In particular, the characteristic color patterns resulting from the two-tailed statistical tests lend themselves to trigger alarms in case of morphological changes along the course of treatment.

Conclusions: The statistical method presented in this work is a promising method to apply to PET monitoring data to reveal interfractional morphological changes in patients, occurring over the course of treatment. Based on simulated in-beam PET treatment monitoring images, we showed that with our method it was possible to correctly identify the regions that changed. Moreover we could quantify the changes, and visualize them superimposed on the CT scan. The proposed method can possibly help clinical personnel in the replanning procedure in adaptive proton therapy treatments.

KEYWORDS
in-beam PET monitoring, proton therapy, voxel-based morphometry

1 | INTRODUCTION

Particle therapy is the treatment of solid tumors by charged particles, mostly protons or carbon ions. Thanks to their characteristic dose deposition profile (Bragg peak), charged particles deposit their dose much more locally than the conventionally used photons, allowing for a better sparing of healthy tissue.[1] However, the steep dose gradients also make it more sensitive to errors in dose delivery.[2,3] Among the main error sources are interfractional patient variations, occurring because of setup errors and anatomical modifications in patients during the treatment course.[3,4] Such errors can possibly compromise the accuracy and affect therapeutic outcomes of particle therapy treatments in particular.

For some patient groups, like head-and-neck patients, typically one or more control computed tomography (CT) scans are made along the course of treatment in order to ensure that the treatment plan is still appropriate.[3–9] This patient group is particularly prone to interfractional morphological changes mostly from presence of cavities, tumor regression, inflammation, organ motion, and weight loss.[5,8,9] If the plan is judged no longer appropriate because of dose deteriorations, a new plan is created ("replanning" or "adaptive planning").

The control CTs are typically made at predefined intervals over the course of treatment, driven by clinical experience for a given patient group and personal assessment of the radiation oncologist. However, given the wide variability in morphological changes between patients throughout radiation therapy,[5] the need for dose optimization for each patient, and the time- and resource-consuming nature of adaptive interventions, it would be preferred to decide in an automated manner when to

perform the control CT and eventually adapt the plan.[4] Undoubtedly, cone beam computed tomography (CBCT) can provide valuable information about the anatomy and setup of the patient before treatment,[10,11] however it does not give indications about possible dose modifications. Online imaging techniques, that is, techniques that can be used during dose delivery, that allow to raise red flags in case of suspected morphological changes associated with dose modifications are thus highly desired to pursue patient-tailored treatment optimization.[3]

In this context, various noninvasive in vivo treatment monitoring techniques have been proposed in the past decades that are based on detecting reaction products of nuclear interactions of the beam in the patient, including PET, prompt gammas, and charged fragment detection (see Refs. 12–14 and references therein). Treatment monitoring with positron emission tomography (PET) imaging is based on the measurement of the β^+ activity, induced by nuclear interactions between the beam and the patient tissue. Typically, measured PET images are compared with predicted images, either from Monte Carlo (MC) simulations or from previous measurements, to estimate whether the dose was delivered correctly or whether a new repeat CT is needed.

Among all PET data acquisition modalities, in-beam PET monitoring is the only option allowing to assess the dose delivery in real time.[12,15–18] In-beam PET imaging is based on detecting the β^+ activity from isotopes like ^{15}O (half life: 122.24 s), ^{13}N (half-life: 2.84 min), ^{10}C (half life: 19.29 s), and ^{11}C (half life: 20.334 min). At the National Center for Hadrontherapy (CNAO) in Pavia,[19] currently a dual imaging modality system (an in-beam dual head PET scanner and a charged fragment tracker) named INSIDE (INovative Solutions for In-beam DosimEtry in hadrontherapy) is being used.[20] INSIDE is currently in a clinical trial phase to test the clinical feasibility and effectiveness in detecting morphological changes in head-and-neck patients.

Despite the broad interest in PET monitoring, there is still no straightforward method to translate the information obtained from the PET images acquired during a treatment session into useful and easy-to-interpret information for clinical personnel. In fact, the obtained PET data, arising from nuclear physics processes, cannot be directly related to the delivered dose, determined mostly by electromagnetic processes. Moreover PET data cannot provide direct information about changes in the anatomy of the patient. Various methods to solve this issue have been proposed. The first attempts focused on translating measured activity into dose.[21–24] Later, several empirical methods were developed to detect variations in particle range by studying the activity profile, including the middle-point analysis,[25] the shift approach,[26] the most-likely-shift approach,[27] and range evaluation based on iso-activity surface extraction.[28–30] The latter has resulted in the clinical validation of a data-driven method to identify the pencil beams that pass through regions where morphological changes were identified using the INSIDE PET data.[30]

The goal of the current work is to propose an innovative and fundamentally different approach to apply to in-beam PET monitoring images, that is inspired by the voxel-based morphometry (VBM) method widely used in neuroimaging. In contrast to all previously published works, the proposed method does not consider particle range, but gives clear voxel-by-voxel statistical information about the probability that a morphological changes occurred. We test the method on a series of simulated in-beam PET monitoring images based on the INSIDE system, with artificially introduced morphological changes.

2 | MATERIALS AND METHODS

2.1 | Patient data

This work included data from a 70-year-old patient treated in 2018 for squamous cell carcinoma (SCC) with proton therapy at CNAO.[19] A planning CT scan was available with corresponding structure set. The resolution of the CT was $0.98 \times 0.98 \times 2$ mm^3. Figure 1(a) displays a slice of the planning CT scan, with the clinical target volume (CTV) highlighted. The CTV was divided in a low dose region (CTV-low) receiving 60 Gy in 30 fractions and a high dose region (CTV high) receiving 66 Gy in 33 fractions. Figure 1(b) shows the same CT slice, but now with the region-of-interest, including the sinonasal cavity, highlighted. An intensity modulated proton therapy (IMPT) treatment plan based on the planning CT scan was available (Syngo Treatment Planning system version VB10, Siemens), which had three orthogonal irradiation fields. The field at 270° IEC (International Electrotechnical Commission) was selected for this work, delivering 2.1×10^{10} protons of the total number of 7.0×10^{10}. The relevant characteristics of the treatment plan are summarized in Table 1.

Furthermore a first control CT, taken after two fractions, and a second control CT, taken after 22 fractions, were available. A slice of the first and second control CT is displayed in Figure 1(c) and (d), respectively. While the morphological changes seen in the first control CT were minimal and did not require adaptive interventions, the second control CT demonstrated large morphological changes caused by sinonasal cavity emptying, that resulted in serious dose modifications requiring replanning. A new treatment plan was thus created and applied starting from fraction 23.

2.2 | INSIDE system

The INSIDE system is a bimodal in-beam treatment monitoring system combining an in-beam PET system

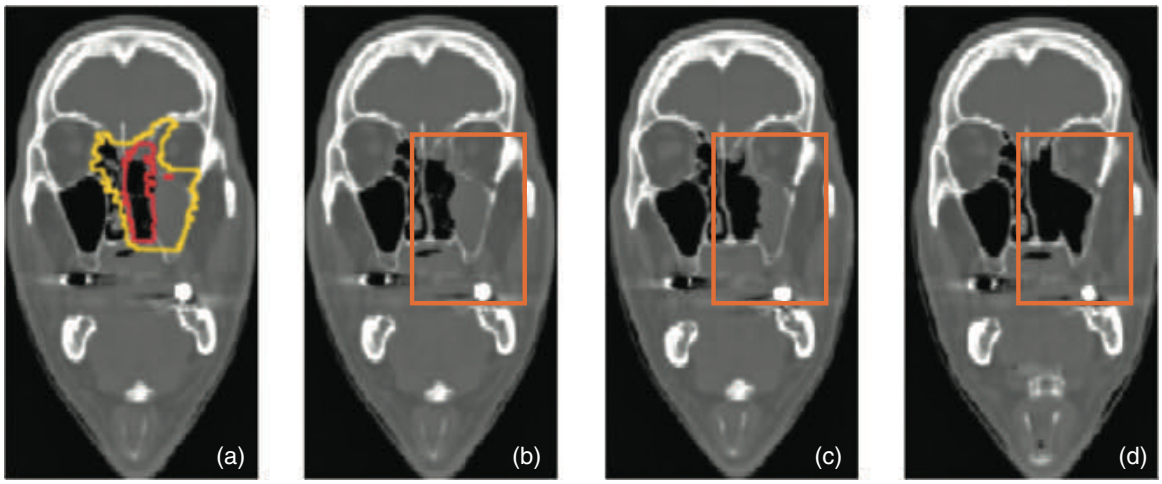

FIGURE 1 (a) Slice of the planning CT of the SCC patient with the CTV-low in yellow and the CTV-high in red. (b) Same, but now without target structures and with the region of interest, indicated as orange rectangle, containing the sinonasal cavity. (c) Corresponding slice of the first control CT. The sinonasal cavity is somewhat emptied, as can be seen by inspecting the region inside the orange rectangle. (d) Corresponding slice of the second control CT, where further emptying of the sinonasal cavity is visible

TABLE 1 Relevant treatment plan characteristics of the SCC patient. Np represents the number of particles

Treatment modality	IMPT
Fractions	33
Fields	270° IEC, 0° IEC, 180° IEC
Tumor volume	CTV-low: 139 mL
	CTV-high: 40 mL
Dose prescription	CTV-low: 60 Gy
	CTV-high: 66 Gy
Dose per fraction	CTV-low: 2 Gy
	CTV-high: 2 Gy
Field 1 nr of protons	2.1×10^{10} Np
Field 1 minimum energy	66 MeV
Field 1 maximum energy	168 MeV
Average duration of field 1	232 s

and a secondary charged particle tracker. This work focuses on the PET system, installed since 2019 at CNAO.[20] It consists of two PET heads with a planar geometry and an active area of 10×25 cm^2 each. Each head has 2×5 detection modules, each made of a 16×16 matrix of lutetium fine silicate (LFS) scintillating crystals, with $3.1 \times 3.1 \times 20$ mm^3 size and 3.2 mm pitch. The readout electronics consists of a 16×16 array of Hamamatsu Silicon Photomultiplier (SiPM) coupled one-to-one to each crystal. The resulting field of view (FOV) is $11.2 \times 22.4 \times 26.4$ cm^3 with a voxel size of $1.6 \times 1.6 \times 1.6$ mm^3, that is, $70 \times 140 \times 165$ voxels. Both the PET heads and the charged particle tracker are mounted on a submillimetric precise mechanical cart. With the current design, the INSIDE in-beam PET is compatible with most head and neck treatments, as well as brain treatments. At present the system is undergoing a clinical trial.[31]

2.3 | Analysis methodology

To determine whether and to which extent the effects of morphological changes occurring over the treatment course are detectable with the INSIDE PET system, a statistical analysis pipeline has been set up. While the current section is limited to the methodology, the technical details of the implementation follow in Section 2.4.

The procedure was inspired by the statistical parametric mapping (SPM) techniques, which are widely used in neuroimaging studies, mainly involving PET and magnetic resonance imaging (MRI) techniques[32] in the study of a large variety of disease conditions.[33–35] SPM maps aiming to highlight possible morphometric differences between two groups of subjects (e.g., subjects affected by a particular pathological condition and control subjects) can be generated according to the VBM approach.[36] In that case, three-dimensional brain scans of two groups of subjects, once properly spatially aligned, are compared at the single-voxel level according to a statistical model. Typically, the general linear model[32] is used as a parametric model to describe the voxel signals in terms of a number of predictors. Statistical tests (e.g., two-sample *t*-test) are conducted at each voxel, and a map of the *p*-value for each voxel to contain significant distinguishing information between the two cohorts is derived. This map, which has the same dimension of the original images, can be overlaid on the brain scan of a group-representative subject to highlight the brain areas where significant differences between the two populations have been detected. The SPM map thus contains an easily viewable information,

whose interpretation depends on the input data and on the applied preprocessing steps.

In our analysis, rather than comparing two groups of subjects, we examined data from a single subject, that changes over time. In particular, we wanted to examine whether a given observed PET treatment-monitoring image, affected by morphological changes, acquired during a certain fraction along the treatment course, was compatible within statistical fluctuations with the PET monitoring image acquired at the beginning of the treatment, that is, in absence of morphological changes. To this purpose, we set the problem as a null hypothesis statistical test, where empirical *p*-values were obtained with the help of MC simulations. We generated the null reference distribution of the voxel intensities (our test statistics) for the PET image in absence of morphological changes, by running the MC simulation of the treatment based on the planning CT with N random seeds. Moreover, we generated a series of observed[1] PET monitoring images, affected by morphological changes. Then, for each voxel the null hypothesis being tested was that the intensity value in that voxel in the observed monitoring PET image was not different from the null reference one, that is, without morphological changes. Rejection of null hypothesis implied that the voxel intensity had significantly changed.

In general, for each voxel in an observed image we can calculate the empirical probability value *p* that its intensity is compatible with the null hypothesis.[37–39] In our problem, the observed voxel intensity could depart from the null distribution in both directions, that is, it could either be more than or less than the average simulated voxel intensity. In fact, a range overshoot, caused for instance by a region that is emptied such as a cavity, could result in less activity than expected in the emptied region, and more activity behind it. A two-tailed implementation of the null hypothesis statistical test was thus necessary. In this case, the null hypothesis H_0 was rejected when the *p*-value was sufficiently extreme and thus unlikely to be the result of chance. When providing a significance level α, the critical regions would exist on the two tail ends of the 1-D distribution of voxel intensity values. Let us define v^i_{sim} as the intensity value in voxel *i* in a PET image, $\bar{v}^i_{sim}$ as the average intensity value over all *N* replicates, and $v^i_{obs_j}$ as the intensity value in the observed PET image *j*. The empirical *p*-values at each voxel *i* could then be computed as [37–39]:

$$p_i = \frac{r_i + 1}{N + 1}, \tag{1}$$

where *N* was the number of replicates and r_i was the number of replicates that produced a test statistic of either $v^i_{sim} \geq v^i_{obs_j}$ if $v^i_{obs_j} \geq \bar{v}^i_{sim}$, or $v^i_{sim} \leq v^i_{obs_j}$ if $v^i_{obs_j} < \bar{v}^i_{sim}$.

The resulting p_i-values evaluated for each voxel *i* of the *j*th PET image can be reported as 3-D voxelized maps. In particular, by setting the conventional significant threshold $\alpha = 0.05$, two separate masks were obtained, highlighting only the voxels for which the null hypothesis was rejected in favor of the alternative hypothesis, that is, that $v^i_{obs_j}$ was either significantly larger or significantly smaller than the reference intensity. These could be overlaid onto the CT of the subject. The two masks indicated the areas, where the activity was less-then-expected or more-than-expected with respect to the reference treatment situation.

The proposed statistical analysis can thus be summarized as follows:

- Obtain the null reference distribution of intensities for each voxel in the PET image by repeating the simulation of the treatment on the planning CT *N* times with different random seeds.
- Generate a series of observed PET images with MC, simulating the PET signal from intermediate treatment fractions *j* based on artificially modified CT scans. As mentioned before, we stress that these observed images are generated by MC simulations. Although the method should naturally be extended to observed PET images from experimental data, for the scope of validating the statistical method and assessing its sensitivity in detecting morphological changes, the simulated observed PET images are sufficient.
- For each observed PET image *j*, do a voxel-wise two-tailed null hypothesis test, yielding for each voxel a *p*-value. Significant *p*-values, that is, $p \leq \alpha/2 = 0.025$, can be visualized with appropriate colors for the case where $v^i_{obs_j} \geq \bar{v}^i_{sim}$ (significantly more activity than expected) and $v^i_{obs_j} \leq \bar{v}^i_{sim}$ (significantly less activity than expected), respectively. These colored 3D maps can be overlaid onto the original CT image, allowing directly to localize the areas where changes in patient's anatomy occurred, that may have caused an over/under dose exposure.
- Apply the above method to the real first and second control CT of the patient that were available, and identify the regions that changed.

2.4 | Technical implementation

2.4.1 | MC simulations of null reference distributions of PET voxel intensities

The null reference distributions of PET intensities in each voxel were obtained from MC simulations by running the same patient simulation many times

[1] We use the term "observed" throughout this work in order to be in line with the terminology adopted in statistical analyses, despite that the PET images including morphological changes were generated with MC simulations.

independently. The MC simulation tool developed for INSIDE[28,40] was used for this purpose. The tool is based on the FLUKA MC code, suited for particle therapy treatment plan simulations.[41] The following issues were considered in the simulation of a treatment:

- Beam line. The last part of the CNAO beam line, including exit window and nozzle were simulated. The pencil beam transverse dimension and shape were reproduced as in measurements for all pencil beam energies. This was done with a previously developed simulation FLUKA tool, that was also used for the dosimetric commissioning and quality assurance of proton and carbon beams at CNAO.[42–44]
- Time structure of the delivered protons, including time, intensity, and duration of each single spot.[45]
- Patient geometry. This was obtained by importing the CT scan with its original resolution in a voxel FLUKA geometry. Tissue composition and densities were determined according to the CT Hounsfield Units method by Schneider.[46]
- Radiation transport and interactions of protons in the patient. Although it was time consuming, we used the maximum precision defaults of the physics models options of FLUKA, and the energy threshold for charged particle transport was set to 100 keV.
- Generation and decay of the 3-D spatial and time distribution of β^+ emitting isotopes. The slowest part of the simulation was the proton transport from the beam line to the production of the β^+emitter inside the patient. Therefore a two-step strategy was implemented.[40] In the first step, a fraction $f = \frac{1}{100}$ of the entire statistics of the treatment field, that is, $f \times 2.1 \times 10^{10} = 2,1 \times 10^8$ protons, was simulated up to the production of the β^+ emitter. A user scoring routine in FLUKA was adopted to extract the spatial coordinates and production time of the β^+emitting isotope (^{11}C,^{15}O, and so on). This collection of isotopes is referred to as the isotope production map (IPM). Then, in the second step, the IPM was used as positron generator. Each produced β^+ emitting isotope was used $1/f$=100 times to decay into a positron according to the specific β^+ decay spectrum of the given isotope and with random direction. The acollinearity angle of the photons was also considered. Accurate photon transport in the patient was simulated down to 100 keV.
- Geometry of the INSIDE PET detector, including the individual pixelated LSF crystals. The patient geometry with the PET planes is given in Figure 2.
- Energy depositions of the pairs of 511 keV photons in the detector. Coincidences were selected, that occurred between the start of treatment up to about 6 min.[30] Previously developed software was used to model energy depositions in the detector, and energy and time resolutions were included by Gaussian smearing.[40]
- Creation of line-of-responses (LORs) and subsequent image reconstruction. This was done with an iterative maximum likelihood estimation maximization (MLEM) procedure.[47]

A 3-D PET image was obtained, with a FOV of $22.4 \times 11.2 \times 26.4$ cm^3, with $1.6 \times 1.6 \times 1.6$ mm^3 voxels, just like in data (see Section 2.2). Due to the partial angular coverage of the detector, the obtained images suffered from reconstruction artefacts in the direction perpendicular to the two PET detector planes. We applied a median filter of 1.6 mm (1 voxel) in all directions in order to decrease statistical fluctuations.

Since the MC simulations of an entire treatment are computationally heavy, we relied on parallelization. A number of $N = 120$ independent replicates was produced, each time changing the random number in the FLUKA simulation. This resulted in 120 IPM and 120 corresponding PET images, making up the null distribution in all voxels. These will be referred to as IPM$_0$ and PET$_0$ images, respectively. An illustration of the procedure to create the null distribution for one of the voxels in the PET image is given in Figure 3.

2.4.2 | MC simulation of PET images including morphological changes

We had to estimate the response of the in-beam PET monitoring images to the day-by-day evolution of the emptying of the sinonasal cavity between the planning CT and the control CTs. For this purpose, we created a set of observed PET images representing various situations with morphological changes. Thus, a set of artificial CTs had to be created, that mimicked this evolution. The control CTs were used as qualitative references for mimicking volumes and directions of this process. An in-house developed image processing tool in Python was used to modify the CT scans as follows.

First, we used rigid coregistration to align the planning CT with the first and the second control CT. The region where emptying of the sinonasal cavity occurred was identified, referred to hereafter as cavity emptying region. Then, in a first layer of voxels in the sinonasal region, the tissue was replaced with air (Hounsfield Unit= −1000) in the control CT, and an intermediate CT scan was generated. This was repeated gradually, emptying step-wise the entire left sinonasal cavity, each time generating a new intermediate CT scan, up to the moment that the cavity was empty. This latter situation was very similar to the situation of the second control CT. The planning CT together with three examples of the intermediate CTs obtained with the above procedure are given in Figure 4.

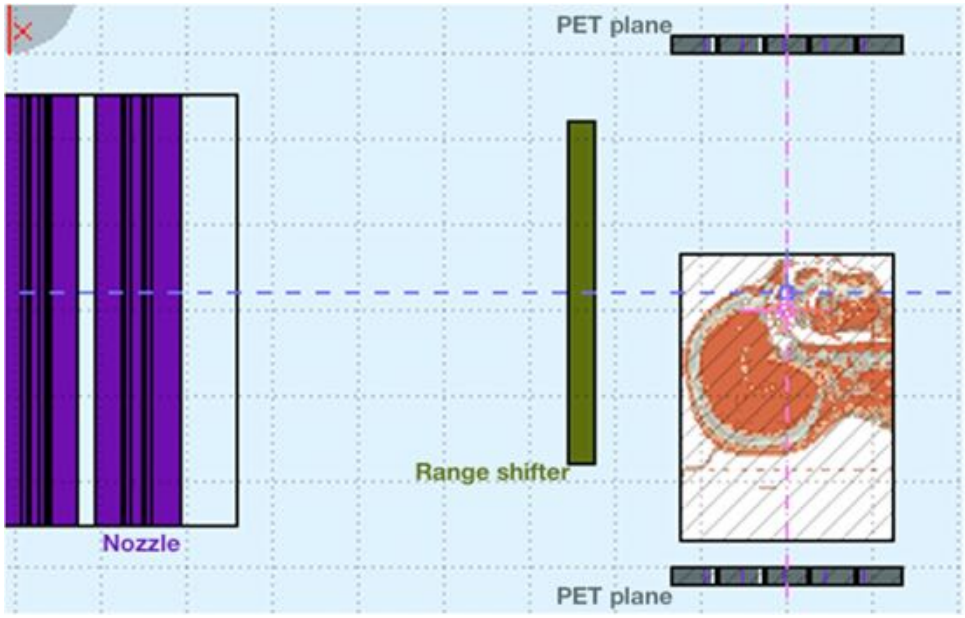

FIGURE 2 Simulation geometry with beam line (purple), range shifter (green), patient and PET planes, obtained by using the FLAIR graphical interface to FLUKA[53]

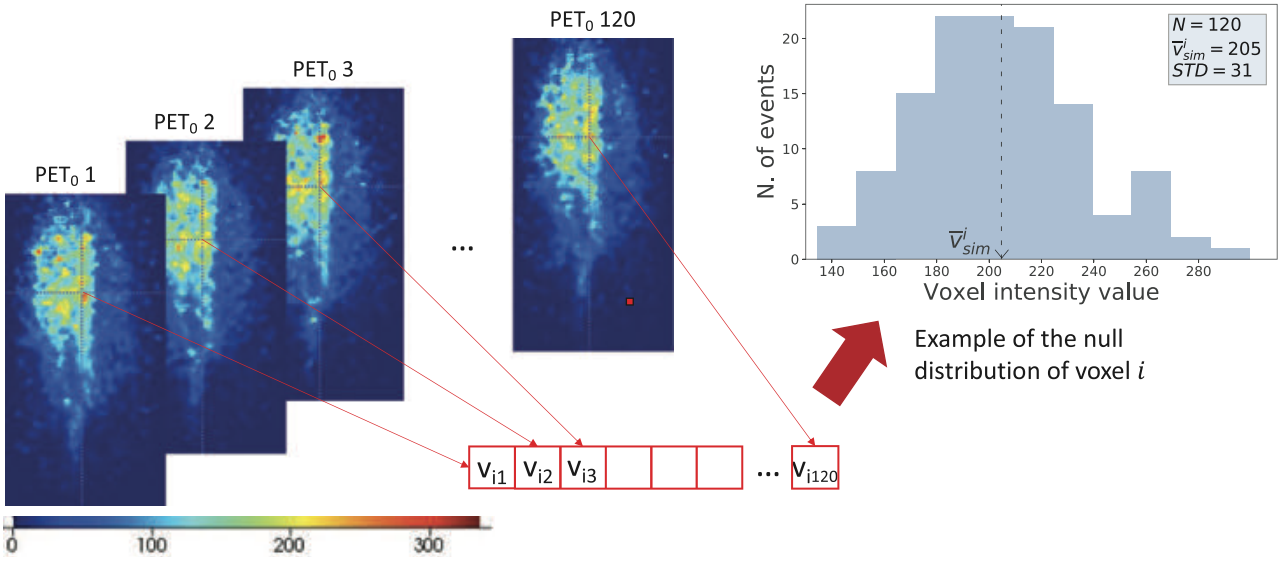

FIGURE 3 Illustration of procedure to create the null distribution in a voxel, starting from 120 PET_0 images, each of which is created with a unique random seed. The color scale used to represent the voxel intensities is also displayed

In total, six intermediate treatment sessions (j= 1 to 6) were simulated. A list of the reduction in tissue volume in the cavity corresponding to each generated intermediate CT is reported in Table 2. Thus, the cavity in the modified CTs evolves from being totally filled (0 mL change in tissue, i.e., no morphological modifications, first table entry) to being totally emptied (13 mL change in tissue, i.e., large morphological changes, last table entry). The sample with $j = 0$ will be referred to hereafter as the control sample. Regarding the size of the volume that was changed, we remind that the CT image voxels had a physical volume of 0.002 mL ($0.98 \times 0.98 \times 2$ mm^3), so the number of CT voxels that were modified ranged from zero to a several thousands (13.1/0.002) voxels when the cavity was emptied. In terms of PET voxels ($1.6 \times 1.6 \times 1.6$ $mm^3 = 0.004$ mL), a complete emptying of the cavity would correspond to about $3{\times}10^3$ voxels (13.1/0.004).

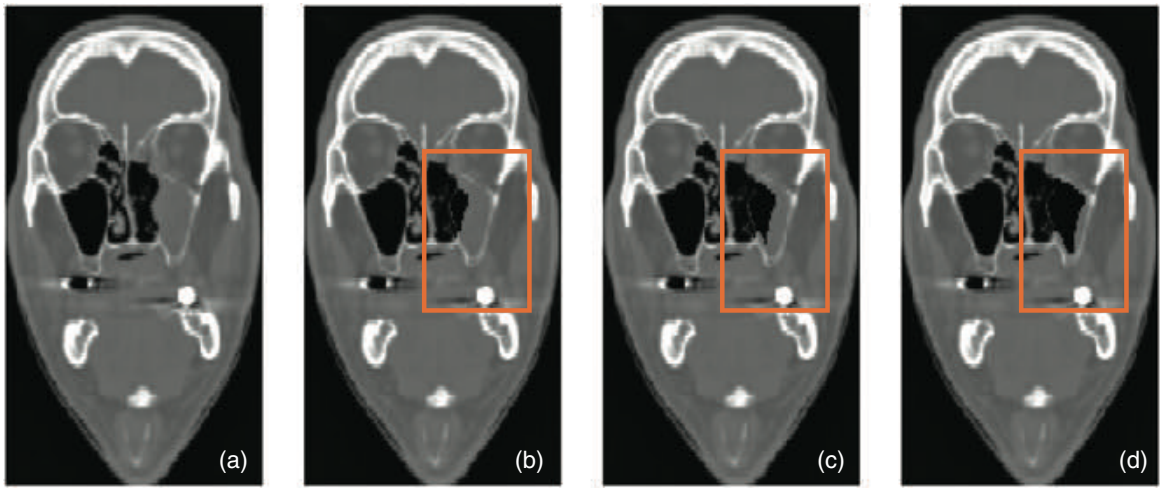

FIGURE 4 Examples of artificially created CTs with different sizes of cavity tissue volume reductions: 0 mL (the planning CT, (a)), 3.8 mL (b), 7.3 mL (c), and 13.1 mL (d). The region of interest containing the left sinonasal cavity is highlighted with an orange rectangle

TABLE 2 List of the generated intermediate CTs, with the corresponding absolute and relative reduction in tissue volume in the cavity

mod. CT *j*	Δ*V*(mL)	Δ*V*(%)
0	0	0
1	1.9	15
2	3.8	29
3	5.7	44
4	7.3	56
5	10.2	78
6	13.1	100

Using the same framework as discussed in Section 2.4.1, for each volume change from Table 2 the IPM and PET images were generated and stored. We will refer to each observed image *j* as $IPM_{obs,j}$ and $PET_{obs,j}$ images, respectively.

To estimate the statistical error in the evaluation of the volume changes due to the MC simulations, the latter were repeated 10 times on each artificially modified CT.

Besides the PET distributions based on the artificially modified CTs, we also simulated the PET signal for the first (no replanning) and second (replanning) control CT. Rigid registration was applied before we simulated the PET signal. These images will be referred to as "PET_{CT1}" and "PET_{CT2}."

2.4.3 | Voxel-wise null hypothesis test

Using the null reference distributions of PET intensities, we calculated for each voxel in each observed monitoring PET image, $PET_{obs,j}$, from Table 2 the *p*-value as in Equation (1). The significant threshold was $\alpha = 0.05$, so for either tail it was $\alpha/2 = 0.025$. For each PET image *j* the two-tailed statistical test lead to two voxel-wise *p*-values maps identifying voxels whose intensity values were unlikely to occur by chance:

- a blue map indicating voxels whose intensity value was classified as significantly low with respect to the null reference distribution;
- a red map indicating voxels whose intensity value was classified as significantly high with respect to the null reference distribution.

The same was done for the PET_{CT1} and PET_{CT2} images.

2.5 | Assessment of the capability to identify regions with morphological changes

2.5.1 | Qualitative assessment

To validate that the statistical method can correctly identify regions where morphological changes occurred, the 3-D distribution of obtained *p*-values related to each intermediate CT *j* was overlaid onto the planning CT. The red and blue areas of the maps were then compared to the artificially modified CT to confirm that the regions identified as significantly different were indeed related to the regions where morphological changes occurred. The same was done for the PET_{CT1} and PET_{CT2} images.

2.5.2 | Quantitative assessment

We also checked whether the method can give quantitative information about the morphological changes. Knowing that the emptying of the cavity should result in

voxels with an intensity value that is significantly reduced with respect to expectations, we investigated whether the size of the regions identified in the blue map were related to the size of the modified cavity region in the planning CT. Thus, for each of the observed PET monitoring images $PET_{obs,j}$ ($j = 1$ to 6) plus the control sample ($j = 0$) we compared the volume that was actually changed (emptied) in the CT scan, V_{mod}^{CT}, with the size of the volume that was classified as significant in the blue (lack of activity) color map of the PET image analysis, V_{obs}^{PET}. The latter was determined by summing the volume of the PET image voxels in the blue p-value map, that fall within the CTV-low. The value for V_{mod}^{CT} was determined by summing the volume of the CT voxels, whose densities were changed to air, and that were contained inside the CTV-low.

2.6 | Influence of PET image artifacts

We also assessed to what extend the method was affected by the typical image artifacts, that were present in our PET images due to the partial angular coverage of the detector and limited statistics. This was done by repeating the analysis from Section 2.5, but using the 3-D IPM maps rather than the PET images, that is, an ideal theoretical situation without detector and image reconstruction effects. Regarding the quantitative analysis in Section 2.5.2, we now compared the size of the volume that was classified as significant in the blue color map of the IPM image analysis, V_{obs}^{IPM}, with V_{mod}^{CT}, calculated as for the PET image analysis. The obtained results were compared with results that were based on the analysis of the PET images.

2.7 | Uncertainty analysis

In order to verify the robustness of the analysis results, the assessments in Sections 2.5 and 2.6 were repeated for each of the 10 repetitions of the simulations performed for each tissue volume change in Table 2. Each time the mean and standard deviation was evaluated. Furthermore, the impact of various additional error sources was evaluated for the case with the largest volume change of Table 2, 13.1 mL, including the following:

- Experimental uncertainties in the cross sections for β^+ emission in the MC simulation. These were changed by +10% and −10% and the analysis of Section 2.5 was repeated. This represents an extremely pessimistic situation, where all cross sections are changed in one direction.
- Uncertainties in the number of histories in the MC simulations. We varied the fraction f of the number of protons in the MC simulation (see Section 2.4.1), which was 1/100. The simulation was repeated with $f = 1/60$ and the analysis of Section 2.5 was repeated.
- The position of the INSIDE dual head PET chart with respect to the patient. In all relevant directions the chart was moved with 1.6 mm and the analysis of Section 2.5 was repeated.

3 | RESULTS

3.1 | Visualization of PET and IPM images

We show in Figure 5 an example of a simulated PET image in case of no sinonasal emptying (one of the 120 replicates) overlaid onto the planning CT in the coronal (a), axial (b), and sagittal (c) planes. The corresponding IPM images are also displayed in (d), (e), and (f), respectively. A few observations can be made. First, strong deformation in the axial and coronal planes can be seen in the PET image, which is a result of the partial angular coverage of the detector. The coronal plane, parallel to the PET planes, is the least affected and therefore in the following selected for visualization of the analysis results. Second, the maximum activity generated was not located in the target region, but much in front with respect to the beam entrance, as expected (see for instance Refs. 25, 30).

3.2 | Identification of regions with morphological changes: Qualitative assessment

In Figure 6 we display examples of the simulated PET signal (one of the 10) for volume changes of 0 mL (a), 3.8 mL (b), 7.3 mL (c), and 13.1 mL (d), overlaid onto the planning CT. In fact, this would represent the kind of information that clinical personnel would observe if looking directly at the PET signal. However, the insets at the left bottom of Figure 6(b), (c), and (d) represent how the situation in the CT actually changed. On one hand, by looking carefully at the PET image, we see that the distribution somewhat changes in the region-of-interest: for instance, the activity in that region is somewhat enlarged in Figure 6(c) and (d) with respect to Figure 6(a) and (b). Thus, we confirm that the activity signal is sensitive to morphological changes. On the other hand, it is very difficult to interpret the information: It is not feasible to estimate the location and cause of the change in the PET signal. It must be noted that simply subtracting the expected PET images from an observed PET image is not appealing, as shown already in various works.[48,49] Since the significant differences in PET signal are often located in regions where the activity is only moderate, statistical fluctuations cause the highlighted regions to appear all over the activated zone, mostly

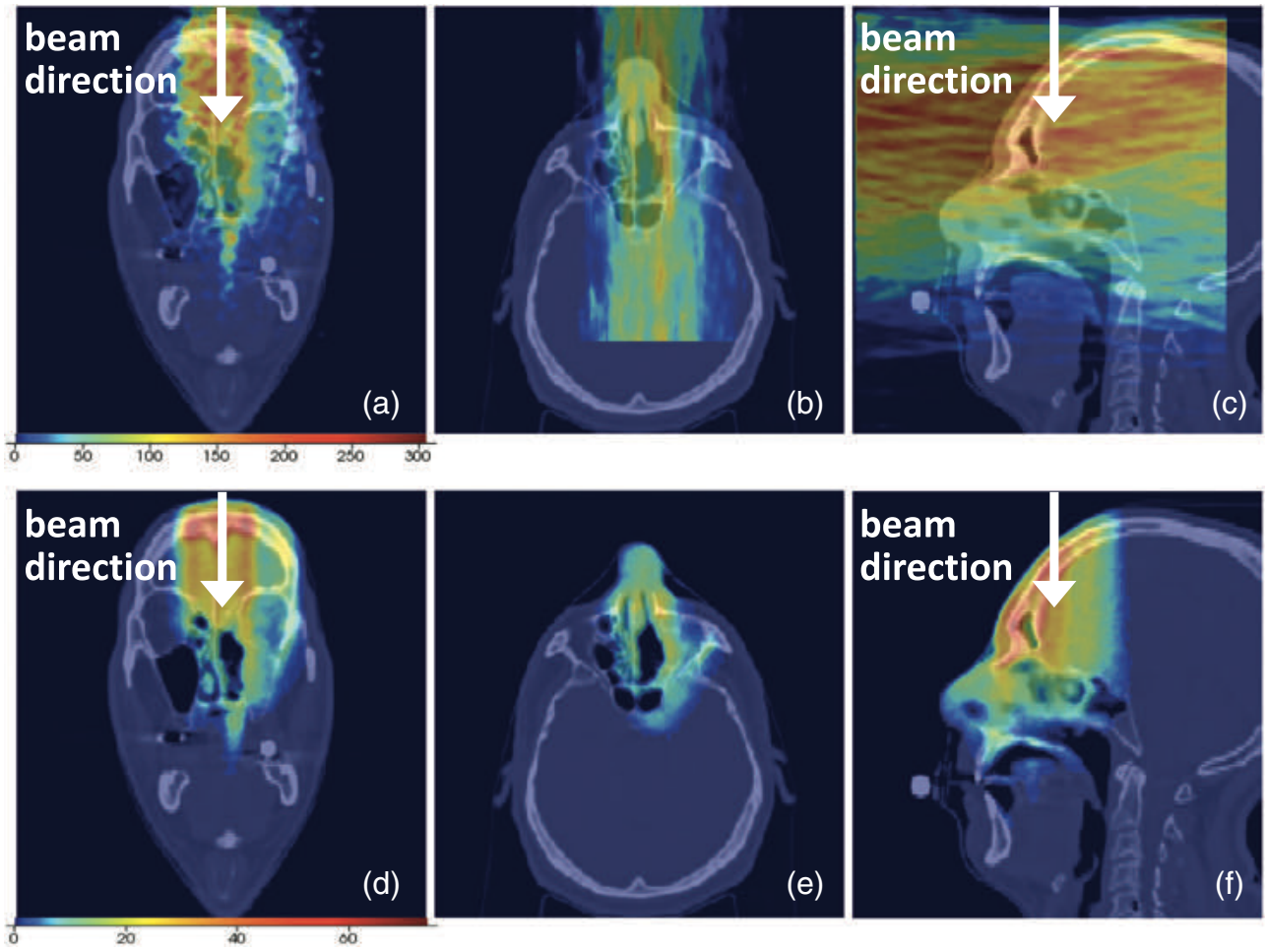

FIGURE 5 Top figures: PET signal in the coronal (a), axial (b), and sagittal (c) planes. Deformation in the axial and coronal plane can be seen. Bottom figures: IPM signal in the coronal (d), axial (e), and sagittal (f) planes. The beam direction is indicated with an arrow in the coronal and sagittal plane, whereas it is orthogonal to the displayed plane in the axial view

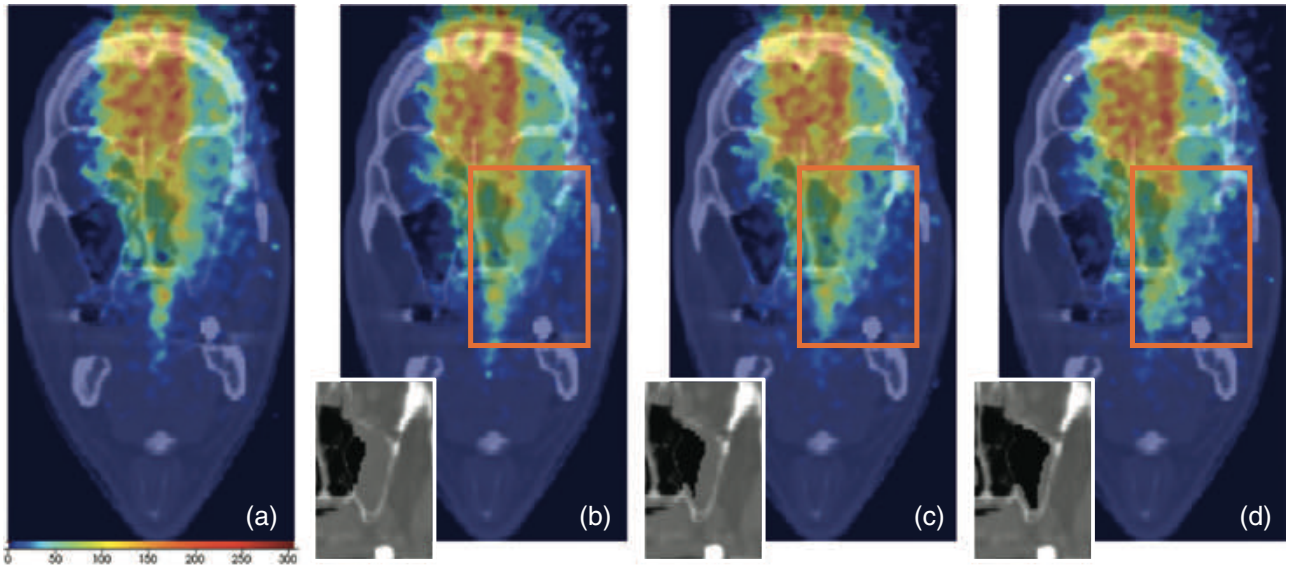

FIGURE 6 Evolution of the PET signal with the gradual emptying of the sinonasal cavity, overlaid on the planning CT, for cavity tissue volume reductions of 0 mL (a), 3.8 mL (b), 7.3 mL (c), and 13.1 mL (d). The small insets on the left bottom in (b), (c), and (d) represent the real underlying situation in the artificially modified CT

uncorrelated with the region where morphological changes occurred.

In Figure 7 we show three different slices of the *p*-value maps overlaid on the planning CT, resulting from the statistical analysis of the control sample $PET_{obs,0}$ (Figure 7(a), (e), and (i)), and three intermediate PET images, $PET_{obs,j}$ for j =2, 4, and 6, corresponding to emptied volumes of 3.8 mL (Figure 7(b), (f), and (j)), 7.3 mL (Figure 7(c), (g), (k)), and 13.1 mL (Figure 7(d), (h), and (l)), respectively. First of all, we observe that the region of morphological changes was well identified through the blue zones (lack of activity), and corresponded well to the regions that were actually emptied in the CT scan. The latter can be seen from the small insets in the left bottom of the slices, showing the corresponding region in the artificially modified CT scan. Second, red zones are visible, located behind the cavity region with respect to the beam entrance, representing voxels with an excess of activity with respect to expectations. This characteristic pattern, first blue and then red when following the beam direction, is a strong indicator for a significant morphological changes resulting in a beam overshoot, here caused by a replacement of tissue by air, resulting in a lower than expected PET signal in the tissue emptying region, and in a large than expected activity behind it due to the increased beam range. Third, we note in Figure 7 that some noise is visible across the slice, that is not associated to anatomical changes. We come back to this in Section 3.3.

Figure 8(a) and (b) show the *p*-value map resulting from the statistical analysis performed on the PET_{CT1} and PET_{CT2} images corresponding to the first and second control CT, respectively. In Figure 8(a), corresponding to minimal changes, we see that the *p*-value map highlights some small red zones behind the emptied region, while no blue voxels can be seen. In Figure 8(b), corresponding to the situation requiring adaptive intervention (see Section 2.1), the significant regions are remarkably well visible through the blue-red pattern. Thus, the method clearly identifies the regions with morphological changes also for the PET signal simulated on the real CTs.

3.3 | Identification of regions with morphological changes: Quantitative assessment

Figure 9 shows the relationship between the volume that was classified as significantly different from the null reference distribution, V_{obs}^{PET}, and the actually changed volume, V_{mod}^{CT}. In this figure, the bullet points represent the average values of the observed volume differences obtained over the 10 simulations made for tissue volume change (see Section 2.4.2), and the error bars are the standard deviation. A linear fit is carried out, obtaining the following parameters: $a = 0.89$ and $b = 2.15$, with an $R^2 = 0.99$. The fitted line is given in the figure, where the line $y = x$ is also given.

TABLE 3 Possible error sources in quantification of the tissue emptying volume, for the case where $V_{mod}^{CT} = 13.1$ mL. The default V_{obs}^{PET} for this case, $V_{default}$, was 12.6 mL

Error source	Size of identified volume (mL)	ΔV(mL)=$V_{obs} - V_{default}$
Cross sections in MC +10%	8.7	−3.9
Cross sections in MC −10%	19.4	7.1
Fraction $f = 1/60$	12.7	0.1
Positioning	13.9	1.4

From Figure 9 we see that even small morphological changes in the simulated images can be detected through the VBM-like analysis we implemented. At the same time we note that most of the points are located above the line $y = x$. This is an expected effect of our statistical analysis, due to the multiple testing problem. A large number of statistical tests are actually carried out simultaneously, as large as the voxel extent of the CTV-low (approximately $N_v = 35\,000$ voxels in our case). Once we fix the significance threshold for each test to α, for the lower-than-expected part of the signals we will obtain up to $\alpha/2 \times N_v$ voxels that are by chance erroneously classified as positive (False Positives or Type I errors). This effect, causing also the noise in Figure 7, can be mitigated by implementing suitable strategies to correct for multiple comparisons which control the proportion of Type I errors.

In addition to the statistical errors related to MC simulations and to the VBM-like analysis, the impact on the maps and the estimated volume change of the following error sources mentioned in Section 2.7 was evaluated. All associated tissue volumes changes, which were estimated for the point with the maximum volume nominal volume change ($V_{mod}^{CT} = 13.1$ mL), are reported in Table 3. The corresponding maps are given in Figure S1 of the Supplementary Material. Although it is clear that the size of the identified volume is sensitive to errors, the identification of the region was seen to be robust. Moreover, most errors from Table 3 can be mitigated by appropriate MC validation and alignment procedures.

3.4 | Influence of PET image artifacts

Figure 10 shows three slices of the *p*-value map overlaid onto the planning CT, based on the statistical analysis of the IPM_{obs} images. The cavity emptying region is very clearly identified and the image hardly contains noise. This is because the images are free of artifacts, and because statistical fluctuations are much smaller than in the PET images displayed in Figure 7.

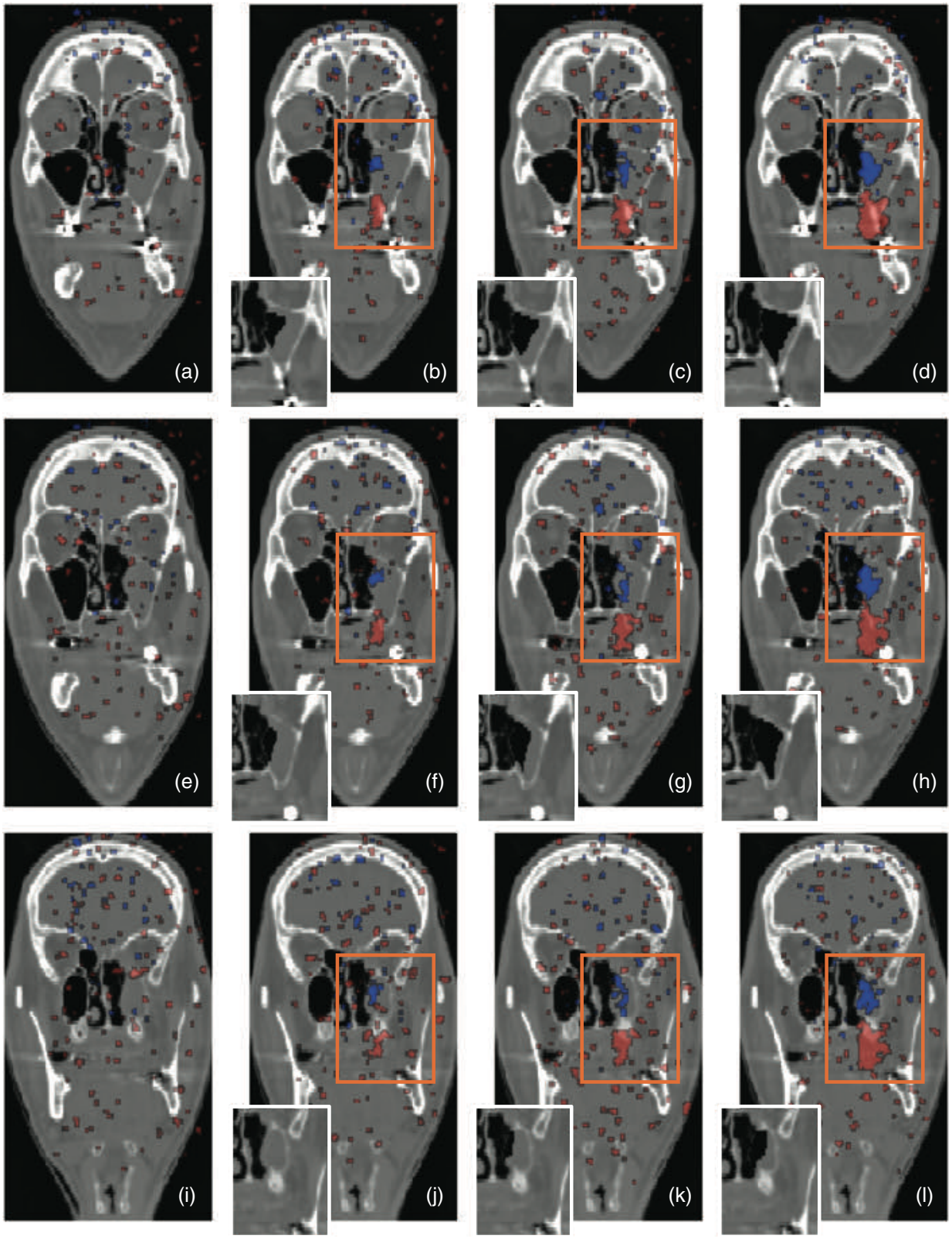

FIGURE 7 *p*-value maps for emptied volumes of 0.0 mL (a, e, i), 3.8 mL (b, f, j), 7.3 mL (c, g, k), and 13.1 mL (d, h, l). The upper, medium, and lower four plots correspond to a different CT slice. The small insets in the slices show the region that corresponds to the observed *p*-value map

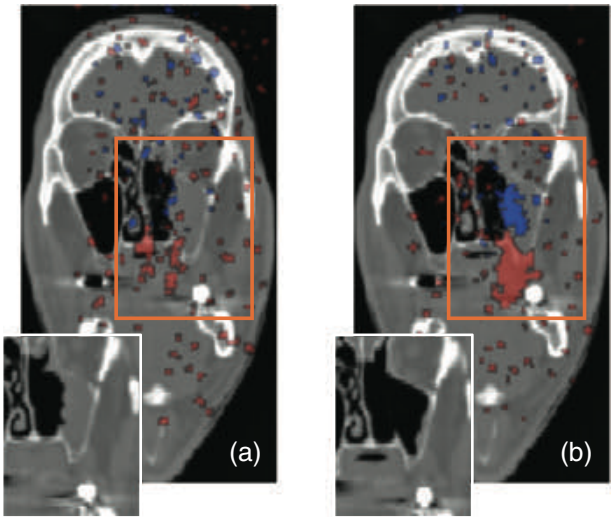

FIGURE 8 *p*-value maps overlaid on the planning CT, resulting from the statistical analysis of the PET images corresponding to the first control CT (a) and the second control CT (b). The small insets in the slices show actual situation in the control CTs that gave rise to the maps

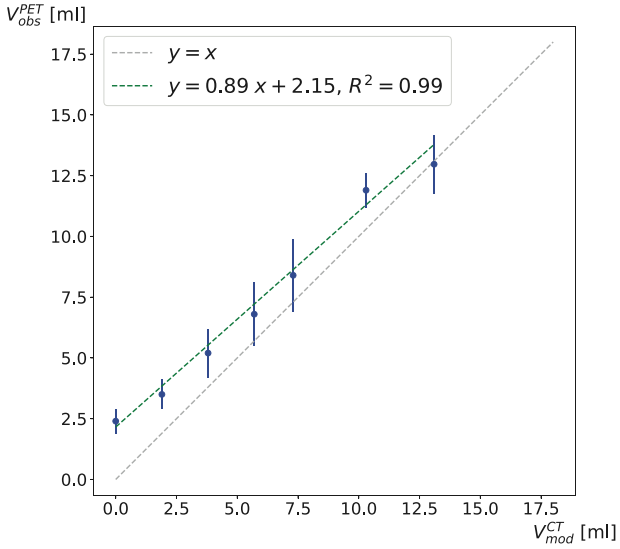

FIGURE 9 Observed volume reduction in PET images, V^{PET}_{obs}, versus actually introduced volume reduction in the intermediate CT scans, V^{CT}_{mod}

In Figure 11 we demonstrate the observed volume change V^{IPM}_{obs} versus the actually introduced volume change, V^{CT}_{mod}. As in Figure 9, the bullet points represent the average values of the observed volume differences obtained over the 10 simulations made for tissue volume change (see Section 2.4.2), and the error bars are the standard deviation. Again a a linear fit with $y = ax + b$ was carried out, yielding a= 0.89 and b = 2.15, with an R^2 = 0.99. The line $y = x$ is also given. When comparing Figure 11 with Figure 9, the points in the former are seen to be distributed much more regularly than those in the latter figure. Thus, the partial angular coverage of the detector and the resulting small geometrical efficiency result in much more fluctuations in the *p*-value maps. Moreover a deviation can be seen from the line $y = x$: the detected volume is biased toward somewhat too high values. This is caused by False Positives, discussed before in Section 3.3, but since the IPM distributions are less fluctuated than the PET distributions, the effect is small.

4 | DISCUSSION

Using MC simulations of a patient treatment with proton therapy, we demonstrated that it was possible to produce 3-D probability maps of morphological changes based on a VBM-inspired method. There are various issues that deserve discussion.

4.1 | Application of VBM method to in-beam PET

We saw from Figure 7 that the quality of the *p*-value maps was affected by the limited statistics and the image artifacts in the PET images (Figure 5) when compared to the ideal situation in Figure 10. To investigate the fluctuations, we tested the effect of different sizes of median filters, but we observed that one voxel in all directions was sufficient to mitigate large fluctuations. To decrease fluctuations from poor statistics, caused by the detector geometry and the limited data acquisition time, it would help to combine the different irradiation fields, so more statistics could be collected.[30] This is challenging, since the PET signal acquired during the second and third field delivery would be a mixed contribution of the fields, but strategies are being developed to overcome this problem.[50] The time passing between the delivery of one field and the next (typically of the order of minutes) and combination of signals would have to be carefully taken into account. This is part of future work. Concerning the image artifacts, these cannot be mitigated and are an intrinsic limitation of dual-head in-beam PET systems. However, they did not hamper the correct identification of the zones with morphological changes.

4.2 | Interpretation of color maps

Concerning the interpretation of the color maps, we remind that in our case, the tissue volume change was caused by an emptying of the cavity as a response to the radiation, so it concerned a replacement of water-like tissue by air. Water and air having a very different activity signal, and given the beam direction, the method resulted in the blue-red pattern indicating in this case an air cavity followed by a beam overshoot. This was

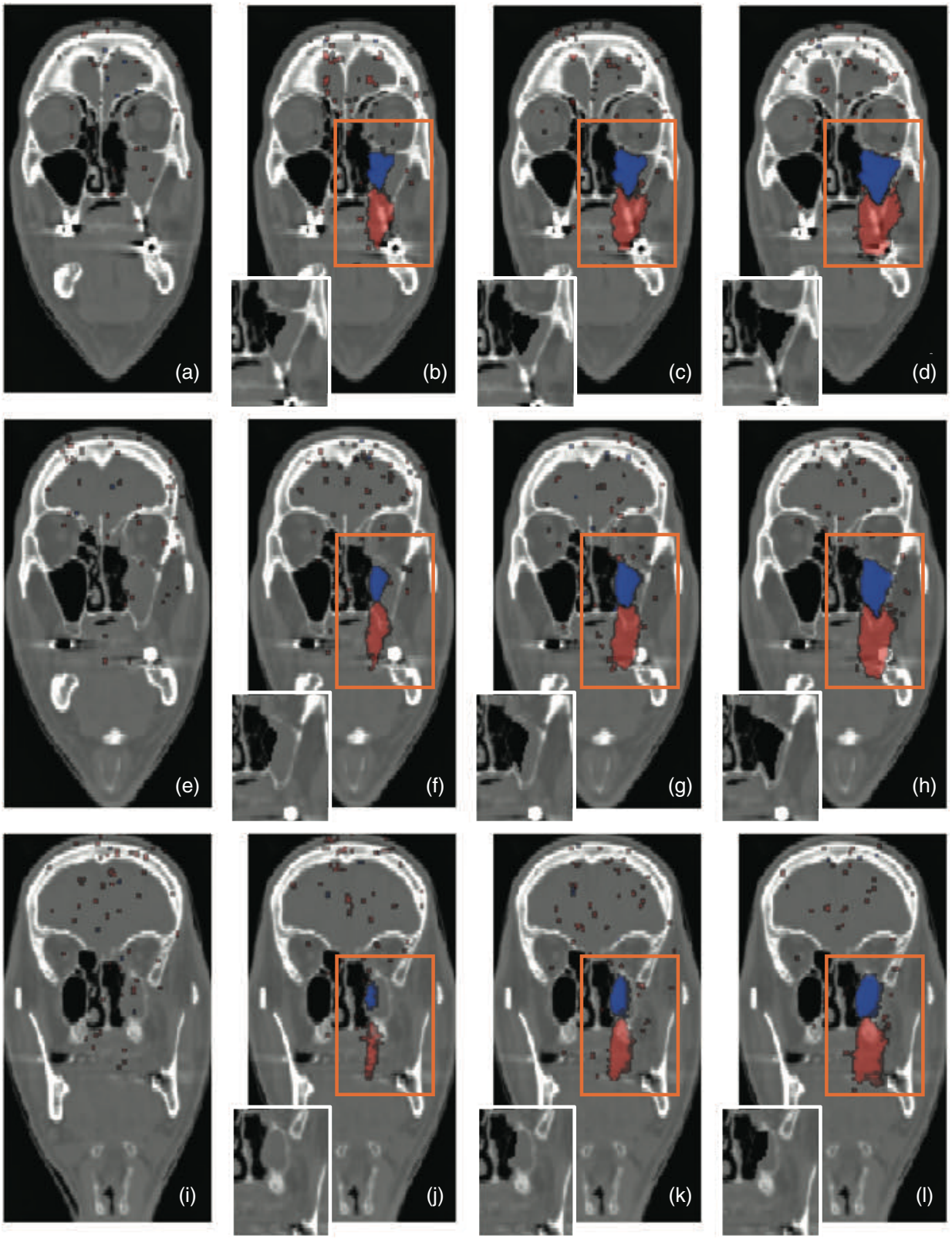

FIGURE 10 *p*-value maps resulting from the IPM analysis, for cavity tissue volume reductions of 0.0 mL (a, e, i), 3.8 mL (b, f, j), 7.3 mL (c, g, k), and 13.1 mL (d, h, l). The upper, medium, and lower four plots correspond to a different CT slice. The small insets in the slices show the region in the modified CT that corresponds to the observed *p*-value map

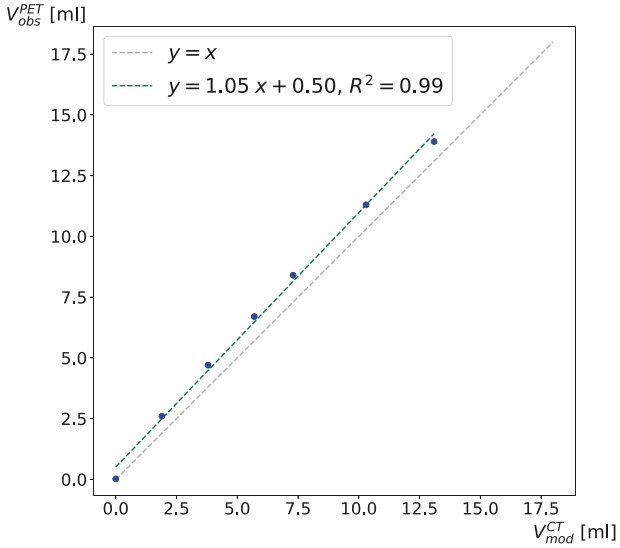

FIGURE 11 Observed volume reduction in IPM maps versus actually introduced volume change in the intermediate CT scans

just one example, we believe that the method would also be sensitive to other kinds of anatomical changes, provided that the anatomical change results in a statistically significantly different PET signal. For instance, let us consider the case of tumor shrinkage, which is a common phenomenon in clinical practice. If the tumor tissue shrinks and is replaced by air, the same blue-red pattern can be expected and the quantification of the volume change would likely be possible. In case the tumor tissue is replaced by water-like material, some differences may be expected too, since (tumor) cell material contains more carbon than water and thus the PET signal would change.[51] However, maybe there would not be a blue-red pattern, but only a local blue, or only a local red pattern. Future research and patient data would be needed to investigate the expected sensitivity of the method to reveal different types of anatomical changes. For the moment, the PET images can serve as a warning ("red flag"), that should trigger the attention of the radiation oncologist if changes are found before a control CT is planned.

Finally, we would like to stress that our method should be seen as complementary to already established imaging techniques like CBCT imaging. In fact, CBCT can give valuable information about the positioning of the patient and the tumor anatomy *before treatment*, but it does not give information about associated dose modifications. For instance, there can be anatomical changes that hardly alter the delivered dose, because the plan is robust enough. In contrast, in-beam PET can give valuable information about the dose delivery *during treatment*. Although PET activity and dose are only indirectly correlated, the presence of significant changes in the in-beam PET signal is undoubtedly a strong indication that the patient dose has changed. We believe that the voxel-by-voxel method to identify the location of changes as presented in this work can bring us a step closer to a real dose reconstruction and support further research in the interpretability of PET monitoring images.

4.3 | Free parameters

Another point of discussion is the presence of free parameters in the *p*-value analysis. One free parameter is the significance level α for defining significant *p*-values (Section 2.3), which was fixed to 0.05, a commonly used value. We repeated the analysis with $\alpha = 0.025$ and saw that the volume and especially the noise decreased somewhat (see Supplementary Material), This is as expected; only the most significant voxels remained.

Another free parameter is the number of replicates N used to build the null reference distributions (Section 2.3). We generated $N = 120$ replicates to build the null distributions of the IPM and PET image voxels, which turned out to be sufficient. We expect a larger N to be somewhat more robust to statistical fluctuations. However, this brings us to a drawback of our method: the large amount of computing resources needed. It took us about 18 h to simulate one replicate, however this aspect can be greatly improved by further parallelization of the simulations or by cloud-based computing resources. A possible implementation would be a web interface that allows medical personnel to upload all the relevant pieces of information, which would then be processed by a cloud computing service, giving back the results for subsequent analysis and interpretation. Another possibility is the usage of fast MC simulation codes, which are being more and more frequently used in clinical practice.[52]

4.4 | Other considerations

Regarding the applicability to real data, the viability and effectiveness of the method in real treatments is still to be tested. However, a few promising results were already obtained in this study, that evidenced some clear advantages of our method with respect to previously used methods:

- The number of free parameters is small. Apart from α and N, there are no free parameters in the analysis method. It relies purely on statistical considerations, with a priori fixed significance thresholds.
- The method gives voxel-by-voxel information. This approach allows not only to detect problems, but also to identify the regions where morphological changes have likely occurred (Figures 7, 8, and 10) and to quantify their volume (Figures 9 and 11), something that is neither immediate nor feasible with other meth-

ods. The voxel-by-voxel information would thus bring us a step closer to eventually translate the results into dose values that are of real clinical interest, using for instance machine learning techniques and artificially created series of customized CT scans including best guesses of morphological changes.

- The method gives specific indications of what actually occurred regarding the morphological change in terms of characteristic blue-red patterns, as we saw from (Figures 7, 8, and 10). Looking from the beam direction, for an air cavity, we expect a pattern of first a blue zone followed by a red zone. And vice versa, in case the beam passes first a zone of increased density with respect to the unmodified CT, one would observe first a red zone, followed by a blue zone. The presence of such patterns are strong indications of morphological changes, giving qualitative and quantitative indications about the changes ongoing in the patient along the beam path. In case of a gradually occurring morphological change, this signal would become more notable over subsequent treatment fractions. The information can be overlaid onto the CT to investigate the cause. Not only such information can be presented easily to medical personnel, but it could also be used to automate adaptive interventions, for instance, by raising red flags in case the pattern reaches a certain size.

Finally, the method was tested for in-beam PET simulations, but can be applied also to other imaging modalities, like in-room or after-beam PET treatment monitoring, or eventually even charged particle or prompt gamma detection.

5 | CONCLUSIONS

We presented a method to detect interfractional morphological changes in a patient during a proton therapy treatment. To the best of our knowledge, this is the first time that a VBM resembling method, used in neuroimaging, has been applied to in-beam PET proton therapy monitoring images. Using a set of simulated PET images based on CT scans with and without morphological variations, we tested the method for a patient with a head-and-neck tumor where morphological changes occurred, that was treated at CNAO. Although the viability and effectiveness of the method in real treatments are still to be confirmed, a few important results were already obtained. First of all, we confirmed that a voxel-wise statistical analysis resembling VBM is appropriate to locate morphological changes through in-beam PET images. This was possible despite the strong image artifacts and limited statistics typically present in in-beam PET images. Moreover, the analysis results could be clearly visualized through characteristic colored patterns that can be effortless overlaid onto the CT scans. We believe that the presence of such patterns makes the method particularly appropriate for raising red flags, that is, warning signs to trigger the attention of the radiation oncologist. Whether the method eventually allows to automate adaptive planning can only be confirmed in future research with large amounts of patient data. Finally, we demonstrated a satisfactory sensitivity to the detection of cavity volume changes and an approximately linear proportionality with the cavity tissue volume reduction, even down to small changes, which is encouraging in view of applying the method to real data.

ACKNOWLEDGMENTS

This research was funded by the Italian Ministry of Education (PRIN MIUR 2010P98A75 INSIDE, 2013–2016), the Italian Institute of Nuclear Physics (RDH and INFN-RT2 PETRA 172800 projects, since 2016), the Historical Museum of Physics and the Enrico Fermi Study and Research Center and the Tuscany Government (POR FSE 2014–2020, INFN-RT2 PETRA 172800 Project, 2018-2019), and the CNAO Foundation (INSIDE2, since 2017).

CONFLICT OF INTEREST

The authors have no conflicts to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

REFERENCES

1. Loeffler JS, Durante M. Charged particle therapy–optimization, challenges and future directions. *Nat Rev Clin Oncol*. 2013;10:411–424.
2. Paganetti H. *Proton Therapy Physics*. CRC Press; 2012.
3. Mackay R. Image guidance for proton therapy. *Clin Oncol*. 2018;30. https://doi.org/10.1016/j.clon.2018.02.004.
4. Albertini F, Matter M, Nenoff L, Zhang Y, Lomax A. Online daily adaptive proton therapy. *Br J Radiol*. 2020;93:20190594.
5. Morgan H, Sher D. Adaptive radiotherapy for head and neck cancer. *Cancers of the Head & Neck*. 2020;5. https://doi.org/10.1186/s41199-019-0046-z.
6. Sonke J, Aznar M, Rasch C. Adaptive radiotherapy for anatomical changes. *Semin Radiat Oncol*. 2019;29:245–257.
7. Kraan AC, van de Water S, Teguh DN, et al. Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors. *Int J Radiat Oncol Biol Phys*. 2013;8:888–896.
8. Placidi L, Bolsi A, Lomax AJ, et al. Effect of anatomic changes on pencil beam scanned proton dose distributions for cranial and extracranial tumors. *Int J Radiat Oncol Biol Phys*. 2017;97:616–623.
9. Cubillos-Mesías M, Troost EG, Lohaus F, et al. Including anatomical variations in robust optimization for head and neck proton therapy can reduce the need of adaptation. *Radiother Oncol*. 2019;131:127–134.
10. Kurz C, Nijhuis R, Reiner M, et al. Feasibility of automated proton therapy plan adaptation for head and neck tumors using cone beam CT images. *Radiat Oncol*. 2016;11. https://doi.org/10.1186/s13014-016-0641-7.

11. Zumsteg Z, DeMarco J, Lee S, et al. Image guidance during head-and-neck cancer radiation therapy: analysis of alignment trends with in-room cone-beam computed tomography scans. *Int J Radiat Oncol Biol Phys*. 2011;83:712–719. https://doi.org/10.1016/j.ijrobp.2011.08.001.
12. Parodi K, Polf J. In vivo range verification in particle therapy. *Med Phys*. 2018;45:e1036–e1050. https://doi.org/10.1002/mp.12960.
13. Knopf A, Lomax A. In vivo proton range verification: a review. *Phys Med Biol*. 2013;58:R131–160.
14. Kraan AC. Range verification methods in particle therapy: underlying physics and Monte Carlo modeling. *Front Oncol*. 2015;5:1–27.150.
15. Enghardt W, Crespo P, Fiedler F, et al. Charged hadron tumour therapy monitoring by means of PET. *Nucl Instrum Methods Phys Res, Sect A*. 2004;525:284–288. [Proceedings of the International Conference on Imaging Techniques in Subatomic Physics, Astrophysics, Medicine, Biology and Industry. ISSN: 0168-9002. https://doi.org/10.1016/j.nima.2004.03.128. URL: http://www.sciencedirect.com/science/article/pii/S0168900204004218.]
16. Sportelli G, Belcari N, Camarlinghi N, et al. First full-beam PET acquisitions in proton therapy with a modular dual-head dedicated system. *Phys Med Biol*. 2013;59:43–60. https://doi.org/10.1088/0031-9155/59/1/43.
17. Nishio T, Miyatake A, Ogino T, Nakagawa K, Saijo N, Esumi H. The development and clinical use of a beam ON-LINE PET system mounted on a rotating gantry port in proton therapy. *Int J Radiat Oncol Biol Phys*. 2010;76:277–286.
18. Dendooven P, Buitenhuis T, Diblen F, et al. Short-lived positron emitters in beam-on PET imaging during proton therapy. *Phys Med Biol*. 2015;60:8923–8947. https://doi.org/10.1088/0031-9155/60/23/8923.
19. Rossi S. The status of CNAO. *Eur Phys J Plus*. 2011;126. https://doi.org/10.1140/epjp/i2011-11078-8.
20. Bisogni GM, Attili A, Battistoni G, et al. INSIDE in-beam positron emission tomography system for particle range monitoring in hadrontherapy. *J Med Imaging*. 2016;4:011005. https://doi.org/10.1117/1.JMI.4.1.011005.
21. Attanasi F, Belcari N, Camarda M, et al. Experimental validation of the filtering approach for dose monitoring in proton therapy at low energy. *Physica Med*. 2008;24:102–106. [Highlights of the "X EFOMP Congress - First European Conference on Medical Physics". ISSN: 1120-1797. https://doi.org/10.1016/j.ejmp.2008.03.001. URL: https://www.sciencedirect.com/science/article/pii/S1120179708000446.]
22. Parodi K, Bortfeld T. A filtering approach based on Gaussian-powerlaw convolutions for local PET verification of proton radiotherapy. *Phys Med Biol*. 2006;51:1991–2009. https://doi.org/10.1088/0031-9155/51/8/003.
23. Remmele S, Hesser J, Paganetti H, Bortfeld T. A deconvolution approach for PET-based dose reconstruction in proton therapy. *Phys Med Biol*. 2011;56:7601–7619. https://doi.org/10.1088/0031-9155/56/23/017.
24. Fourkal E, Fan J, Veltchev I. Absolute dose reconstruction in proton therapy using PET imaging modality: feasibility study. *Phys Med Biol*. 2009;54:N217–N228. https://doi.org/10.1088/0031-9155/54/11/N02.
25. Min CH, Zhu X, Winey BA, et al. Clinical application of in-room positron emission tomography for in vivo treatment monitoring in proton radiation therapy. *Int J Radiat Oncol Biol Phys*. 2013;86:183–189.
26. Knopf A, Parodi K, Paganetti H, Cascio E, Bonab A, Bortfeld T. Quantitative assessment of the physical potential of proton beam range verification with PET/CT. *Phys Med Biol*. 2008;53:4137–4151. https://doi.org/10.1088/0031-9155/53/15/009.
27. Frey K, Unholtz D, Bauer J, et al. Automation and uncertainty analysis of a method forin-vivorange verification in particle therapy. *Phys Med Biol*. 2014;59:5903–5919. https://doi.org/10.1088/0031-9155/59/19/5903.
28. Fiorina E, Ferrero V, Pennazio F, et al. Monte Carlo simulation tool for online treatment monitoring in hadrontherapy with in-beam PET: a patient study. *Physica Med*. 2018;51:71–80.
29. Ferrero V, Fiorina E, Morrocchi M, et al. Online proton therapy monitoring: clinical test of a Silicon-photodetector-based in-beam PET. *Sci Rep*. 2018;8. https://doi.org/10.1038/s41598-018-22325-6.
30. Fiorina E, Ferrero V, Baroni G, et al. Detection of interfractional morphological changes in proton therapy: a simulation and in vivo study with the inside in-beam PET. *Front Phys*. 2021;8:578388. https://doi.org/10.3389/fphy.2020.578388.
31. ClinicalTrials.gov, *INSIDE clinical trial*. https://www.clinicaltrials.gov/ct2/show/study/NCT03662373.
32. Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*. 1994;2:189–210. https://doi.org/10.1002/hbm.460020402.
33. Calderoni S, Retico A, Biagi L, et al. Female children with autism spectrum disorder: an insight from mass-univariate and pattern classification analyses. *NeuroImage*. 2012;59:1013–1022.
34. Curzio O, Calderoni S, Maestro S, et al. Lower gray matter volumes of frontal lobes and insula in adolescents with anorexia nervosa restricting type: findings from a brain morphometry study. *Eur Psychiatry*. 2020;63:e27. https://doi.org/10.1192/j.eurpsy.2020.19.
35. Palumbo L, Bosco P, Fantacci M, et al. Evaluation of the intra- and inter-method agreement of brain MRI segmentation software packages: a comparison between SPM12 and FreeSurfer v6.0. *Physica Med*. 2019;64:261–272.
36. Ashburner J, Friston KJ. Voxel-based morphometry—the methods. *NeuroImage*. 2000;11:805–821.
37. Davison A, Hinkley D. Bootstrap methods and their application. *J Am Statist Assoc*. 1997;94. https://doi.org/10.2307/1271471.
38. Phipson B, Smyth GK. Permutation P-values should never be zero: calculating exact p-values when permutations are randomly drawn. *Stat Appl Genet Mol Biol*. 2010;9. https://doi.org/10.2202/1544-6115.1585.
39. North B, Curtis D, Sham P. A note on the calculation of empirical p values from Monte Carlo procedures. *Am J Hum Genet*. 2002;71:439–41. https://doi.org/10.1086/341527.
40. Pennazio F, Battistoni G, Bisogni MG, et al. Carbon ions beam therapy monitoring with the INSIDE in-beam PET. *Phys Med Biol*. 2018;63. https://doi.org/10.1088/1361-6560/aacab8.
41. Battistoni G, et al.. The FLUKA code: an accurate simulation tool for particle therapy. *Front Oncol*. 2016;6:116.
42. Mirandola A, Molinelli S, Vilches-Freixas G, et al. Dosimetric commissioning and quality assurance of scanned ion beams at the italian national center for oncological hadrontherapy. *Med Phys*. 2015;42:5287. https://doi.org/10.1118/1.4928397.
43. Magro G, Molinelli S, Mairani A, et al. Dosimetric accuracy of a treatment planning system for actively scanned proton beams and small target volumes: Monte Carlo and experimental validation. *Phys Med Biol*. 2015;60:6865–6880. https://doi.org/10.1088/0031-9155/60/17/6865.
44. Molinelli S, Mairani A, Mirandola A, et al. Dosimetric accuracy assessment of a treatment plan verification system for scanned proton beam radiotherapy: one-year experimental results and Monte Carlo analysis of the involved uncertainties. *Phys Med Biol*. 2013;58:3837–3847.
45. Giordanengo S, Garella M, Marchetto F, et al. The CNAO dose delivery system for modulated scanning ion beam radiotherapy. *Med Phys*. 2015;42:263. https://doi.org/10.1118/1.4903276.
46. Schneider W, Bortfeld T, Schlegel W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Phys Med Biol*. 2000;45:459–78. https://doi.org/10.1088/0031-9155/45/2/314.
47. Camarlinghi N, Sportelli G, Battistoni G, et al. An in-beam PET system for monitoring ion-beam therapy: test on phantoms using

clinical 62 MeV protons. *J Instrum*. 2014;9:C04005–C04005. https://doi.org/10.1088/1748-0221/9/04/c04005.
48. Knopf A, Parodi K, Paganetti H, et al. Accuracy of proton beam range verification using post-treatment positron emission tomography/computed tomography as function of treatment site. *Int J Radiat Oncol Biol Phys*. 2011;79:297–304.
49. Helmbrecht S, Kuess P, Birkfellner W, et al. Systematic analysis on the achievable accuracy of PT-PET through automated evaluation techniques. *Z Med Phys*. 2015;25:146–155.
50. Ferrero V, Bisogni MG, Camarlinghi N, et al. Double-field hadrontherapy treatment monitoring with the INSIDE In-Beam PET scanner: proof of concept. *IIEEE Trans Radiat Plasma Med Sci*. 2018;2:588–593. https://doi.org/10.1109/TRPMS.2018.2870478.
51. Cho J, Ibbott G, Gillin M, et al. Determination of elemental tissue composition following proton treatment using positron emission tomography. *Phys Med Biol*. 2013;58:3815.
52. Muraro S, Battistoni G, Kraan AC. Challenges in Monte Carlo simulations as clinical and research tool in particle therapy: a review. *Front Phys*. 2020;8:391. https://doi.org/10.3389/fphy.2020.567800
53. Vlachoudis V FLAIR: A Powerful But User Friendly Graphical Interface For FLUKA, Proc. Int. Conf. on Mathematics, Computational Methods & Reactor Physics (M&C 2009), Saratoga Springs, New York, 2009 URL: http://www.fluka.org/flair/Flair_MC2009.pdf

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kraan AC, Berti A, Retico A, et al. Localization of anatomical changes in patients during proton therapy with in-beam PET monitoring: A voxel-based morphometry approach exploiting Monte Carlo simulations. *Med. Phys.* 2022;49:23–40. https://doi.org/10.1002/mp.15336