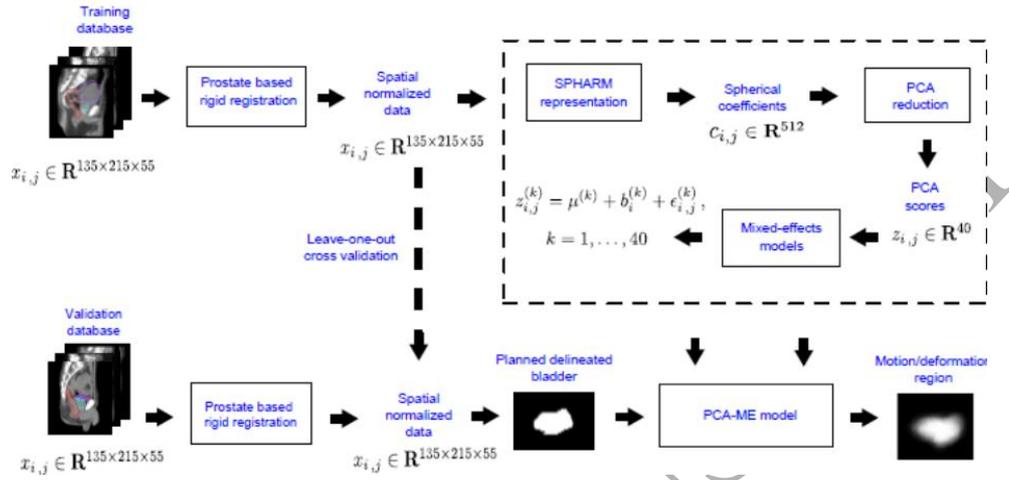


Highlights

- We propose a population model to estimate the probability of the bladder presence on a given region during treatment using only the planning CT scan as input information.
- We train a motion/deformation model, based on longitudinal data, to predict bladder motion and deformation between fractions.
- We propose a longitudinal analysis using mixed-effect models to separate intra- and inter-patient variability in order to control confounding.
- We reduce, in a factor of 10, the number of variables required to represent bladder surface using spherical harmonics (SPHARM).



Population model of bladder motion and deformation based on dominant eigenmodes and mixed-effects models in prostate cancer radiotherapy [☆]

Richard RIOS^{a,b,c,*}, Renaud De CREVOISIER^{a,b,d}, Juan D. OSPINA^{a,b}, Frederic COMMANDEUR^{a,b}, Caroline LAFOND^d, Antoine SIMON^{a,b}, Pascal HAIGRON^{a,b}, Jairo ESPINOSA^c, Oscar ACOSTA^{a,b}

^aINSERM, U1099, F-35000 Rennes, France

^bUniversité de Rennes 1, LTSI, F-35000 Rennes, France

^c Universidad Nacional de Colombia, Facultad de Minas, GAUNAL, Medellín, Colombia

^dCRLCC Eugène Marquis, Département de Radiothérapie, F-35000 Rennes, France

Abstract

In radiotherapy for prostate cancer irradiation of neighboring organs at risk may lead to undesirable side-effects. Given this setting, the bladder presents the largest inter-fraction shape variations hampering the computation of the actual delivered dose vs. planned dose. This paper proposes a population model, based on longitudinal data, able to estimate the probability of bladder presence during treatment, using only the planning computed tomography (CT) scan as input information. As in previously-proposed principal component analysis (PCA) population-based models, we have used the data to obtain the dominant eigenmodes that describe bladder geometric variations between fractions. However, we have used a longitudinal analysis along each mode in order to properly characterize patient's variance from the total population variance. We have proposed a mixed-effects (ME) model in order to separate intra- and inter-patient variability, in an effort to control confounding cohort effects. Other than using PCA, bladder shapes are represented by using spherical harmonics (SPHARM) that additionally enables data compression without information lost. Based on training data from repeated CT scans, the ME model was thus implemented following dimensionality reduction by means of SPHARM and PCA. We have evaluated the model in

[☆]This document is a collaborative effort between the LTSI-INSERM U1099, Université de Rennes 1 and GAUNAL research group from the Universidad Nacional de Colombia.

*Corresponding author: Richard RIOS (rriospa@unal.edu.co)

a leave-one-out cross validation framework on the training data but also using independent data. Probability maps (PMs) were thus generated with several draws from the learnt model as predicted regions where the bladder will likely move and deform. These PMs were compared with the actual regions using metrics based on mutual information distance and misestimated voxels. The prediction was also compared with two previous population PCA-based models. The proposed model was able to reduce the uncertainties in the estimation of the probable region of bladder motion and deformation. This model can thus be used for tailoring radiotherapy treatments.

Keywords: principal component analysis, motion/deformation models, mixed-effects models, prostate cancer radiotherapy.

Prostate cancer is one of the most common cancer types worldwide. One of the standard treatment methods is external radiotherapy, which involves delivering ionizing radiation to a clinical target, namely the prostate and seminal vesicles. Due to the uncertain location of organs during treatment, consisting around 40 radiation fractions delivering a total dose ranging from 70 to 80Gy, safety margins are defined around the tumor target upon treatment planning. This also leads to portions of healthy organs neighboring the prostate or organs at risk – the bladder and rectum – to be included in the target volume, potentially resulting in adverse events affecting patients' urinary (hematuria and cystitis, among others) or rectal (rectal bleeding, fecal incontinence, etc.) functions. Several studies have shown that increasing dose delivery to the prostate leads to improved local cancer control, up to approximately 80Gy (Zietman et al., 2010). However, such dose increases are limited by their associated risks of treatment-related toxicity involving the organs at risk (Fiorino et al., 2009; Fonteyne et al., 2008).

The bladder is particular for presenting the largest inter-fraction shape variations during treatment, caused by continuous volume changes (Roeske et al., 1995; Turner et al., 1997). These shape variations introduce geometric uncertainties that render assessment of the actual dose delivered to the bladder during treatment difficult, thereby leading to dose uncertainties that limit the possibility of modeling dose-volume response for late genitourinary (GU) toxicity (Fiorino et al., 2009; Thor et al., 2013b; Nassef et al., 2016; Palorini et al., 2016a; Landoni et al., 2016). The Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) project has stated that a similar dose-response to that of late gastrointestinal (GI) toxicity is far from being established. The dosimetric variables obtained from the planning CT prove to be very poor surrogates for the real-delivered dose

(Viswanathan et al., 2010; Rosewall et al., 2010; Palorini et al., 2016b). As a result, it appears crucial to quantify uncertainties produced by inter-fraction bladder variations in order to determine dosimetric factors that affect late GU complications.

Several approaches have been developed in order to quantify and characterize the geometrical uncertainties produced by organ motion and deformation between fractions, including: serial imaging measurement of the organ during treatment course (Roeske et al., 1995; van Herk et al., 1995; Ten Haken et al., 1991), fiducial markers (Crook et al., 1995; Balter et al., 1995), margins of organs at risk (OAR) (Thor et al., 2010; Mageras et al., 1999; Stroom et al., 1999), rigid-body motion (Thor et al., 2013a; Killoran et al., 1997; Craig et al., 2001), parametrization of the organ structure (Hoogeman et al., 2002; Mageras et al., 1996; Pavel-Mititean et al., 2004), as well as biomechanical (Yan et al., 1999) and statistical models (Fontenla et al., 2001a,b; Söhn et al., 2012; Budiarto et al., 2011; Thörnqvist et al., 2013; Bondar et al., 2014; Rios et al., 2016). Recently, a method based on weighted scenarios of the fundamental directions of the patient's geometric variability was implemented (Söhn et al., 2005; Budiarto et al., 2011). These directions of variability, called motion/deformation modes, were obtained by applying principal component analysis (PCA) to a data set of (pre-)treatment organ geometries, usually parameterized by a set of *corresponding surface points*. This approach known as a *point-distribution model* (PDM) was introduced in statistical shape modelling and organ segmentation (Cootes et al., 1995; Lorenz and Krahnstöver, 2000; Davatzikos et al., 2001; Pekar et al., 2004).

PCA-based motion/deformation models were initially used by Söhn et al. (2005) for modeling individual geometric variations between fractions of the rectum, bladder, and prostate. The main goal was to estimate the patient's variance along each mode of geometric variability. These modes were derived by diagonalization of a data matrix, called the *covariance matrix*, whose columns were made of patient's motion/deformation vectors (i.e., vectors that describe the observed geometric variation of the organ at the on-treatment CTs/cone-beam CTs (CBCTs) relative to the mean shape). New organ geometries were then generated by adding a weighted sum of a few dominating modes to the mean organ shape, where each weight obeyed a Gaussian distribution with mean of zero and the corresponding eigenvalue as variance. PCA has also been applied to generate individual 4D statistical models for organs undergoing respiratory motion like the lung, liver, and heart (Zhang et al., 2007a; Badawi et al., 2010; McClelland et al., 2013). The same method was in turn applied to describe the geometric variations of organs like prostate, seminal vesicles, and pelvic lymph nodes (Thörnqvist et al., 2013).

However, this individual-based methodology was limited by the required number of CT scans per patient for training the model. In clinical practice, only one CT scan is available at the planning treatment for a typical patient, and on-treatment CTs/CBCTs are seldom available during treatment. Budiarto et al. (2011) thus proposed a population-based method that learned the geometric variations from a population database seeking to infer the possible organ variations for a typical patient despite his small number of CTs/CBCTs. Budiarto et al. (2011) applied his method to describe clinical target volume (CTV) motion and deformation between fraction in prostate cancer using a longitudinal database. It is worth mentioning that longitudinal data set consists of repeated observations of a set of homologous objects or variables at several time points (Durrleman et al., 2013a; Pinheiro and Bates, 2000). In our context, it corresponds to repeated observations of the organ at different fractions for a patient population. This methodology was also applied to model motion and deformation of the CTVs and bladder in rectal and prostate cancer radiotherapy, respectively (Bondar et al., 2014; Rios et al., 2016). In addition, Hu et al. (2015) proposed a population-PCA model to describe prostate deformation in magnetic resonance (MR)-tumor-targeted biopsies using a longitudinal database of prostates obtained from MR images and biomechanical models. Similarly, several researchers have introduced so-called 4D population-based or cross-population models for organs undergoing respiratory motion (He et al., 2010; McClelland et al., 2013; Preiswerk et al., 2014). In population-based models, modes were derived from the population covariance matrix made by the patients' motion/deformation vectors, i.e., all the rigid and non-rigid organ displacements observed for each patient around his mean bladder were stacked as columns. For a given patient, new organ samples were also generated by a weighted sum of a few dominating modes to the mean shape, whereby each weight also obeyed a Gaussian distribution with mean of zero and the corresponding eigenvalue as variance (i.e., the total population variance). However, this may lead to a suboptimal estimation of individual-specific variability, given that the mode variance is considered to be shared by all the individuals. This is an assumption that is not necessarily verified as the patients may exhibit heterogeneous organ variability.

In this paper, we sought to properly characterize patient's variance along each mode of geometric variability from the population variance. Next, we have proposed a longitudinal analysis using mixed-effects (ME) models to separate intra- and inter-patient variability along each mode in order to control confounding effects. We have also proposed an early step of dimensionality reduction using a SPHARM parameterization followed by PCA. We have thus found the modes of

geometric variation in the SPHARM coordinates rather than in the surface points as PDM-based shape models. This step enabled us to reduce the number of parameters required to represent the bladder surface, further encoded with the dominant modes derived by PCA. This step also allowed us to further reduce data storage and processing time, among others aspects.

Our main contribution regarding groundwork was to introduce ME models in order to estimate individual-specific variations between fractions from a reduced-order model (SPHARM/PCA based model). Using the ME model the following aspects were addressed for each mode: i) the estimation and separation of individual-specific variance from that of the population, i.e., the inherent correlation of intra-individual observations; ii) the reduction of the observed population variance by grouping repeated observations per patient by means of a “patient factor”; iii) the characterization of the individual-specific directions of motion and deformation from the population; iv) the study of longitudinal observations ordered in time, which may reflect temporal trajectories in the latent space. This study is also the first explicitly addressing the issue of predicting a region of probable inter-fraction bladder motion/deformation using solely the planning CT. A quantitative comparison of predicted *vs* observed probability maps was proposed using our model and other previous motion/deformation models.

Our proposed model is based on the assumptions stated in the methodology developed by Budiarto et al. (2011): firstly, that the bladder moves and deforms in a limited number of directions, constrained by the body's anatomy; secondly, that the inter-fraction geometric organ variations are randomly distributed along the set of variability directions; thirdly, that despite the potential variations in organ size and shape across the population, principal directions of geometric variation prove to be similar in all patients; fourthly, that geometric variations between fractions are independent random variables, where the time sequence of the observations does not matter. Hence, each bladder within the population can be characterized by its variation along each direction.

The paper is organized as follows. Section 1 describes the two databases (training and testing) included in this study. Section 2 provides a detailed description of the proposed method together with the validation framework. Section 3 presents the results and comparison to groundwork methods. Lastly, Section 4 discusses the experimental findings, with a conclusion given in Section 5 in addition to final considerations and future perspectives.

1. Data

We included two population databases of patients treated for prostate cancer with external beam radiotherapy (RT). We trained the model using data from a database of 20 patients, which included a planning CT and several on-treatment CTs (5-8) for each patient. We also used an independent validation database of 28 patients described as follows: 25 patients had a planning CT as well as 6-9 on-treatment CTs, and 3 patients with a planning CT as well as 35-39 on-treatment CBCTs for each. The contours of the bladder, rectum, prostate and seminal vesicles (SV) were manually delineated following the same protocol. The prescribed dose was computed in a standard treatment planning system (TPS) step then re-sampled into the CT native space. We used $135 \times 215 \times 55$ CTs and CBCTs with $1 \text{ mm} \times 1 \text{ mm} \times 2 \text{ mm}$ voxels. From the delineated CT and CBCT scans, observed-probability maps called PM_{obs} were computed for comparison with the prediction obtained using the proposed method, as explained in the validation section.

2. Methods

2.1. Population model training

The training steps are depicted in Fig. 1. Firstly, rigid-spatial normalization was performed in order to anatomically align the pelvic structures (prostate and bladder); secondly, bladder surface parameterization was conducted using SPHARM, followed by dimensionality reduction using PCA, thereby yielding a limited set of directions of maximum geometric variability. These directions, or modes, were validated by leave-one-out cross validation, aiming to determine whether patients exhibited common directions of bladder motion and deformation; finally, a ME model was fit to each direction to model the inter and intra-individual geometric variations obtained.

2.1.1. Prostate-based rigid registration

This was done to spatially align all the bladders in the same spatial referential. After defining the prostate barycenter of a patient as common coordinate system, henceforth called the template, all the patient structures were thus registered by aligning the prostate barycenters, as in Rios et al. (2016), thus yielding a transformation T . The template was a patient from the training database whose prostate exhibited the lowest volume variation and the most similar mean volume to the population mean. The individual bladders were then propagated to the common template using the obtained transformation T . This enabled bladder motion and

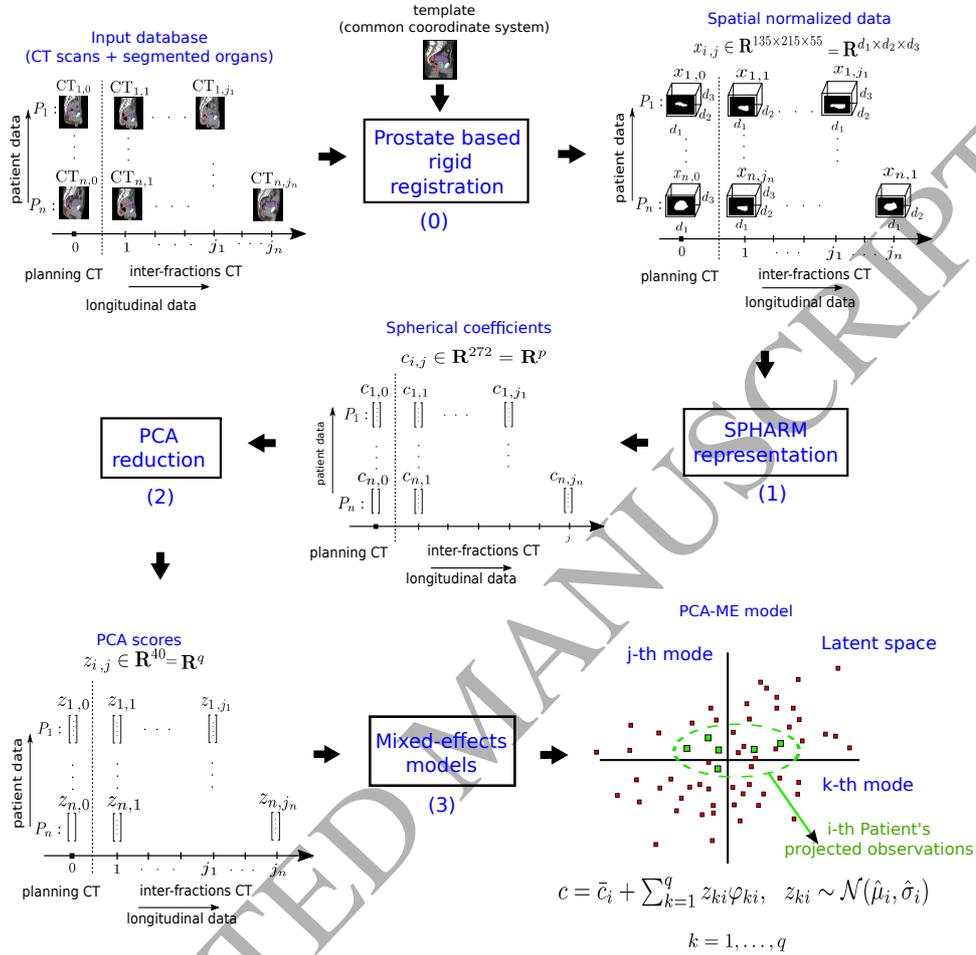


Figure 1: Workflow of the method used to train the population-based model. 0) Prostate-based registration, 1) spherical harmonics (SPHARM) representation, 2) principal component analysis (PCA) reduction, 3) ME model.

deformation to be quantified with respect to the template prostate barycenter. Fig. 2 depicts an example of the performed rigid registration/propagation step.

2.1.2. Spherical harmonics

An essential objective in shape modeling has been the reduction of parameters needed to describe a shape instance. As a first step towards this goal, PDM method was proposed to parameterize 3D shapes by means of a set of boundary points (Cootes et al., 1995; Lorenz and Krahnstöver, 2000). However, this

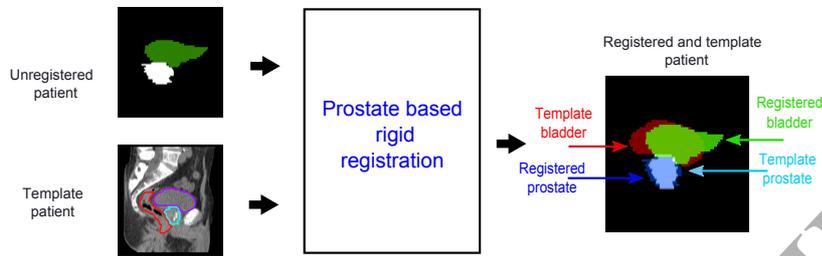


Figure 2: Overlay of a typical individual bladder registered to the template.

approach required thousands of points to properly describe shape surfaces and a one-to-one correspondence between these boundary points (Yan et al., 1999; Lorenz and Krahnstöver, 2000; Söhn et al., 2005; Budiarto et al., 2011; Bondar et al., 2014). SPHARM, the extension of 2D Fourier techniques to three dimensions, was proposed to obtain smooth, accurate and fine parametrization of 3D shapes with spherical topology with a fewer number of parameters (Ballard and Brown, 1982; Brechbühler et al., 1995; Székely et al., 1996; Kelemen et al., 1999). This method consisted in expressing a surface as a linear combination of a reduced number of spherical harmonics basis functions, where different levels of approximation errors can be obtained by truncating the spherical harmonic series. As opposed to other methods like PDM, it offers the advantage of delivering a shape parametrization with an implicit correspondence between shapes on the boundary. Besides, the spherical harmonics define an orthogonal basis functions whose spanned space has a reduced dimension that enables us to describe and quantify shape variations. This implies that the number of parameters (i.e., SPHARM coefficients) needed to describe a shape instance is lower than other shape modeling methods like PDM.

In medical image, SPHARM has played an crucial role in 3D-shape representation and modeling of anatomic structures in several applications, including computer assisted diagnosis (Shen et al., 2004; Styner et al., 2004, 2006; Shen et al., 2009b,a), rigid registration (Dillenseger et al., 2006; Shen et al., 2009b) and organ segmentation (Brechbühler et al., 1995; Székely et al., 1996; Kelemen et al., 1999; Gerig et al., 2001; Tateyama et al., 2012), among others. In most of these applications, statistical shape models were developed by applying PCA to a training population of organ surfaces parameterized with SPHARM coefficients, where eigenmodes of maximum variation around the mean shape were calculated. In these applications, the surfaces of these binary organs were initially parameterized by surface boundary points, and subsequently expanded into series

of spherical harmonics.

In our study, SPHARM was used to parameterize the surface of each bladder into a vector of SPHARM coordinates. Each bladder surface had the same radial surface sampling in spherical coordinates, maintaining the correspondences across all the points. Following mesh parameterization, each bladder represented by function $f \in \mathcal{R}^3$ was encoded as a reduced number of spherical harmonic functions, as follows:

$$\begin{aligned} f(\theta, \phi) &= \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi), \\ &\approx \sum_{l=0}^L \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi), \quad 0 < L < \infty \end{aligned} \quad (1)$$

where θ and ϕ denoted the spherical coordinates, and with

$$Y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l-m)!}{(l+m)!}} P_l^m(\cos(\theta)) e^{im\phi} \quad (2)$$

defining a spherical harmonic function of degree l and order m , where P_l^m was an associated Legendre polynomial; L was the number of spherical harmonic functions used to approximate f ; and coefficients c_l^m corresponded to the coordinates in the space spanned by the spherical harmonic basis. The c_l^m coefficients were obtained as the projection of the mesh function f onto the spherical harmonic basis, i.e., $c_l^m = \langle f(\theta, \phi), Y_l^m(\theta, \phi) \rangle$. By fixing a harmonic degree to L , the bladder surface at the j -th CT scan of the i -th patient, denoted by $x_{i,j} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, was uniquely represented in spherical harmonic coordinates as follows:

$$c_{ij} = [c_{ij,0}^0 \quad c_{ij,1}^0 \quad c_{ij,1}^1 \quad c_{ij,2}^0 \quad \cdots \quad c_{ij,L}^L \quad c_{ij,1}^{-1} \quad c_{ij,2}^{-1} \quad \cdots \quad c_{ij,L}^{-L}]^T \quad (3)$$

where

$$c_{ij} \in \mathbb{R}^p \quad \wedge \quad p = 2 \left[(L+1)^2 - \frac{(L+1)^2 - (L+1)}{2} \right] \ll d = d_1 \times d_2 \times d_3 \quad (4)$$

Observe that we have taken only the SPHARM coefficients that were non-zero. The value of L was optimally chosen by analyzing its impact on the approximation

error in the training cohort. The approximation error was assessed as a function of L in terms of dice score (DS) and the Hausdorff distance (HD).

2.1.3. Principal component analysis

This step was aimed to represent, in a reduced space, all possible rigid and non-rigid bladder displacements. After obtaining SPHARM coefficients, an empirical covariance matrix \mathbf{C} of local patient motion-deformation vectors was thus defined in order to calculate the directions of maximum variability by using matrix diagonalization. Here, a local patient motion-deformation vector defined the difference between the observed bladder and patient mean bladder. Thus, denoting the observed bladder of the i -th patient at the j -th CT as c_{ij} , the empirical covariance matrix \mathbf{C} was defined as follows:

$$\begin{aligned}\mathbf{C}_{empirical} &= \frac{1}{n} \sum_{i=1}^n \frac{1}{j_i - 1} \sum_{j=1}^{j_i} (c_{ij} - \bar{c}_i)(c_{ij} - \bar{c}_i)^T \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{j_i - 1} \sum_{j=1}^{j_i} d_{ij}d_{ij}^T = \mathbf{U}\mathbf{D}\mathbf{U}^T\end{aligned}\quad (5)$$

where $d_{ij} = c_{ij} - \bar{c}_i$, n was the number of patients, j_i was the number of observations available for the i -th patient, and \bar{c}_i the average shape of the i -th patient. Similarly, \mathbf{U} was a matrix composed of eigenvectors φ_k of $\mathbf{C}_{empirical}$, and \mathbf{D} was the diagonal matrix constructed from its corresponding eigenvalues λ_k , *i.e.*, $\mathbf{C}_{empirical}\varphi_k = \lambda_k\varphi_k$. Eigenvectors φ_k defined the directions of geometric variability of the bladder, called modes. As a result, each bladder in the spherical harmonic space was expressed as a linear combination of the modes, as follows:

$$c_{ij} = \bar{c}_i + z_{ij1}\varphi_1 + \cdots + z_{ijr}\varphi_r \quad (6)$$

where $z_{ijk} = d_{ij}^T\varphi_k$, and r was the rank of covariance matrix $\mathbf{C}_{empirical}$, and equal to the number of available images in the population database. Modes $\{\varphi_k\}_k$ defined a new coordinate system, where each vector d_{ij} was represented by new r coordinates z_{ij1}, \dots, z_{ijr} . The eigenvalues of matrix \mathbf{D} defined a significance ranking for the data variability represented by each mode, attesting $var(\varphi_1) = \lambda_1 \geq var(\varphi_2) = \lambda_2 \geq \cdots \geq var(\varphi_r) = \lambda_r$, and showing that data variability was dominated by the first q modes with $q \ll r$.

In order to test whether the modes described common directions of bladder motion and deformation among all patients, we performed a leave-one out cross validation on the training database. Thus, the error assessing the difference between the original and reconstructed motion-deformation vector was defined as follows:

$$e_i = \frac{\|d_i - \hat{d}_i\|}{\|d_i\|} \quad (7)$$

where d_i and \hat{d}_i were the i -th bladder motion-deformation vector and its approximation, respectively. If the reconstruction of d_i was exact, then e_i was equal to 0; otherwise, the value of e_i was between 0 and 1. The mean of the reconstruction error e_i in all the motion-deformation vectors was also calculated, i.e., $\mathcal{M} = (1/N) \cdot \sum_{i=1}^N e_i$, where e_i denoted the approximation error of an i -th motion-deformation vector d and N the number of the available images, either in the PCA model or left-out patients.

2.1.4. Linear mixed-effects model

Several longitudinal studies have been designed to investigate anatomical or functional changes over time in objects or variables that are observed repeatedly at several time points (Durrleman et al., 2013b; Laird and Ware, 1982; Donald Hedeker, 2006; Fitzmaurice et al., 2011; Frees, 2004). For instance, these repeated observations may be brain volume, response to clinical treatments, disease progression, blood pressure, heart motion or tumor evolution, among others (Prabhu et al., 2014; Hanlon et al., 1998; Bastogne et al., 2010; Bernal-Rusiel et al., 2013; Ziegler et al., 2015). In these longitudinal studies, observations may be obtained under changing experimental conditions that are not possible to fully control, leading to considerable variations among subjects in the number and time points of the observations. The resulting data sets, which are known as unbalanced data, have often analyzed using ME models that offer to simultaneously model both the random effects that differentiate one subject from a population and the inherent correlation of repeated measurements in the same subject (Liu, 2016; Frees, 2004; Pinheiro and Bates, 2000; Fitzmaurice et al., 2011; Donald Hedeker, 2006; Liang and Zeger, 1986; Laird and Ware, 1982). A linear ME model is then a statistical model that incorporates both “fixed effects”, which are parameters related to the entire population or a certain level of classification factor, as well as “random effects”, which are parameters associated with differences among individual subjects drawn at random from a population.

The use of basis functions and ME have already proposed in the literature to account for random effects in longitudinal data. These basis function methods include: smoothing splines, wavelet basis, polynomial basis, and functional principal components (Guo, 2002; Morris and Carroll, 2006; Wu and Zhang, 2002; Aston et al., 2010). In our study, the longitudinal data was encoded in the PCA modes, where the observed motion-deformation vectors in the training database were projected over each mode while a linear ME model was fit to each. The score resulting from each projection could be interpreted as a measure of how much a direction of geometric variability was presented in a motion-deformation vector. The implemented ME model thus aimed the characterization of population and individual bladder variations along each mode by introducing a “patient” factor modeled as a random effect.

Let z be the measure of motion/deformation of an organ along a mode. Considering the patient as the unit of observation state (subject), we used an index i , ranging from 1 to n , to differentiate among patients, and an index j to differentiate between observation times in a patient. The encoded longitudinal data was thus defined as follows:

$$\{z_{11k}, \dots, z_{1j_1k}\}, \{z_{21k}, \dots, z_{2j_2k}\}, \dots, \{z_{n1k}, \dots, z_{nj_nk}\}, \quad k = 1, \dots, q \quad (8)$$

with the linear ME model proposed for each mode described as follows:

$$\begin{aligned} z_{ijk} &= \mu_k + b_{ik} + \varepsilon_{ijk}, \quad i = 1, \dots, n, \quad j = 1, \dots, j_i, \\ b_{ik} &\sim \mathcal{N}(0, \sigma_{bk}^2), \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma_k^2), \end{aligned} \quad (9)$$

where μ_k was the mean projection along the k -th mode for the whole population, b_{ik} was a random variable representing the deviation of the i -th patient mean from the population mean, and ε_{ijk} was a random variable representing the deviation for j -th projection of the i -th patient from the i -th patient mean. In addition, variances σ_{bk}^2 and σ_k^2 denote inter and intra-patient variability, respectively. The population based model is thus described as follows:

$$c = \bar{c}_i + Wz \quad (10)$$

$$c = [c_0^0 \quad c_1^0 \quad \dots \quad c_L^{-L}], \quad W = [\varphi_1 \quad \varphi_2 \quad \dots \quad \varphi_q], \quad z = [z_1 \quad z_2 \quad \dots \quad z_q]$$

where \bar{c}_i was the patient mean shape, $z_k \sim \mathcal{N}(\mu_k, \sigma_{bk}^2 + \sigma_k^2)$, for $k = 1, \dots, q$.

2.2. Out-of-sample problem

We sought to predict a likely motion-deformation region of the bladder for a new patient using only his planning CT information. We described regions of motion/deformation in terms of 3D probability maps (PM) that express the probability of a voxel being occupied by the bladder during treatment. As shown in Fig. 3, the first step was to spatially normalize the out-of-sample patient P_l to match the template patient; secondly, the SPHARM and PCA representations of his delineated bladder $c_{l,0}$ and $z_{l,0}$ were obtained; and thirdly, each ME model was adapted to the new patient by estimating his patient-specific mean and variance $\hat{\mu}_{lk}$ and $\hat{\sigma}_{ljk}^2$.

In ME models, random subject-specific effects are summarized to predict the response of a new subject. The predictor is a linear combination of both fixed and random effects, and it can be interpreted as a *shrinkage estimator*. Let us suppose that the l -th patient was included in the data set and we want to estimate the following linear model:

$$z_{ljk} = \mu_k + b_{lk} + \epsilon_{ljk} \sim \mathcal{N}(\hat{\mu}_{lk}, \hat{\sigma}_{ljk}^2), \quad j = 1, \dots, j_l, k = 1, \dots, q \quad (11)$$

A shrinkage estimator of $\hat{\mu}_{lk} = \mu_k + b_{lk}$ is then given by:

$$\hat{\mu}_{lk} = \mu_k + \zeta_k (\mu_{lk} - \mu_k), \quad \zeta_k = \frac{j_l}{j_l + \sigma_k^2 / \sigma_{bk}^2}, \quad k = 1, \dots, q \quad (12)$$

where μ_{lk} is the mean of the l -th patient at the k -th mode, and j_l is the number of patient's observations. Term ζ is known as the credibility factor. Similarly, the variance of the prediction error $\hat{\sigma}_{ljk}^2 = \text{var}(\epsilon_{ljk})$ is given by:

$$\hat{\sigma}_{ljk}^2 = \left(1 - \frac{\sigma_{bk}^2}{\sigma_k^2 + \sigma_{bk}^2}\right) \frac{n}{\sigma_k^2 + \sigma_{bk}^2} \left(1 - \frac{\sigma_{bk}^2}{\sigma_k^2 + \sigma_{bk}^2}\right) - \frac{(\sigma_{bk}^2)^2}{\sigma_k^2 + \sigma_{bk}^2} + \sigma_k^2 + \sigma_{bk}^2 \quad (13)$$

In our study, $\bar{c}_l = c_{l,0}$, $\mu_l = W^T(\bar{c}_l - c_{l,0}) = z_l = 0$ and $j_l = 1$ due to we only had the information obtained from the planning CT. In this way, probable structures of the bladder during treatment were thus inferred after sampling S times the Normal distributed vector z . Subsequently, each sample of z was mapped to a vector c in the spherical harmonic space using matrix W . Then, binary images representing probable bladder configurations were obtained from the parameterized surfaces that were coded in vectors c . Following this, an estimation of the motion/deformation region was calculated by the ratio between the number of times that a voxel was occupied by the bladder and the S generated samples.

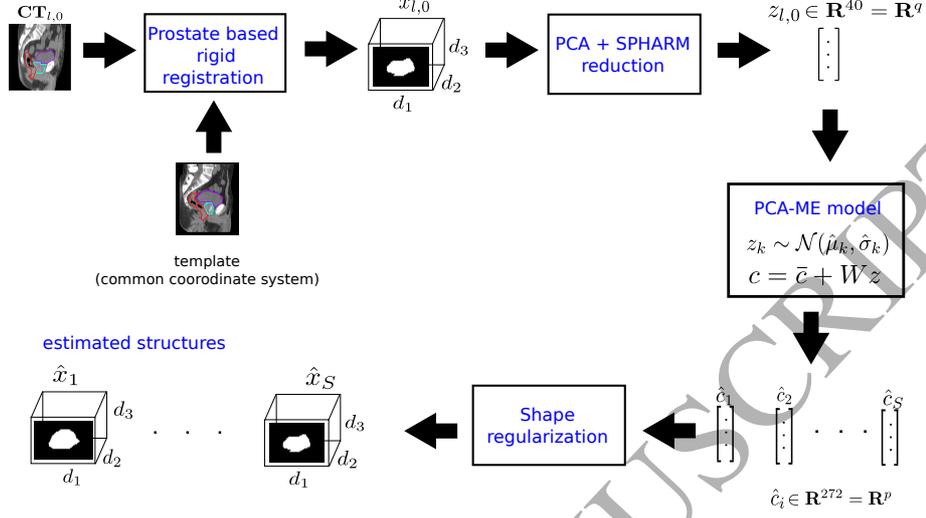


Figure 3: Schematic representation for estimating new structures based on the planning CT for a new patient.

2.3. Evaluation of prediction performance

Two additional population PCA-based models were used to compare the model's prediction performance. Considering motion-deformation vectors computed between the observed bladders and population mean bladder, the PCA model implemented by Söhn et al. (2005) was extended to the population data as follows:

$$c = \bar{c}_{pop} + \sum_{k=1}^q z_k \psi_k, \quad (14a)$$

$$\text{Global PCA model} := z_k \sim \mathcal{N}(0, \alpha_k) \quad \wedge \quad \bar{c}_{pop} = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^{j_i} c_{ij} \quad (14b)$$

$$\frac{1}{N-1} \sum_{i=1}^n \sum_{j=1}^{j_i} (c_{ij} - \bar{c}_{pop})(c_{ij} - \bar{c}_{pop})^T \psi_k = \alpha_k \psi_k, \quad (14c)$$

where $N = \sum_{i=1}^n j_i$ was the total number of observations in the training database, \bar{c}_{pop} was the population mean bladder, and vectors ψ_k were the eigenvectors of the empirical covariance matrix centered in terms of the population mean bladder. Observe that in this model it was thus considered that all the observed bladders in

the training set came from a single individual. The second model proposed was taken from Budiarto et al. (2011), defined as follows:

$$c = \bar{c}_i + \sum_{k=1}^q z_k \varphi_k \quad (15a)$$

$$\text{Local PCA model} := z_k \sim \mathcal{N}(0, \lambda_k) \quad \wedge \quad \bar{c}_i = \frac{1}{J_i} \sum_{j=1}^{J_i} c_{ij} \quad (15b)$$

$$\mathbf{C}_{\text{empirical}} \varphi_k = \lambda_k \varphi_k, \quad (15c)$$

where \bar{c}_i was the mean bladder of the i -th patient, and vectors φ_i were the eigenvectors of empirical covariance $\mathbf{C}_{\text{empirical}}$. Then, given a new patient with planned delineated bladder z_l , both PCA models were adapted as follows:

$$\text{Global PCA model} : c = c_{l,0} + \sum_{k=1}^q z_k \psi_k, \quad \wedge \quad z_k \sim \mathcal{N}(z_l, \alpha_k) \quad (16)$$

$$\text{Local PCA model} : c = c_{l,0} + \sum_{k=1}^q z_k \varphi_k, \quad \wedge \quad z_k \sim \mathcal{N}(z_l, \lambda_k) \quad (17)$$

Both PCA based models used the modes to transform the delineated bladder on the planning CT and predict new bladder structures. It is worth mentioning that both sets of modes $\{\psi_k\}$ and $\{\varphi_k\}$ expressed different directions of organ motion and deformation. For example, modes $\{\psi_k\}$ described motion/deformation deltas regarding the population mean bladder whereas modes $\{\varphi_k\}$ described local patient motion/deformation deltas.

The models' prediction performance was evaluated by following leave-one-out cross validation in the training database and using the independent database. We thus estimated probability maps with each model and compared with the observed PM in both databases. Estimated PMs were derived for each model and patient, as follows: firstly, 40 bladder structures were generated using the bladder observed at the planning CT (i.e. $S = 40$, see Fig. 3); and secondly, a distribution of 15 PMs were derived from the 40 estimated structures, where each PM was obtained from a sub-set of estimated bladders (without replacement) whose cardinality was equal to the number of available CT/CBCTs.

Four PMs were thus calculated for each out-sample patient: one PM estimated by means of the proposed model, labeled $PM_{\text{Local PCA-ME}}$; one PM obtained

from the patient's observed available images, labeled PM_{obs} ; one PM estimated by means of the global PCA model, labeled $PM_{Global\ PCA}$; and one PM estimated by using the local PCA model, labeled $PM_{Local\ PCA}$.

2.4. Metrics of similarity between PMs

Two metrics were proposed for assessing the similarity between two PMs X and Y .

2.4.1. Mutual information-based metric

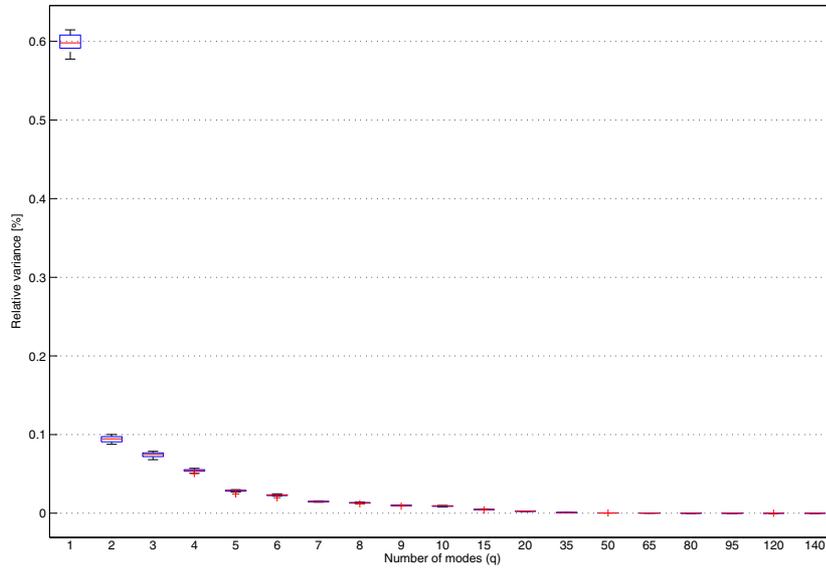


Figure 4: Relative mode variances obtained from the PCA in the training database (relative values: sum of all eigenvalues normalized to 100%).

$$dist_{intensity}(X, Y) = 1 - \frac{I(X, Y)}{H(X, Y)} \quad (18)$$

$$H(X, Y) = - \sum_{x,y} p_{XY}(x, y) \log p_{XY}(x, y)$$

$$I(X, Y) = H(X) + H(Y) - H(X, Y)$$

where $p(X, Y)$ is the joint probability distribution, and $I(X, Y)$ and $H(X, Y)$ denoted mutual information (MI) and joint entropy between X and Y , respectively. Similarly, x and y corresponded to intensity values in the voxels of X and Y , respectively. $H(X)$ and $H(y)$ denoted marginal entropies of each image. MI measures the amount of information that both images share (redundant information) based on the intensities of corresponding voxels, which is assumed to be maximal if both images are geometrically aligned. Similarly, joint entropy measures the uncertainty or dispersion between images. $dist_{intensity}$ also provided a distance based on the ratio of dependence and dispersion between two PMs, attesting $dist_{intensity}(X, Y) = 0$ iff $X = Y$, and $dist_{intensity}(X, Y) \leq 1$ for all pairs (X, Y) (Kraskov and Grassberger, 2009; Russakoff et al., 2004; Maes et al., 1997). It offers the advantage of not making any prior assumptions regarding the functional relationship between images. Since we were dealing with binary images, each voxel was considered as a discrete random variable. Joint and marginal entropies were calculated by normalizing the joint, and marginal histograms, i.e. a (joint) probability distribution of intensity values were estimated by counting the number of times each value occurred in the images and dividing this value by the total number of voxels.

2.4.2. Metric of misestimated voxel

We also proposed an error metric to assess whether a voxel was overestimated or underestimated. Denoting v as a voxel, let $X_b = \{v/X(v) > 0\}$ be the binary image obtained by those voxels with a probability greater than 0. The proposed metric is defined as follows:

$$\text{Accuracy}(X, Y) = \frac{|X_p \cap Y_p| + |\bar{X}_p \cap \bar{Y}_p|}{d} \quad (19)$$

where d denoted the total number of voxels (see Eq. 4), $|\cdot|$ denoted the cardinality, and \bar{X}_p and \bar{Y}_p denoted the complement of X_p and Y_p , respectively. As a result, this metric provided the proportion of the properly estimated voxels (i.e. both true positives and negatives voxels in the motion/deformation region) among the total number of voxels.

3. Results

The analysis was carried out in MATLAB (R2014a). The eigenvectors φ_k and ψ_k were obtained using the function *eigs* while the ME model of each score z_{ijk}

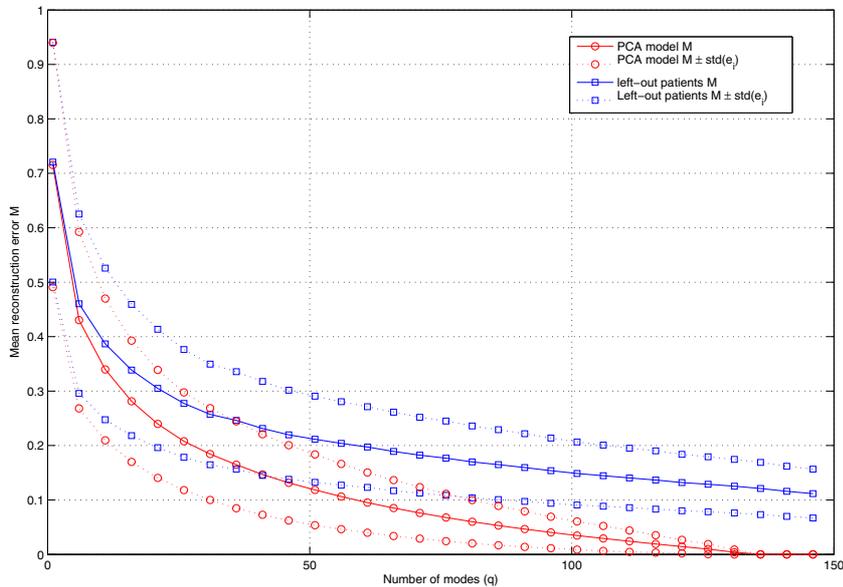


Figure 5: Mean approximation error M as a function of the number of modes in the PCA model and left-out patients. There are bands (dotted-lines) that represent the mean approximation error M (solid-lines) plus and minus one standard deviation of e_i (see Eq. 7).

was fit using the function *fitlme* with *restricted maximum likelihood estimation* (REML) as estimation method (Pinheiro and Bates, 1996, 2000).

3.1. SPHARM and PCA representation

In the first phase of dimensionality reduction, an L of 15 spherical harmonics was selected to approximate bladder surfaces in the population database. With $L = 15$ in Eq. 1, the average DS and HD values in the training database were 0.954 ± 0.0522 and 6.112 ± 4.997 mm, respectively. Conversely, validation database equivalents were 0.97 ± 0.0099 and 4.32 ± 2.27 mm, respectively. Any bladder in the databases can thus be represented by $p = 272$ coefficients instead of $d = 1596375$ voxel variables.

Fig. 4 shows the eigenvalue spectrum of all 20 combinations of splitting the training database from the leave-one-out cross validation. The number of columns of the covariance matrix $\mathbf{C}_{empirical}$ varied between 151 and 154 (i.e. the available images after removing the left-out patient). The first mode contributed the

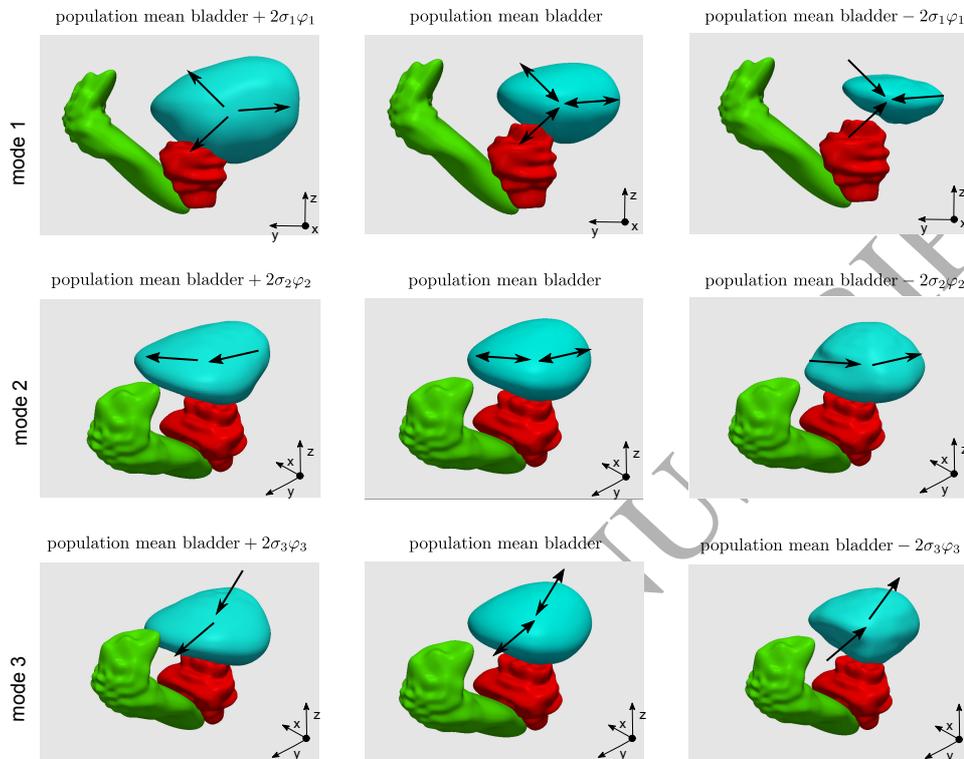


Figure 6: 3D sagittal views of the first three motion/deformation modes applied on the population mean bladder (center column). The left and right columns correspond to geometric variations along the respective mode φ_i , defined as $population\ mean\ bladder \pm 2\sigma_i\varphi_i$, where σ_i^2 is the mode variance. The prostate (red color) and rectum (green color) of the template patient were also included as spatial reference.

largest variance (approximately 60% of the variability); similarly, the second, third, and fourth modes contributed, on average, 9%, 7.8% and 5% of the variability, respectively. The first eight modes contributed 90% of the cumulative variance, on average. Meanwhile, Fig. 5 shows that both mean and standard deviation of the approximation error \mathcal{M} decreased as the number of modes was increased in both the PCA model and left-out patients. However, mean and standard deviation of the reconstruction error \mathcal{M} did not reach the value of zero in the left-out patients, and they exhibited a lower decreasing rate after the first eight modes. Considering only the first 60 modes, the mean reconstruction error \mathcal{M} was 10% and 20% in the PCA model and left-out patient, respectively. When truncating the first 40, it still obtained an average accumulated variability of 98%

and an average reconstruction error of 15% and 23% for the PCA model and left-out patients, respectively. Therefore, after performing two phases of dimensionality reduction, any bladder was represented as a vector of $q = 40$ scores ($q = 40 \ll p = 272 \ll d = 1596375$). Fig. 5 also shows bands (dotted lines) around the mean approximation error \mathcal{M} , which represent the mean plus and minus one standard deviation of e_i (see Eq. 7). It can be observed that the variance also decreased as the number of modes was increased in both the PCA model and left-out patients.

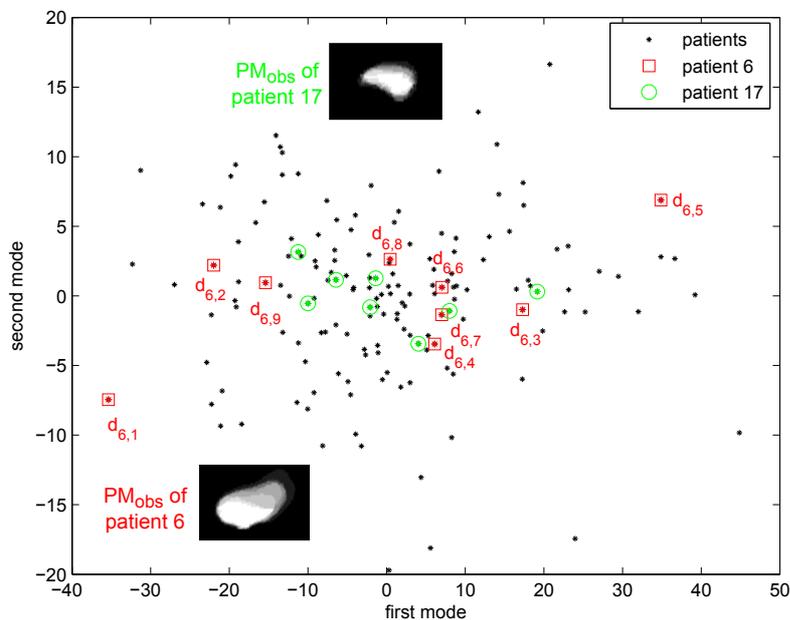


Figure 7: Scatterplot of the patients from the training database projected in the first two modes (black points). Color circles and squares correspond to the vectors of two individuals. $d_{6,j}$ denotes the observed motion/deformation vectors of Patient 6's bladder.

Fig. 6 presents sagittal views of the first three modes over the mean bladder of the training database. The first mode was associated with bladder volume. Similarly, modes with greater indices, like the second and third, represented directions of dilation and contraction in some regions. For example, the second mode indicated dilation/contraction in the left-dorsal and right-ventral regions of the bladder whereas the third mode indicated dilation/contraction in the inferior-ventral and

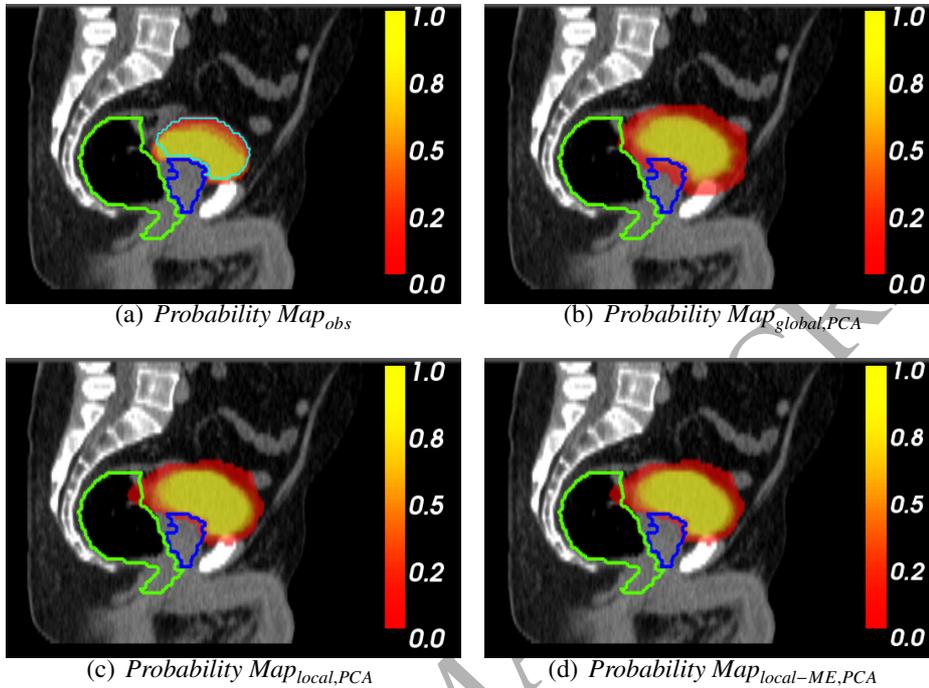


Figure 8: 3D sagittal views of PMs obtained for one patient in the training database as an out-of-sample patient. Segmented prostate (blue) and rectum (green) at the planning CT scan are overlaid. Top left figure also depicts the segmented bladder (cyan) at the planning CT.

dorsal regions.

Fig. 7 depicts the projection of the motion/deformation vectors of the training patients on the space generated by the first two modes, where each point represents the observed geometric variations along these two modes. This figure also presents the observed motion/deformation vectors of Patient 6, indicating that the geometric variation followed a random trajectory. In addition, we can observe that Patient 6 and 17 had variance that differed both between them and from the mode's population variance. For example, Patient 6's bladder exhibited greater changes in volume during treatment than that of Patient 17, as shown in the PMs of both patients, proving that Patient 6's specific-variance along the first mode was greater than that of Patient 17. In addition, we also observed that Patient 17's variance along the second mode was very slight due to the geometric variations of his bladder being greater in the inferior-medial and dorsal regions.

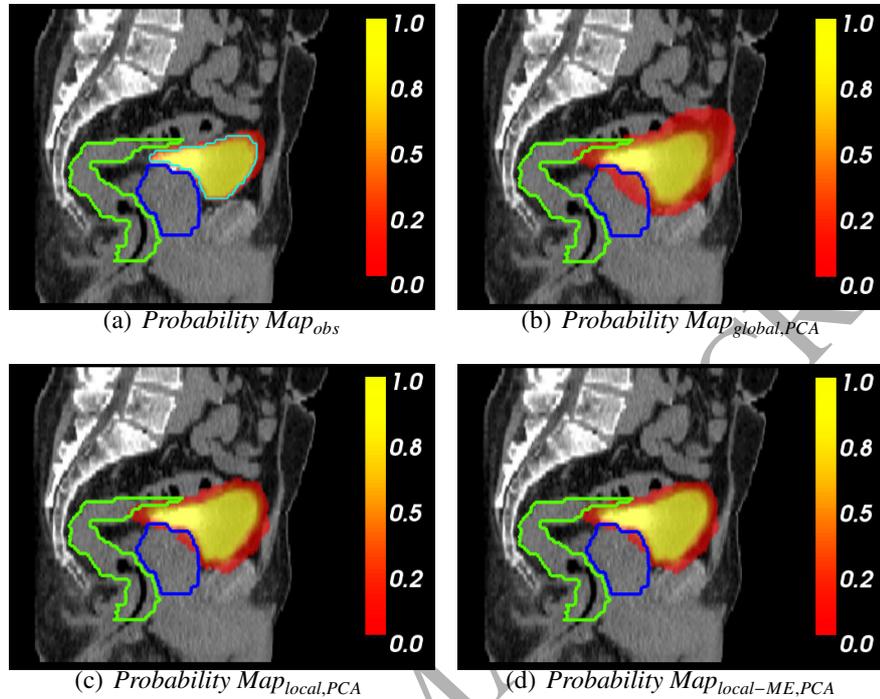


Figure 9: 3D sagittal views of PMs obtained for one patient in the validation database as an out-of-sample patient. Segmented prostate (blue) and rectum (green) at the planning CT scan are overlaid. Top left figure also depicts the segmented bladder (cyan) at the planning CT.

3.2. Evaluation of performance prediction with groundwork models

Fig. 8 and 9 demonstrates the PMs obtained by all three models and the available observations of two patients as out-sample patients (one for each database). In both patients, the Global and Local PCA model overestimated more voxels than our proposed model. Tables 1 and 2 present the averages obtained by these metrics for these patients.

Fig. 10 depicts the averages of the metric $dist_{intensity}$ for patients in both databases. On average, the global PCA model obtained the highest measures of dissimilarity in both databases while our proposed model obtained the lowest measures. Similarly, Fig. 11 also shows the averages of the metric $Accuracy$ for patients in both databases. On average, the global PCA model also obtained the lowest accuracy in both databases while our proposed model obtained the highest. For each patient and metric, a significant test was made between the PCA models using the Wilcoxon rank sum test for all 15 $dist_{intensity}$ and $Accuracy$ va-

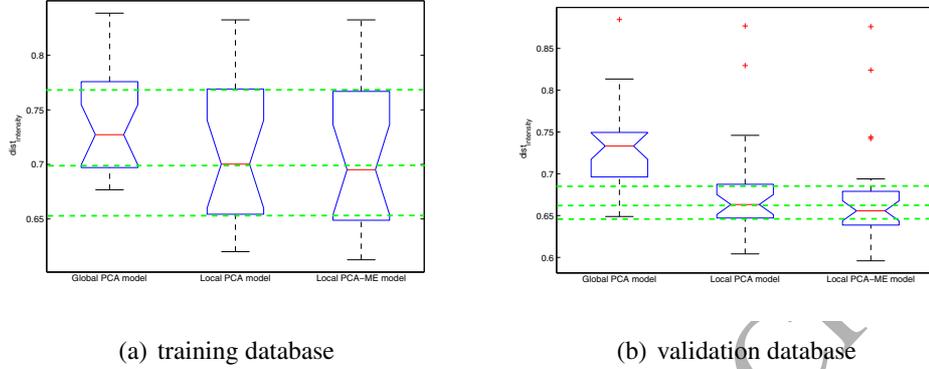


Figure 10: Average $dist_{intensity}$ values of observed and estimated PMs for patients in both databases.

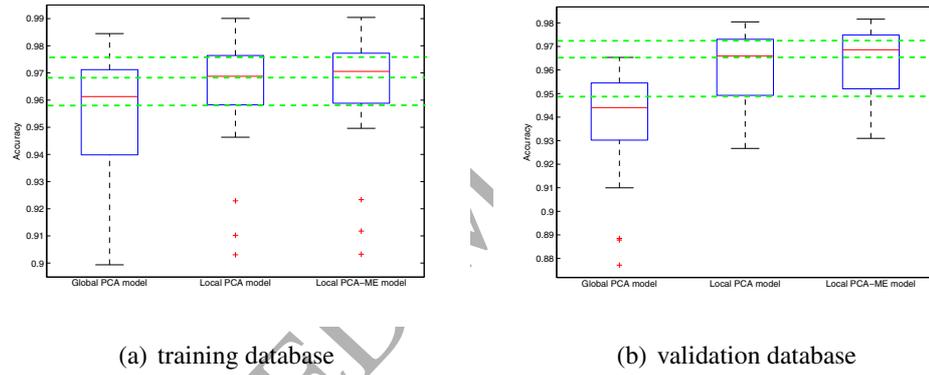


Figure 11: Average $Accuracy$ values of observed and estimated PMs for patients in both databases.

lues. The Table 3 shows the number of patients that had significant difference in both databases. Besides, we also obtained four variance distributions by deriving the variance of the 15 values for each metric and patient in both databases. Table 4 presents the p-values of the significant tests that were made between these variance distributions of the PCA models using the Wilcoxon rank sum test.

Fig. 12 also presents the average joint histograms of the voxel intensities of PM_{obs} vs $PM_{Local\ PCA-ME} - PM_{Global\ PCA}$ and PM_{obs} vs $PM_{Local\ PCA-ME} - PM_{Local\ PCA}$ in both database. This presents the number of times that a joint intensity value was repeated between the observed and estimated PMs. For example, Figs. 12a and 12c reveal that the global PCA overestimated more voxels than our proposed model in both databases in probabilities p ranging between 0 and 0.6 instead of

Table 1: Averaged values from the metric $dist_{intensity}$ for two patients between estimated and observed PMs.

Patient	PM_{obs}	PM_{obs}	PM_{obs}
	VS	VS	VS
	$PM_{Global\ PCA}$	$PM_{Local\ PCA}$	$PM_{Local\ PCA-ME}$
Training database	0.691	0.676	0.66
Validation database	0.74	0.66	0.65

Table 2: Averaged values from the metric $Accuracy$ for two patients between estimated and observed PMs.

Patient	PM_{obs}	PM_{obs}	PM_{obs}
	VS	VS	VS
	$PM_{Global\ PCA}$	$PM_{Local\ PCA}$	$PM_{Local\ PCA-ME}$
Training database	0.968	0.972	0.98
Validation database	0.946	0.97	0.975

the observed value of $p = 0$. Meanwhile, Figs. 12b and 12d demonstrate that our proposed model was also more capable of estimating voxels observed with probability $p = 0$.

4. Discussion

We have hereby proposed a new methodology for predicting bladder motion and deformation between fractions in prostate cancer radiotherapy via population analysis. We calculated bladder motion and deformation regions based solely on the planning CT scan. In order to predict which regions were likely to exhibit motion/deformation, we first performed a prostate-based rigid-registration

Table 3: Number of patients exhibiting significance difference using the Wilcoxon rank sum test in both databases.

Database	Local PCA-ME model vs Global PCA model	Local PCA model vs Global PCA model	Local PCA-ME model vs Local PCA model
	$dist_{intensity}/Accuracy$	$dist_{intensity}/Accuracy$	$dist_{intensity}/Accuracy$
Training	17/18	17/14	0/0
Validation	20/27	20/26	5/3

Table 4: p-values of the Wilcoxon rank sum tests for the variance distributions between PCA models in both databases.

Database	Local PCA-ME model vs Global PCA model	Local PCA model vs Global PCA model	Local PCA-ME model vs Local PCA model
	$dist_{intensity}/Accuracy$	$dist_{intensity}/Accuracy$	$dist_{intensity}/Accuracy$
Training	0.0411/0.0001	0.0720/0.003	0.6/0.32
Validation	$0.037/3 * 10^{-5}$	0.10/0.0001	0.51/0.44

alignment, two dimensionality reduction steps, and fitting regression. The first dimensionality reduction step consisted of surface parameterization of the bladder in the spherical harmonic space. The second step of dimensionality reduction was to determine the directions of bladder geometric variability by means of PCA. These directions, called modes, were validated by leave-one-out cross validation. Following this, an ME model was fitted along each mode to the longitudinal data defined by the projections of motion/deformation vectors observed in the train-

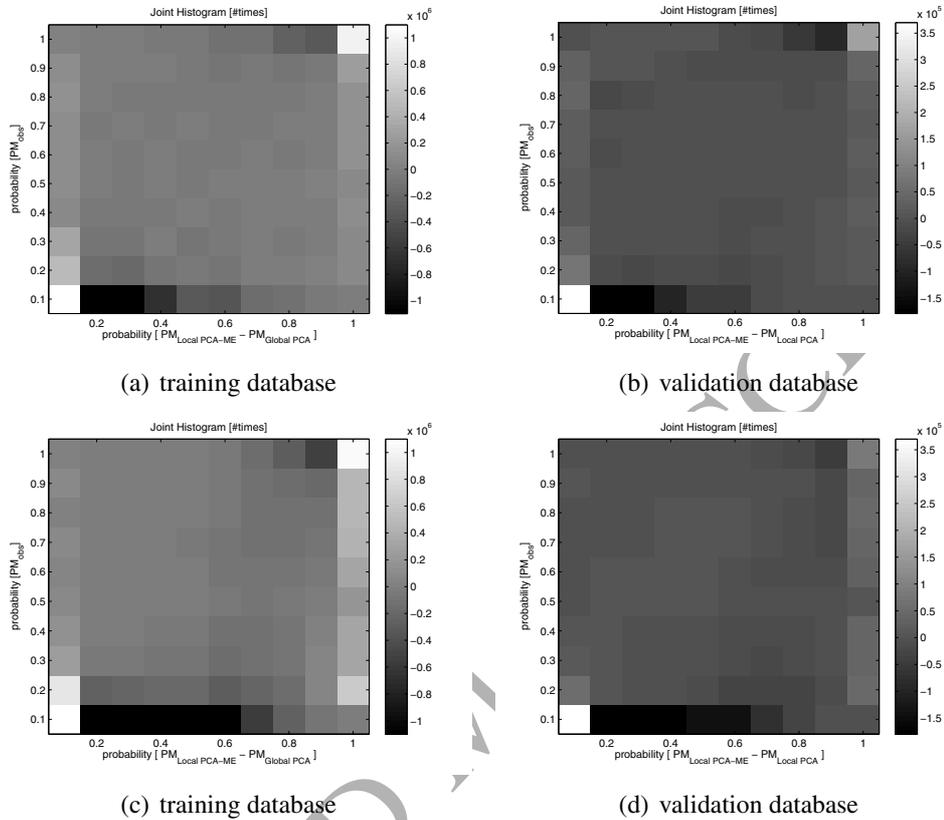


Figure 12: Average joint histograms of observed and estimated PMs for patients in the training and validation database. Left column shows the average joint histograms between PM_{obs} vs $PM_{Local\ PCA-ME} - PM_{Global\ PCA}$ while right column shows the average joint histograms between PM_{obs} vs $PM_{Local\ PCA-ME} - PM_{Local\ PCA}$.

ing database. We used leave-one-out cross validation on the training database and an external database to validate the model. The proposed model was then compared to two additional population PCA based models. Motion/deformation regions were characterized by means of probability maps (PMs) and metrics were provided in order to measure similarities between observed and estimated PMs.

Previous studies have matched CT scans for modeling geometric variations of the pelvic organs based on body anatomy, soft-tissue registration, or fiducial markers. For example, Söhn et al. (2005) used bone anatomy to match for each patient the repeat CT scans to the planning CT in order to obtain an intra-patient model for the prostate/bladder/rectum. Thörnqvist et al. (2013) used fiducial markers to

model individual geometric variations of the prostate. In contrast, Hu et al. (2015) proposed anatomical apex and base points of the prostate. Meanwhile, Budiarto et al. (2011) defined the prostate barycenter as the origin for each CT scan in order to derive a population-based model of the CTV for prostate cancer. However, we believe that using the barycenter as a reference point is not feasible for all pelvic organs to obtain population-based models. For this reason, Bondar et al. (2014) used a setup based on bone anatomy for each patient, to calculate intra-patient motion-deformation vectors and develop a population model of CTVs for rectum cancer, as did (Budiarto et al., 2011). Nevertheless, this method may not be the most suitable setup for analyzing in the same framework all the different sizes and shapes that an organ can manifest in patients (Rios et al., 2016). In our study, we did not use bone anatomy as it can significantly differ from one patient to another, affecting the rigid registration and distribution of the pelvic organs. On the other hand, the bladder is attached to the prostate across the urethra, and the surface in contact with the prostate is generally stable, or is at least considered the least mobile part, while the top and anterior part can expand. We thereby sought to align this fixed bladder section in order to model both bladder motion and deformation. We believe that the ideal setup could be to match CTs using the boundary between the bladder and prostate, yet that type of rigid-registration is not possible. For intra-patient analysis, we think that the prostate barycenter offers a better means to perform this alignment due to the prostate volume being near constant during treatment, despite its rigid displacements (Roeske et al., 1995; Melian et al., 1997) (see Fig. A.14a in appendix A).

For inter-patient analysis, we consider that differences in prostate size and shape in the population can be reduced by defining the template as a patient presenting a prostate with the lowest volume variation and the most similar mean volume to the population mean (see Fig. A.14a in appendix A). In our training database, this patient prostate volume was $42.14 \pm 3.33\text{cm}^3$, whereas the population volume was $42.65 \pm 17.43\text{cm}^3$. For example, Fig. A.14b shows 3D sagittal views of some segmented patient prostates that were plotted as outside the first and third quartiles of the boxplot of the average prostate volumes. In Fig. A.14a, it can be observed that the size and shape differences were not significant for patients 17, 7, and 20, particularly in the upper anterior part of the prostate (the surface in contact with the bladder). The highest dissimilarity was obtained for Patient 8, an outlier like Patient 17. We thus consider that size and shape differences among prostates in the population are not sufficiently greater than those observed in bladders to introduce disturbances, at least in the first eigenmodes. For example, it can be observed in Fig. 6 that the first three modes did not reveal geometric variations

in the lower posterior bladder section, *i.e.*, the region that presented the geometric variations in the bladder produced by the barycenter alignment (see Figs. A.13 and A.14). Also, Fig. A.13 demonstrates that for intra-patient analysis, the first mode was also dominated by the bladder volume, meaning that volume changes in the prostate do not affect at least the most important mode in the bladder. It is worth mentioning that we did not introduce any rotation, nor conduct affine transformation, as the bladder does not present rotations as the prostate does, nor do volume changes in the prostate correlate with volume changes in the bladder. Besides, if all the bladders in the population are transformed using affine or non-rigid registration, inter-patient geometric variability can disappear, which is the main advantage of using ME models, where individual effects are separated from the mean. We therefore considered that the prostate barycenter also provided a suitable reference point in the pelvic zone to model bladder motion and deformation for inter-patient analysis, despite the potential variations of prostate in size and shape across the population.

Similar to other SPHARM-based shape models, our study was significantly aided by encoding the information of the bladder shape into vectors of SPHARM coefficients with orders of 10^2 , instead of vectors of boundary points with orders of 10^3 obtained from a meshing process, as used in (Söhn et al., 2005; Budiarto et al., 2011). The SPHARM parameterization thus enabled us to reduce not only computation time but also numerical errors, by affording finite arithmetic operations, and data storage while preserving shape information. However, it is important to highlight that not all the bladder shapes were properly parameterized using the radial sampling in spherical coordinates. In our training database, some bladders with non-convex dilations in the inferior-ventral region resulted in higher reconstruction errors. Future work should thus be dedicated to improve bladder-shape modeling. One way would be to extend the SPHARM methodology to model a much larger class of simply-connected closed shapes using an area-preserving, distortion minimizing spherical mapping, as expressed in (Brechtbühler et al., 1995; Shen et al., 2009a). Alternatively, a second possibility is to use a different computational technique called isometric analysis, which generalizes and improves on the standard finite element method (Zhang et al., 2007b; Nguyen et al., 2015). In population data, this technique offers the advantage of obtaining a surface parameterization with corresponding points in all organ instances without requiring a mesh relaxation procedure. This methodology defines a transformation to a canonical domain, where parametrization is carried out and extrapolated to any organ instance.

When considering population-based models developed for pelvic organs, we

found that 15 modes were required to describe 90% of CTV motion and deformation for prostate cancer (Budiarto et al., 2011). For rectum cancer, 28 modes were found to be required to describe 90% of the accumulated variability in CTVs (Bondar et al., 2014). In contrast, eight modes were sufficient in our study to describe the same percentage of accumulated variability. On analysis of the three methods' results, we noted that there were less fundamental directions of geometric variability in the bladder compared to other pelvic organs, despite its great volume and shape variations. Nevertheless, we were interested to note that more directions were observed in the prostate/SV composition, potentially due to changes in the bladder and rectum volumes, or due to patient-specific variations and delineation errors. Regarding the rectum, its greater number of directions may result from its flexible and vertical structure, enabling it to deform in any direction along its major axis (Hoogeman et al., 2002), in addition to patient-specific variations and delineation errors. Future studies using other population databases would be required to confirm or contradict the number of principal directions of motion and deformation reported in these organs.

On the other hand, as in (Budiarto et al., 2011), the reconstruction error in the PCA model and left-out patients in our study tended to decrease towards 0 as the number of modes was increased. However, the error did not completely disappear in the left-out patients, meaning that not all individual directions were accounted for by the PCA modes. In addition, as observed in Fig. 5, the rate in the left-out patient curve decreased after the first eight modes, implying that the order of the modes was not preserved in all patients. As a result, both of these issues may significantly affect the performance of any population PCA model. For example, the lowest measures of similarity were observed in Patients 6, 12, and 20 in the training database, whose bladders exhibited geometric variations in the superior-ventral region (see Patient 6's PM in Fig. 7). These variations were not represented by at least the first four modes (81% of the accumulated variability). It would therefore be interesting for future studies to first make a classification of bladders with similar shapes (Rios et al., 2016), then design the proposed model locally for each subgroup. For example, Bishop and Tipping (1998) introduced a methodology for developing hierarchical PCA models where each observation is plotted in the latent space to determine the number of sub-PCA models in the inferior levels.

It is also crucial to emphasize that we applied a quadratic relative error in the SPHARM space to assess the modes in our study, unlike Budiarto et al. (2011), who used a quadratic error normalized by the number of points. Nevertheless, we considered it possible to compare the behavior of both errors. For example,

(Budiarto et al., 2011) required 37 modes to reduce the initial error by 62% in the left-out patients, whereas 60 modes were needed in our study to reduce it by 80%. Both results may indicate there was significant intra-patient variability that was not sufficiently detected by the first modes obtained from the population. This also means that intra-patient variability may be higher in the bladder, potentially due to anatomical differences among patients and delineation errors. In this way, a trade-off was made between accumulated variability and the reconstruction error for determining the number of modes (40) in our PCA model, where an additional 20 modes only reduced the reconstruction error by 3%.

As established in the model assumptions, each mode corresponded to a direction of geometric variability that dilated or contracted certain regions of the bladder. Unlike the first mode, not all the rest were found to define an affine transformation that may be initially applied in the setup, such as rotation along an axis or a uniform dilation/contraction. In addition, we showed that the bladder motion and deformation between fractions follows random trajectory, as the data projections showed in the latent space (see Fig. 7). It was also possible, in this space, to characterize the geometric variations of each patient, where some directions were clearly found to be more dominant than others.

Although several studies have previously addressed the estimation of motion/deformation regions of the pelvic organs by means of the coverage matrix concept (Stroom et al., 1999; Mageras et al., 1999; Craig et al., 2001; Price and Moore, 2007), this is the first, to our knowledge, to introduce a quantitative and qualitative comparison between estimated and observed regions using metrics of similarity and joint histograms. As observed in Figs. 10, 11, and 12, our proposed model obtained the lowest measures of dissimilarity and misestimated voxels. Previous population PCA models have thus overestimated or underestimated more voxels, primarily those not occupied by the bladder. As a result, our proposed model reduced uncertainties in estimations of the probable region of motion and deformation. This was achievable as the ME models enable us to reduce the observed population variance along each mode by grouping repeated observations per patient. For example, we obtained more patients presenting significant differences between the Global PCA model and our proposed model in both databases than the Local PCA model; moreover, we significantly reduced uncertainties in estimations of motion/deformation region regarding the Global PCA model (see Table 4). We therefore concluded that the Global, Local, and Local PCA-ME models provided progressive decrease in uncertainties for estimating likely motion/deformation regions, while our proposed model led to the lowest reduction in all the patients (see Fig. 8 and 9). When more CTs/CBCTs scans become

available for a patient during treatment, this information could be used to improve estimation of the motion/deformation region by personalizing modes, variance, and mean bladder shape. It is worth mentioning that one limitation to use PCA models for modeling organ motion and deformation is their strong dependence on the organ's mean shape, as planning CT scans may not provide a good estimation. Future studies should be focused on analyzing the methodology's robustness for handling outlier patients whose variance along a mode is almost similar to the population variance. This kind of patient may lead to a misestimation of the intra-patient variability σ^2 affecting the performance of the ME model (see Patient 6 in the training database in Fig. 7). A robust formulation of PCA, the so-called robust PCA, could help overcome the problems associated with outlier patients.

PCA-based shape models are typically efficient for organs whose shape variations can be captured by a reasonable number of modes (Heimann and Meinzer, 2009). For instance, two studies have shown that 15 and 16 modes were required to describe most of the inter- and intra-patient motion and deformation of CTVs in prostate and rectum cancer, respectively (Budiarto et al., 2011; Bondar et al., 2014). However, in population databases concerning highly-varying soft-tissue volumes, a large number of modes may be required to properly capture the complex shape variations observed in the training database. In our study, for instance, we required 40 modes to generate an inter-patient PCA model for the bladder. We believe that the large number of modes may render interpretation of higher modes difficult, as they may not describe directions of geometric variability observed in the population rather than patient-specific directions. In addition, the large number of modes may also hinder inference and analysis of the results for each ME model. However, as in Aston et al. (2010), the modes define an orthogonal basis that enables us to consider each score z_{ijk} as an independent random variable across k that can be easily modeled as a response variable (Hotelling, 1933; Jolliffe, 2002).

We have previously mentioned that the scores can be interpreted as measures of geometric variability along the modes. Future studies should therefore focus on determining covariates that may help to describe patient-specific bladder deformation. The results displayed in Fig. 7 and published by (Rios et al., 2016) appear to indicate that there are some covariates that may underlie the motion/deformation region in the latent space or mode's variance for any arbitrary patient. For example, bladder size or volume could be considered as fixed parameters that may determine changes in volume and directions of geometric variations. (Rios et al., 2016) and (Palorini et al., 2016a) reported, in line with Fig. 7, that bladders with large volumes exhibited geometric variations in the superior-ventral region and

higher variance along the first mode. In an adaptive radiotherapy scheme, a time covariate can also be added, as a random or fixed variable, to describe the daily volume changes of the bladder.

As previously stated, Hu et al. (2015) proposed a population-PCA model to describe prostate deformation in MR-tumor-targeted biopsies. Concerning this methodology, we found two shared characteristics: firstly, that the scores were also considered as parameters that describe intra-subject organ motion; secondly, the added objective of characterizing subject-specific probability density functions (SSPDFs) of motion and deformation (parameterized by means of multivariate Gaussian distributions). However, Hu et al. (2015) proposed a fixed-effects model to describe SSPDF of the scores across the population, while such a model consists of a multivariate non-linear regression involving the scores of the reference shapes defined as covariates. Hu et al. (2015) also decided to model the SSPDF parameters instead of the subject-specific scores, as proposed in our study. Nevertheless, we believe that it could be interesting for future studies to add the scores of the bladder segmented at the planning CT as fixed variables, as this may help to correlate the subject-specific variance and directions of geometric variability. For future works, we wish to underline that both studies have suggested the possibility of including other parameters in the learning framework, such as organ size or any temporal information.

For future developments, it is worth noting that our main goal has been to predict late GU toxicity based only on planning treatment information, and we thus only used the planning CT scan to predict bladder motion/deformation regions. Nonetheless, we have also performed some simulations using two and three observations of the out-of-sample patient, with the aim of simulating an adaptive radiotherapy treatment, and the results achieved better estimation of the region. We think that the number of observations for the out-of-sample patient helps to decrease the variance of the estimate for the shrinkage predictor. Thus, the more observations we have for the out-of-sample patient, the more certain we can be about how the patient differs from the population mean. In these simulations, it is also worth mentioning that we only used these observations to improve the estimation of the patient's average bladder and shrinkage estimator of each mode, *i.e.* to improve the model adaptation yet not retrain the model's parameters (modes φ_k , and inter- intra-patient variances). We considered that more observations should be available to fit the model's parameters again, as five CTs/CBCTs has been established as the minimum number of observations required for intra-patient models (Söhn et al., 2005).

In this study, we described regions of motion/deformation in terms of 3D pro-

bability maps (PM) that were obtained by sampling the model's distribution or using patient images. We have thus considered motion/deformation regions as histograms that can be compared using the proposed metrics. We believe that the next step should be focused on estimating the uncertainty of the motion/deformation region using the model's joint distribution. However, we considered that the large number of modes may hinder inference of the associated joint distribution, as the number of combinations of scores at the boundaries of the confidence intervals will significantly grow when we add modes to the joint distribution. Nonetheless, we consider that the uncertainty region can also be estimated using only those modes that have a strong influence on variables, like the accumulated delivered dose, *i.e.*, that the uncertainty region can be restricted to those modes that have a strong influence on the mean accumulated dose. Finally, it may also be interesting to add linearly modes that describe geometric variations between the mean bladders of the patients as the modes obtained from the covariance matrix $\mathbf{C}_{empirical}$ only described directions observed on the intra-patient level. These models are known as multi-level PCA models (Di et al., 2009; Greven et al., 2010).

5. Conclusion

We proposed a population based model to predict bladder motion and deformation between fractions using solely the planning CT. In comparison with previous studies, our proposed model was able to decrease uncertainty in the estimation of the region where the bladder will likely move and deform. We also demonstrated that, by following a longitudinal study with ME models, it was possible to separate the patient-specific variance from population variance and thus reduce the total variance. The potential applications of this model include margin evaluation, delivered dose estimation, toxicity prediction, and the design of robust treatment plans, among others. Future studies should now be conducted to validate the proposed model by means of a large cohort of patients who have undergone prostate cancer radiotherapy. It could also be valuable to focus on the relationship between inter-fraction bladder motion/deformation, delivered dose, and late GU toxicity in the prostate cancer radiotherapy context.

6. Acknowledgements

This research was supported by the Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS) – Colombia with the “Becas de Doctorado en Colombia 2010 – Convocatoria 511” grant, and also by the Institut National du Cancer (INCa) – France with the “STIC IGRT-P” project.

Appendix A. Supplementary material

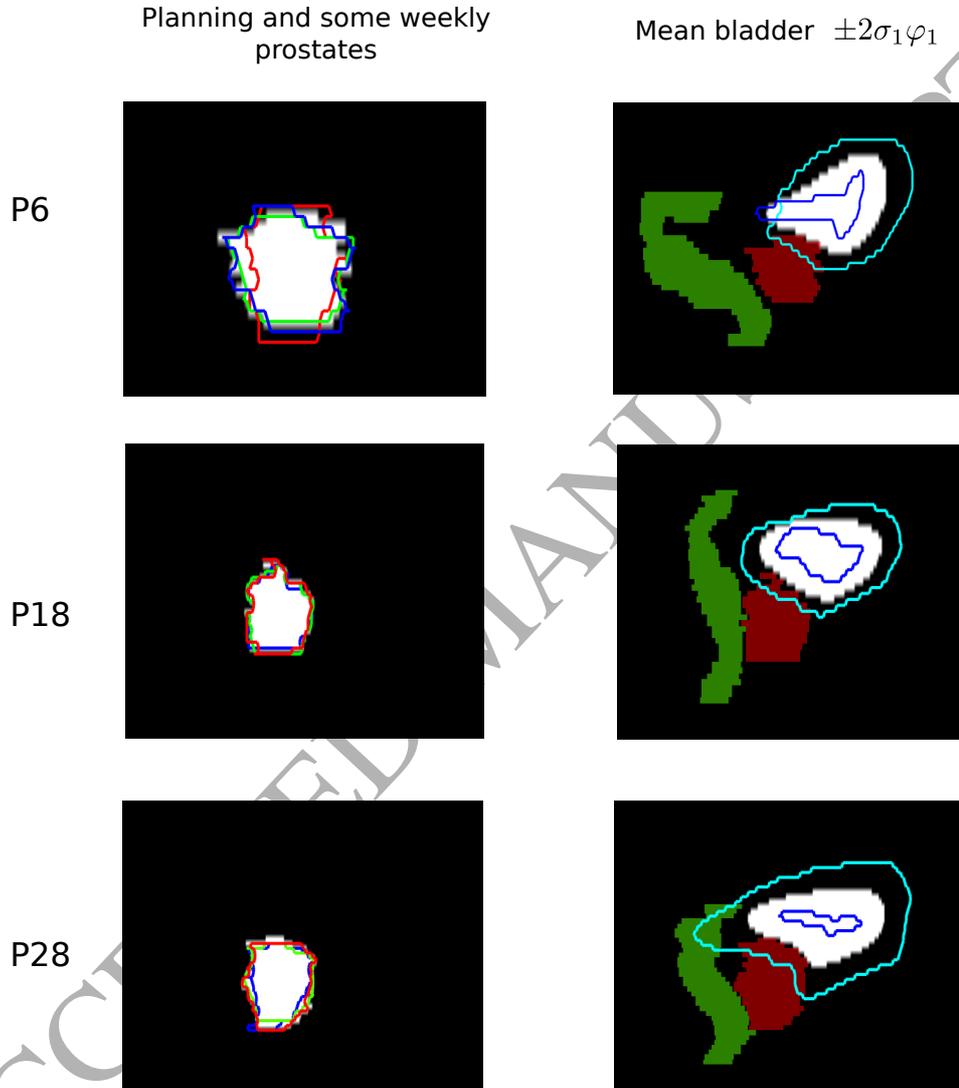
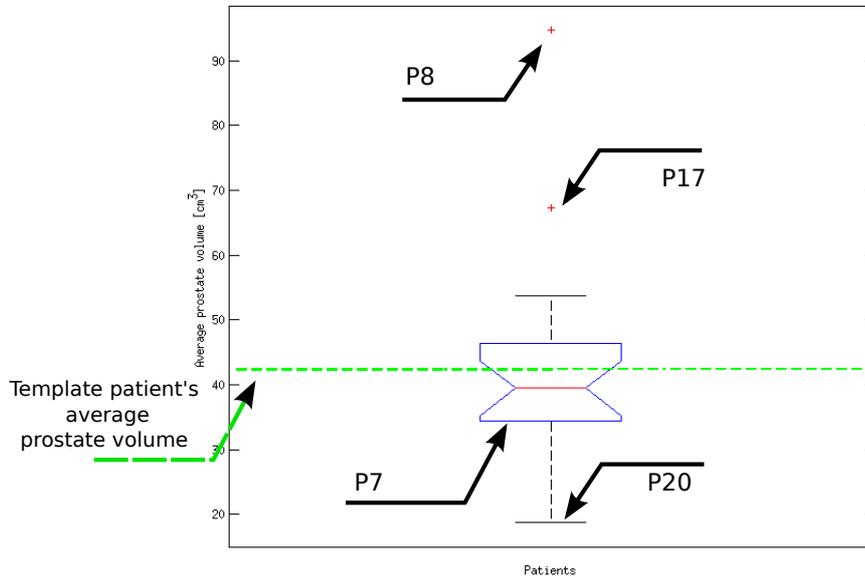
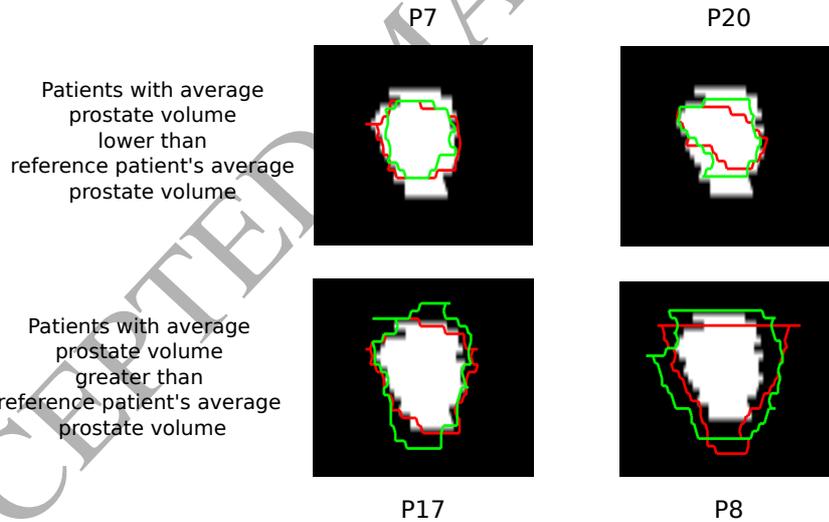


Figure A.13: Intra-patient analysis for patients with on-treatment CBCTs. Left column depicts 3D sagittal views of the segmented prostate (white) at the planning CT with two observations (color contours) during treatment. Right columns shows 3D sagittal views of geometric variations (blue and cyan) along the first mode applied on the mean bladder (white), where σ_1^2 is the mode variance. The prostate (red) and rectum (green) of the patient were also included as spatial reference.



a) Average prostate volume in the training database



b) Template prostate (white) vs observed prostates of patients outside from first and third quartiles

Figure A.14: (a) Boxplot of average prostate volume in the training database. Green line shows the average volume of the template prostate. (b) Overlays of some observed prostates (green and red) of patients outside from first and third quartiles vs template prostate (white).

References

- Aston, J. A. D., Chiou, J.-M., Evans, J. P., Mar. 2010. Linguistic pitch analysis using functional principal component mixed effect models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 59 (2), 297–317.
- Badawi, A. M., Weiss, E., Sleeman, W. C., Yan, C., Hugo, G. D., Sep. 2010. Optimizing principal component models for representing interfraction variation in lung cancer radiotherapy. *Med Phys* 37 (9), 5080–5091.
- Ballard, D. H., Brown, C. M., May 1982. *Computer Vision*, first edition Edition. Prentice Hall, Englewood Cliffs, N.J.
- Balter, J. M., Sandler, H. M., Lam, K., Bree, R. L., Lichter, A. S., Ten Haken, R. K., Jan. 1995. Measurement of prostate movement over the course of routine radiotherapy using implanted markers. *Int J Radiat Oncol Biol Phys* 31 (1), 113–118.
- Bastogne, T., Samson, A., Vallois, P., Wantz-Mzires, S., Pinel, S., Bechet, D., Barberi-Heyob, M., 2010. Phenomenological modeling of tumor diameter growth based on a mixed effects model. *Journal of Theoretical Biology* 262 (3), 544–552.
- Bernal-Rusiel, J. L., Reuter, M., Greve, D. N., Fischl, B., Sabuncu, M. R., November 2013. Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. *Neuroimage* 81, 358–370.
- Bishop, C., Tipping, M., Mar. 1998. A hierarchical latent variable model for data visualization. *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (3), 281–293.
- Bondar, L., Intven, M., Burbach, J. P. M., Budiarto, E., Kleijnen, J.-P., Philippens, M., van Asselen, B., Seravalli, E., Reerink, O., Raaymakers, B., November 2014. Statistical modeling of CTV motion and deformation for IMRT of early-stage rectal cancer. *Int J Radiat Oncol Biol Phys* 90 (3), 664–672.
- Brechbühler, C., Gerig, G., Kübler, O., Mar. 1995. Parametrization of Closed Surfaces for 3-D Shape Description. *Computer Vision and Image Understanding* 61 (2), 154–170.
- Budiarto, E., Keijzer, M., Storchi, P. R., Hoogeman, M. S., Bondar, L., Mutanga, T. F., de Boer, H. C. J., Heemink, A. W., Feb. 2011. A population-based model

- to describe geometrical uncertainties in radiotherapy: applied to prostate cases. *Phys. Med. Biol.* 56 (4), 1045–1061.
- Cootes, T. F., Taylor, C. J., Cooper, D. H., Graham, J., Jan. 1995. Active Shape Models-Their Training and Application. *Computer Vision and Image Understanding* 61 (1), 38–59.
- Craig, T., Battista, J., Moiseenko, V., Van Dyk, J., Jan. 2001. Considerations for the implementation of target volume protocols in radiation therapy. *Int J Radiat Oncol Biol Phys* 49 (1), 241–250.
- Crook, J. M., Raymond, Y., Salhani, D., Yang, H., Esche, B., Oct. 1995. Prostate motion during standard radiotherapy as assessed by fiducial markers. *Radiother. Oncol.* 37 (1), 35–42.
- Davatzikos, C., Shen, D., Mohamed, A., Kyriacou, S. K., Aug. 2001. A framework for predictive modeling of anatomical deformations. *IEEE Transactions on Medical Imaging* 20 (8), 836–843.
- Di, C.-Z., Crainiceanu, C. M., Caffo, B. S., Punjabi, N. M., 2009. Multilevel functional principal component analysis. *The Annals of Applied Statistics* 3 (1), 458–488.
- Dillenseger, J.-L., Guillaume, H., Patard, J.-J., Nov. 2006. Spherical harmonics based intrasubject 3-D kidney modeling/registration technique applied on partial information. *IEEE Trans Biomed Eng* 53 (11), 2185–2193.
- Donald Hedeker, R. D. G., 2006. *Longitudinal Data Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey, USA.
- Durrleman, S., Pennec, X., Trouvé, A., Braga, J., Gerig, G., Ayache, N., 2013a. Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. *International Journal of Computer Vision* 103 (1), 22–59.
- Durrleman, S., Pennec, X., Trouv, A., Braga, J., Gerig, G., Ayache, N., May 2013b. Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. *Int J Comput Vis* 103 (1), 22–59.
- Fiorino, C., Valdagni, R., Rancati, T., Sanguineti, G., Nov. 2009. Dose volume effects for normal tissues in external radiotherapy: Pelvis. *Radiother. Oncol.* 93 (2), 153–167.

- Fitzmaurice, G., Laird, N., Ware, J., 2011. *Applied Longitudinal Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, New Jersey, USA.
- Fontenla, E., Pelizzari, C. A., Roeske, J. C., Chen, G. T. Y., Sep. 2001a. Numerical analysis of a model of organ motion using serial imaging measurements from prostate radiotherapy. *Phys. Med. Biol.* 46 (9), 2337.
- Fontenla, E., Pelizzari, C. A., Roeske, J. C., Chen, G. T. Y., Sep. 2001b. Using serial imaging data to model variabilities in organ position and shape during radiotherapy. *Phys. Med. Biol.* 46 (9), 2317–36.
- Fonteyne, V., Villeirs, G., Speleers, B., De Neve, W., De Wagter, C., Lumen, N., De Meerleer, G., 2008. Intensity-modulated radiotherapy as primary therapy for prostate cancer: Report on acute toxicity after dose escalation with simultaneous integrated boost to intraprostatic lesion. *Int. J. Radiat. Oncol. Biol. Phys.* 72 (3), 799–807.
- Frees, E. W., 2004. *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*, 1st Edition. Cambridge University Press, New York.
- Gerig, G., Styner, M., Jones, D., Weinberger, D., Lieberman, J., 2001. Shape analysis of brain ventricles using SPHARM. In: *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis- MMBIA*. pp. 171–178.
- Greven, S., Crainiceanu, C., Caffo, B., Reich, D., 2010. Longitudinal functional principal component analysis. *Electron J Stat* 4, 1022–1054.
- Guo, W., 2002. Functional Mixed Effects Models. *Biometrics* 58 (1), 121–128.
- Hanlon, A. L., Moore, D. F., Hanks, G. E., 1998. Modeling postradiation prostate specific antigen level kinetics. *Cancer* 83 (1), 130–134.
- He, T., Xue, Z., Xie, W., Wong, S. T. C., 2010. Online 4-D CT estimation for patient-specific respiratory motion based on real-time breathing signals. *Med Image Comput Comput Assist Interv* 13 (Pt 3), 392–399.
- Heimann, T., Meinzer, H.-P., Aug. 2009. Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis* 13 (4), 543–563.

- Hoogeman, M. S., van Herk, M., Yan, D., Boersma, L. J., Koper, P. C. M., Lebesque, J. V., Oct. 2002. A model to simulate day-to-day variations in rectum shape. *Int J Radiat Oncol Biol Phys* 54 (2), 615–625.
- Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24 (6), 417–441.
- Hu, Y., Gibson, E., Ahmed, H. U., Moore, C. M., Emberton, M., Barratt, D. C., Dec. 2015. Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration. *Med. Image Anal.* 26 (1), 332–344.
- Jolliffe, I. T., Oct. 2002. *Principal Component Analysis*. 2nd ed, Springer, NY,USA.
- Kelemen, A., Szekely, G., Gerig, G., Oct. 1999. Elastic model-based segmentation of 3-D neuroradiological data sets. *IEEE Transactions on Medical Imaging* 18 (10), 828–839.
- Killoran, J. H., Kooy, H. M., Gladstone, D. J., Welte, F. J., Beard, C. J., Jan. 1997. A numerical simulation of organ motion and daily setup uncertainties: Implications for radiation therapy. *Int J Radiat Oncol Biol Phys* 37 (1), 213–221.
- Kraskov, A., Grassberger, P., 2009. MIC: Mutual information based hierarchical clustering. In: Emmert-Streib, F., Dehmer, M. (Eds.), *Information Theory and Statistical Learning*. Springer US, pp. 101–123.
- Laird, N. M., Ware, J. H., Dec. 1982. Random-effects models for longitudinal data. *Biometrics* 38 (4), 963–974.
- Landoni, V., Fiorino, C., Cozzarini, C., Sanguineti, G., Valdagni, R., Rancati, T., Mar. 2016. Predicting toxicity in radiotherapy for prostate cancer. *Physica Medica* 32 (3), 521–532.
- Liang, K.-Y., Zeger, S. L., Apr. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73 (1), 13–22.
- Liu, X., 2016. Chapter 3 - linear mixed-effects models. In: Liu, X. (Ed.), *Methods and Applications of Longitudinal Data Analysis*. Academic Press, Oxford, pp. 61 – 94.

- Lorenz, C., Krahnstöver, N., 2000. Generation of point-based 3d statistical shape models for anatomical objects. *Computer Vision and Image Understanding* 77 (2), 175–191.
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 16 (2), 187–198.
- Mageras, G. S., Fuks, Z., Leibel, S. A., Ling, C. C., Zelefsky, M. J., Kooy, H. M., van Herk, M., Kutcher, G. J., Jan. 1999. Computerized design of target margins for treatment uncertainties in conformal radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 43 (2), 437–445.
- Mageras, G. S., Kutcher, G. J., Leibel, S. A., Zelefsky, M. J., Melian, E., Mohan, R., Fuks, Z., May 1996. A method of incorporating organ motion uncertainties into three-dimensional conformal treatment plans. *Int. J. Radiat. Oncol. Biol. Phys.* 35 (2), 333–342.
- McClelland, J. R., Hawkes, D. J., Schaeffter, T., King, A. P., Jan. 2013. Respiratory motion models: A review. *Med. Image Anal.* 17 (1), 19–42.
- Melian, E., Mageras, G. S., Fuks, Z., Leibel, S. A., Niehaus, A., Lorant, H., Zelefsky, M., Baldwin, B., Kutcher, G. J., 1997. Variation in prostate position quantitation and implications for three-dimensional conformal treatment planning. *Int J Radiat Oncol Biol Phys* 38 (1), 73–81.
- Morris, J. S., Carroll, R. J., Apr. 2006. Wavelet-based functional mixed models. *J R Stat Soc Series B Stat Methodol* 68 (2), 179–199.
- Nassef, M., Simon, A., Cazoulat, G., Dumnil, A., Blay, C., Lafond, C., Acosta, O., Balosso, J., Haignon, P., de Crevoisier, R., Apr. 2016. Quantification of dose uncertainties in cumulated dose estimation compared to planned dose in prostate IMRT. *Radiotherapy and Oncology* 119 (1), 129–136.
- Nguyen, V. P., Anitescu, C., Bordas, S. P. A., Rabczuk, T., November 2015. Iso-geometric analysis: An overview and computer implementation aspects. *Mathematics and Computers in Simulation* 117, 89–116.
- Palorini, F., Botti, A., Carillo, V., Gianolini, S., Improta, I., Iotti, C., Rancati, T., Cozzarini, C., Fiorino, C., Mar. 2016a. Bladder dosesurface maps and urinary toxicity: Robustness with respect to motion in assessing local dose effects. *Physica Medica* 32 (3), 506–511.

- Palorini, F., Cozzarini, C., Gianolini, S., Botti, A., Carillo, V., Iotti, C., Rancati, T., Valdagni, R., Fiorino, C., Apr. 2016b. First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy. *Radiother Oncol* 119 (1), 123–128.
- Pavel-Mititean, L. M., Rowbottom, C. G., Hector, C. L., Partridge, M., Bortfeld, T., Schlegel, W., Jun. 2004. A geometric model for evaluating the effects of inter-fraction rectal motion during prostate radiotherapy. *Phys. Med. Biol.* 49 (12), 2613–29.
- Pekar, V., McNutt, T. R., Kaus, M. R., Nov. 2004. Automated model-based organ delineation for radiotherapy planning in prostatic region. *Int. J. Radiat. Oncol. Biol. Phys.* 60 (3), 973–980.
- Pinheiro, J., Bates, D., 2000. *Mixed-Effects Models in S and S-PLUS*, 1st Edition. Springer, New York.
- Pinheiro, J. C., Bates, D. M., 1996. Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing* 6 (3), 289–296.
- Prabhu, R. S., Won, M., Shaw, E. G., Hu, C., Brachman, D. G., Buckner, J. C., Stelzer, K. J., Barger, G. R., Brown, P. D., Gilbert, M. R., Mehta, M. P., 2014. Effect of the addition of chemotherapy to radiotherapy on cognitive function in patients with low-grade glioma: Secondary analysis of RTOG 98-02. *Journal of Clinical Oncology* 32 (6), 535–541.
- Preiswerk, F., De Luca, V., Arnold, P., Celicanin, Z., Petrusca, L., Tanner, C., Bieri, O., Salomir, R., Cattin, P. C., Jul. 2014. Model-guided respiratory organ motion prediction of the liver from 2d ultrasound. *Med. Image Anal.* 18 (5), 740–751.
- Price, G. J., Moore, C. J., Apr. 2007. A method to calculate coverage probability from uncertainties in radiotherapy via a statistical shape model. *Phys Med Biol* 52 (7), 1947–1965.
- Rios, R., Ospina, J. D., Lafond, C., Acosta, O., Espinosa, J., de Crevoisier, R., 2016. Characterization of bladder motion and deformation in prostate cancer radiotherapy. *IRBM Journal*. <http://dx.doi.org/10.1016/j.irbm.2016.09.001>.
- Roeske, J. C., Forman, J. D., Mesina, C. F., He, T., Pelizzari, C. A., Fontenla, E., Vijayakumar, S., Chen, G. T. Y., 1995. Evaluation of changes in the size and

- location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy. *Int J Radiat Oncol Biol Phys* 33 (5), 1321–1329.
- Rosewall, T., Catton, C., Currie, G., Bayley, A., Chung, P., Wheat, J., Milosevic, M., 2010. The relationship between external beam radiotherapy dose and chronic urinary dysfunction – a methodological critique. *Radiother. Oncol.* 97 (1), 40–47.
- Russakoff, D. B., Tomasi, C., Rohlfing, T., Jr, C. R. M., 2004. Image similarity using mutual information of regions. In: *Computer Vision - ECCV 2004*. Vol. 3023 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, pp. 596–607.
- Shen, L., Farid, H., McPeck, M. A., Apr. 2009a. Modeling three-dimensional morphological structures using spherical harmonics. *Evolution* 63 (4), 1003–1016.
- Shen, L., Ford, J., Makedon, F., Saykin, A., 2004. A surface-based approach for classification of 3d neuroanatomical structures. *Intelligent Data Analysis* 8 (6), pp.519–542.
- Shen, L., Kim, S., Saykin, A. J., Jun. 2009b. Fourier method for large scale surface modeling and registration. *Comput Graph* 33 (3), 299–311.
- Söhn, M., Birkner, M., Yan, D., Alber, M., Dec. 2005. Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation. *Phys. Med. Biol.* 50 (24), 5893–5908.
- Söhn, M., Sobotta, B., Alber, M., Jun. 2012. Dosimetric treatment course simulation based on a statistical model of deformable organ motion. *Phys. Med. Biol.* 57 (12), 3693–3709.
- Stroom, J. C., de Boer, H. C. J., Huizenga, H., Visser, A. G., Mar. 1999. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *Int J Radiat Oncol Biol Phys* 43 (4), 905–919.
- Styner, M., Lieberman, J. A., Pantazis, D., Gerig, G., Sep. 2004. Boundary and medial shape analysis of the hippocampus in schizophrenia. *Medical Image Analysis* 8 (3), 197–203.

Styner, M., Oguz, I., Xu, S., Brechbühler, C., Pantazis, D., Levitt, J. J., Shenton, M. E., Gerig, G., 2006. Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM. *Insight J* (1071), 242–250.

Székel, G., Kelemen, A., Brechbühler, C., Gerig, G., Mar. 1996. Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *Medical Image Analysis* 1 (1), 19–34.

Tateyama, T., Okegawa, M., Uetani, M., Tanaka, H., Kohara, S., Han, X., Kanasaki, S., Sato, S., Wakamiya, M., Furukawa, A., Jiang, H., Chen, Y. W., Nov. 2012. Efficient shape representation and statistical shape modeling of the liver using spherical harmonic functions (SPHARM). In: 2012 Joint 6th International Conference on Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS). pp. 428–431.

Ten Haken, R. K., Forman, J. D., Heimburger, D. K., Gerhardsson, A., McShan, D. L., Perez-Tamayo, C., Schoepfel, S. L., Lichter, A. S., Jun. 1991. Treatment planning issues related to prostate movement in response to differential filling of the rectum and bladder. *Int. J. Radiat. Oncol. Biol. Phys.* 20 (6), 1317–1324.

Thor, M., Apte, A., Deasy, J. O., Muren, L. P., Apr. 2013a. Statistical simulations to estimate motion-inclusive dose-volume histograms for prediction of rectal morbidity following radiotherapy. *Acta Oncol.* 52 (3), 666–675.

Thor, M., Bentzen, L., Hysing, L. B., Ekanger, C., Helle, S.-I., Karlsdóttir, A., Muren, L. P., May 2013b. Prediction of rectum and bladder morbidity following radiotherapy of prostate cancer based on motion-inclusive dose distributions. *Radiother. Oncol.* 107 (2), 147–152.

Thor, M., Vaeth, M., Karlsdóttir, A., Muren, L. P., Oct. 2010. Rectum motion and morbidity prediction: Improving correlation between late morbidity and DVH parameters through use of rectum planning organ at risk volumes. *Acta Oncol* 49 (7), 1061–1068.

Thörnqvist, S., Hysing, L. B., Zolnay, A. G., Shn, M., Hoogeman, M. S., Muren, L. P., Bentzen, L., Heijmen, B. J. M., Dec. 2013. Treatment simulations with a statistical deformable motion model to evaluate margins for multiple targets in radiotherapy for high-risk prostate cancer. *Radiother Oncol* 109 (3), 344–349.

Turner, S. L., Swindell, R., Bowl, N., Marrs, J., Brookes, B., Read, G., Cowan, R. A., 1997. Bladder movement during radiation therapy for bladder cancer: implications for treatment planning. *Int J Radiat Oncol Biol Phys* 39 (2), 355–360.

van Herk, M., Bruce, A., Guus Kroes, A. P., Shouman, T., Touw, A., Lebesque, J. V., Dec. 1995. Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. *Int J Radiat Oncol Biol Phys* 33 (5), 1311–1320.

Viswanathan, A. N., Yorke, E. D., Marks, L. B., Eifel, P. J., Shipley, W. U., 2010. Radiation dose-volume effects of the urinary bladder. *Int. J. Radiat. Oncol. Biol. Phys.* 76 (3), S116–122.

Wu, H., Zhang, J.-T., 2002. Local Polynomial Mixed-Effects Models for Longitudinal Data. *Journal of the American Statistical Association* 97 (459), 883–897.

Yan, D., Jaffray, D., Wong, J., Jun. 1999. A model to accumulate fractionated dose in a deforming organ. *Int J Radiat Oncol Biol Phys* 44 (3), 665–675.

Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y.-C., Rosenzweig, K. E., Ling, C. C., Mageras, G. S., Dec. 2007a. A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med Phys* 34 (12), 4772–4781.

Zhang, Y., Bazilevs, Y., Goswami, S., Bajaj, C. L., Hughes, T. J. R., May 2007b. Patient-specific vascular NURBS modeling for isogeometric analysis of blood flow. *Computer Methods in Applied Mechanics and Engineering* 196 (29), 2943–2959.

Ziegler, G., Penny, W. D., Ridgway, G. R., Ourselin, S., Friston, K. J., Nov. 2015. Estimating anatomical trajectories with Bayesian mixed-effects modeling. *NeuroImage* 121, 51–68.

Zietman, A. L., Bae, K., Slater, J. D., Shipley, W. U., Efstathiou, J. A., Coen, J. J., Bush, D. A., Lunt, M., Spiegel, D. Y., Skowronski, R., Jabola, B. R., Rossi, C. J., 2010. Randomized trial comparing conventional-dose with high-dose conformal radiation therapy in early-stage adenocarcinoma of the prostate: long-term results from proton radiation oncology group/american college of radiology 95-09. *Journal of Clinical Oncology* 28 (7), 1106–1111.