STREAMLINE.
OPTIMIZE.
TRUST.

ADAPTIVO | LINACVIEW | DOSEVIEW | PIPSPRO | QA PILOT | IMSURE | STRUCTSURE AI QA

COMPLETE INTEGRATED QA

STANDARD IMAGING®

CLICK HERE TO LEARN MORE

Received: 17 June 2021 | Revised: 24 August 2021 | Accepted: 16 September 2021

DOI: 10.1002/mp.15262

RESEARCH ARTICLE

MEDICAL PHYSICS

Reducing systematic errors due to deformation of organs at risk in radiotherapy

Øyvind Lunde Rørtveit[1,2] | **Liv Bolstad Hysing**[1,2] | **Andreas Størksen Stordal**[3,4] | **Sara Pilskog**[1,2]

[1] Department of Oncology and Medical Physics, Haukeland University Hospital, Bergen, Norway

[2] Department of physics and technology, University of Bergen, Bergen, Norway

[3] NORCE Research, Bergen, Norway

[4] Department of Mathematics, University of Bergen, Bergen, Norway

Correspondence
Øyvind Lunde Rørtveit, Department of Oncology and Medical Physics, Haukeland University Hospital, P.O.B. 1400, 5021 Bergen, Norway.
Email: oyvind.lunde.rortveit@helse-bergen.no

Funding information
Trond Mohn stiftelse, Grant/Award Number: BFS2017TMT07

Abstract

Purpose: In radiotherapy (RT), the planning CT (pCT) is commonly used to plan the full RT-course. Due to organ deformation and motion, the organ shapes seen at the pCT will not be identical to their shapes during RT. Any difference between the pCT organ shape and the organ's mean shape during RT will cause systematic errors. We propose to use statistical shrinkage estimation to reduce this error using only the pCT and the population mean shape computed from training data.

Methods: The method was evaluated for the rectum in a cohort of 37 prostate cancer patients that had a pCT and 7–10 treatment CTs with rectum delineations. Deformable registration was performed both within-patient and between patients, resulting in point-to-point correspondence between all rectum shapes, which enabled us to compute a population mean rectum. Shrinkage estimates were found by combining the pCTs linearly with the population mean.

The method was trained and evaluated using leave-one-out cross validation. The shrinkage estimates and the patient mean shapes were compared geometrically using the Dice similarity index (DSI), Hausdorff distance (HD), and bidirectional local distance. Clinical dose/volume histograms, equivalent uniform dose (EUD) and minimum dose to the hottest 5% volume (D5%) were compared for the shrinkage estimate and the pCT.

Results: The method resulted in moderate but statistically significant increase in similarity to the patient mean shape over the pCT. On average, the HD was reduced from 15.6 to 13.4 mm, while the DSI was increased from 0.74 to 0.78. Significant reduction in the bias of volume estimates was found in the DVH-range of 52.5–65 Gy, where the bias was reduced from −1.3 to −0.2 percentage points, but no significant improvement was found in EUD or D5%,

Conclusions: The results suggest that shrinkage estimation can reduce systematic errors due to organ deformations in RT. The method has potential to increase the accuracy in RT of deformable organs and can improve motion modeling.

KEYWORDS
organ deformation, organ motion, rectum, systematic errors, treatment uncertainties

1 | INTRODUCTION

The planning CT (pCT) is used as a representation of the patient anatomy during treatment. Changes to the patient anatomy and its organ shapes during treatment cause differences between the planned and the delivered doses.[1] Discrepancies between the pCT anatomy and the average anatomy during treatment are called systematic errors, since they affect every fraction. While this term is often used to describe the discrepancy in setup position, systematic errors also occur when an organ of interest has a different shape in the pCT as compared to its average shape during treatment. This study focuses on this latter systematic shape error. Potential systematic errors in organs-at-risk (OARs) can be accounted for through robust optimization,[2] or with margins (planning OAR volumes, PRVs).[3] Unlike these methods, the presented method aims to predict and correct for the systematic error. Correcting for systematic setup error (i.e., without deformation) by using several scans, typically taken during the first few fractions of RT, has been investigated.[4] A similar method can be used to handle deformation, using deformable registration from multiple scans to find an average shape,[5] but this would require adaptive re-planning unless all scans are taken before the first treatment. The presented method, on the other hand, requires only a single scan. Image-guidance has been successful in reducing the systematic errors for the target volumes, but for shape changes and for many OARs, mitigation strategies based on rigid re-alignment are insufficient, calling for more resource demanding adaptive RT.[6,7] Even adaptive RT with replanning at every fraction is not a perfect remedy, since intra-fraction motion can be considerable.[8,9]

For the rectum, one of the dose-limiting OARs in both prostate and cervical cancer RT, the shape of the organ seen in the pCT directly influences the dose that can be safely administered to the tumor. Dose objectives used for planning come from dose-response models, and the majority of these models are based on the dose to the pCT-shape of the rectum.[10–12] Stronger response prediction has been achieved by accounting for rectal motion,[13–15] but estimates of the average delivered dose to the rectum are resource expensive, demanding both frequent images and complex software, limiting its use.

In this study, we aim to derive a model that enables estimation of a patient's mean rectal shape from the pCT scan only. To solve the obstacle of requiring multiple imaging input to assess the average shape, we apply the statistical method of shrinkage estimation combined with information from a deformable population model of the rectum.

2 | MATERIALS AND METHODS

In statistics, shrinkage estimation is a well-known technique used to reduce variance in estimates,[16] not to be confused with physical/anatomical shrinkage. The organ shape used in planning can be considered an estimate of the mean organ shape. Reducing the variance of this estimate means reducing the average difference between the planning shape and the average shape during treatment and therefore corresponds to reducing systematic shape errors. One important aspect of the proposed method is the calculation and use of a mean organ shape across the population – the population mean shape (PMS). Intuitively, the shrinkage method can be understood through the principle of "regression toward the mean": if a single sample taken from an individual is extreme, the next sample taken from the same individual is likely to be closer to the population mean. Our hypothesis is therefore that the true mean shape of an organ can be estimated with lower variance by combining information from the pCT shape (pS) and the PMS.

For computation of the PMS, we rely on deformable registration of organ contours. We used a variant of the thin plate spline - robust point matching (TPS-RPM) algorithm.[17,18] Each shape is represented by a set of points on the organ surface (a mesh representation). The algorithm finds a transform function (a TPS function in this case) which can be used to transform a reference structure into a shape as similar as possible to a target structure. Transforming the reference structure is commonly referred to as «warping». By replacing the target structure by the warped reference structure, we end up with a one-to-one correspondence of points in the two shapes.

To find point-to-point correspondence between the surface points on all rectum shapes in the dataset, we used the method first presented in Budiarto et al.[19] The method is illustrated in Figure 1. First, we performed registration from a reference patient's pCT rectum structure to each of the other patients' pCT rectum structures (inter-patient registration). Since our data set contains multiple scans for each patient, we then performed registration from the rectum structure in each pCT to each of the same patient's other rectum shapes (intra-patient registration). Since there is now a point-to-point correspondence between all shapes, we were able to average the coordinates of the points across all CTs to find a population mean rectum shape.

After the (preliminary) PMS had been built, the procedure was repeated, with the preliminary PMS taking the place of the global reference shape. This was done to increase the accuracy of the intra-patient registrations, and similar to the procedure used in Budiarto et al.[19] In practice, the deformable registration was split in two steps – finding the transform function and applying it (warping). In the intra-patient step, the transform function was found between the original pS (not the warped reference shape) and the treatment CT shapes. To find the final treatment CT shapes, the global reference was warped twice, first from its original shape into the pCT shape of the specific patient,

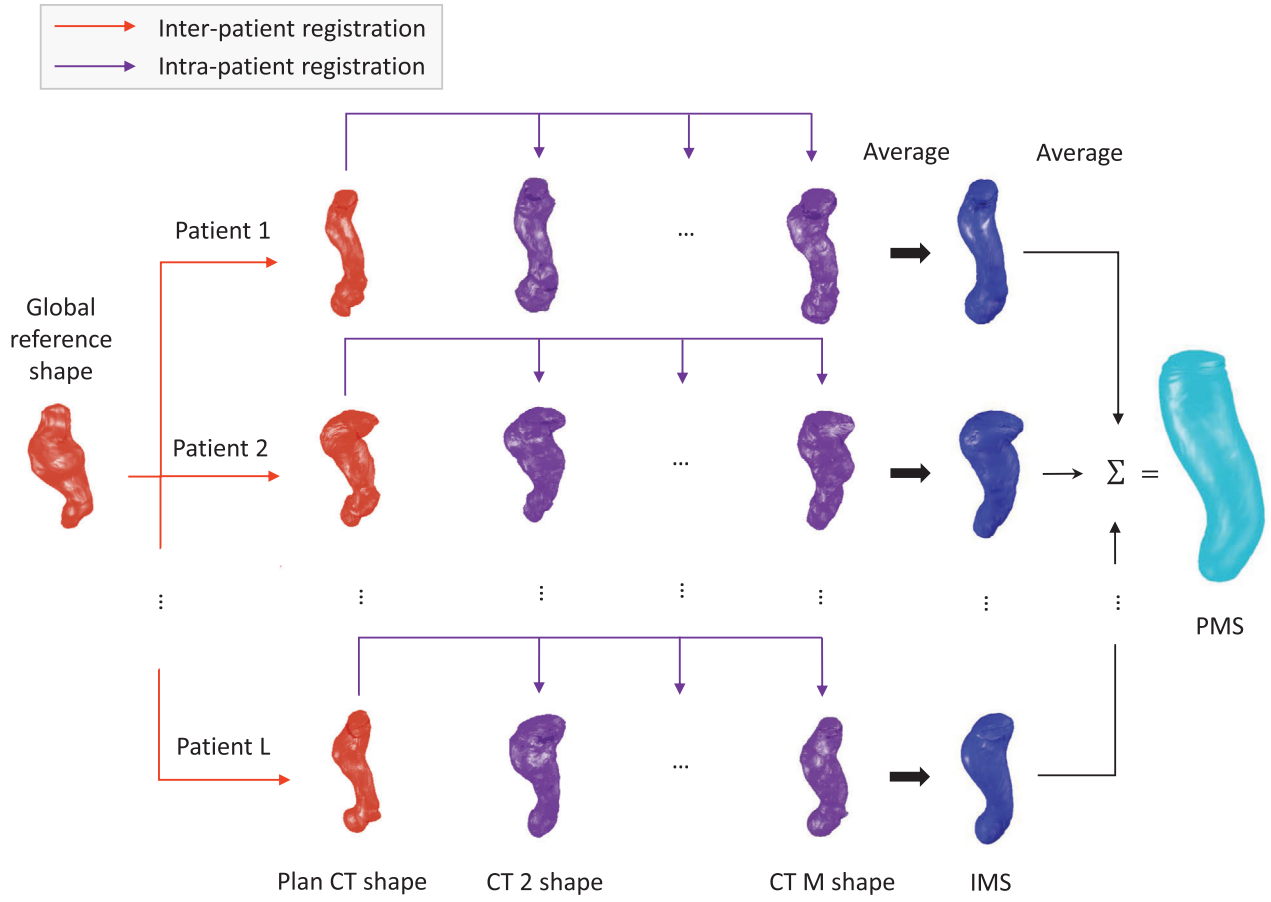

FIGURE 1 Procedure for using inter- and intra-patient registration to compute the individual mean shapes (IMS) and the population mean shape (PMS)

and then from the pCT shape to the treatment CT shape.

For a new patient, deformable registration needs only be performed between the patient's pS and the PMS. Thus, the procedure in Figure 1 is only needed in order to build the PMS and to evaluate the method.

With point-to-point correspondence, we can compare and combine rectal shapes. For each shape, the coordinates of the points were gathered into a vector:

$$[x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_N, y_N, z_N]^T.$$

Let the PMS and the pS be described by such vector representations. The shrinkage estimated shape (SES) of the patient mean shape can then be computed as:

$$\text{SES} = pS + \lambda\,(\text{PMS} - pS).$$

Here, λ is the shrinkage factor—a value between 0 and 1 that determines weighing between the pS and the PMS. A small λ means little shrinkage, that is, that the SES is close to the pS, while a large λ means that the SES is closer to the PMS.

Matching the coordinate systems of the PMS and the pS was secured by a rigid shift of the PMS before combining with the pS. The shift was performed such that the center of gravity, that is, the coordinate average, matched between the PMS and the PS. We saw that the cranial region, where the variation and uncertainty is greatest, negatively affected the matching; therefore, we left out the 50% most cranial points when calculating the center of gravity.

2.1 | Computing the shrinkage factor

The value of the shrinkage factor λ that minimizes the mean squared error (MSE) over the training set can be found analytically. The treatment CT rectum shapes (but not the pCT shapes) was used to estimate each patient's individual mean shape (IMS), which was used as the target in the optimization of the shrinkage factors. Given the PMS vector, the pS_i vector, and the IMS_i vector for patient i, the MSE is given by

$$\text{MSE}\ (\lambda) = \frac{1}{LN} \sum_{i=1}^{L} \|pS_i + \lambda\,(PMS - pS_i) - IMS_i\|^2,$$

with a minimum at

$$\lambda_{\mathrm{MMSE}} = \frac{\sum_{i=1}^{L} (IMS_i - pS_i)^T (PMS - pS_i)}{\sum_{i=1}^{L} \|PMS - pS_i\|^2}.$$

2.2 | Validation

The method was evaluated using the rectum shapes of 37 prostate cancer patients, who each had a pCT and seven to 10 treatment CTs acquired during their course of RT – a total of 373 CT scans. The CT resolution was 0.7 mm in-plane and 2 mm in the z-direction. The rectum shapes for all CTs were contoured by expert physicists and quality assured by two independent physicists.

The SESs used for validation were produced using leave-one-out cross validation (CV), where for each patient, the data were separated into a test set containing this patient only, and a training set comprising the remaining 36 patients. The training set was used to compute the PMS and the shrinkage factor for the test patient's SES. The SESs computed this way were used in all validation.

2.2.1 | Volumetric similarity

To reduce the systematic error, the SES must resemble the IMS more than the pS does. We used the Dice similarity index (DSI) to assess the similarity between shapes. The DSI between two shapes X and Y is defined as $\frac{2|X \cap Y|}{|X|+|Y|}$, where $|\cdot|$ indicates volume, and $\cap$ is the intersection operator. The DSI ranges from 0 to 1, with a higher value indicating more overlapping shapes. For each patient, both the pS and the SES were compared to the IMS to evaluate improvement in the novel method.

2.2.2 | Surface similarity

The DSI is related to the proportion of the volume that is shared between two structures, but the actual numbers can be hard to interpret. A more tangible measure might be to compare distances between surfaces, in our case between the pS and the IMS and between the SES and the IMS, with results in mm. We used the bidirectional local distance (BLD) as a distance metric. The BLD, introduced by Kim et al.,[20] is an extension to the Hausdorff distance (HD) to include local, pointwise distances. If the one-directional local distance from a point a in mesh A to a mesh B is defined as

$$\mathrm{OLD}\ (a, B) = \min_{b \in B} \|a - b\|,$$

then the bi-directional local distance from a to B is

$$\mathrm{BLD}\ (a,\ B) = \max\left(OLD\,(a,\ B),\ \max S_{a,\ B}\right).$$

Where $S_{a,B}$ is the set of all local distances $OLD(b \in B, A)$ where the endpoint in A is a.

We present results for the median, mean, and maximum BLD, where the latter is the same as the HD.

We also used the BLD to study the spatial distribution of the systematic error. We averaged the distance from the pS to the IMS for each point on the organ surface across the population. Changes to the spatial distribution from the shrinkage estimate were analyzed by comparing the pointwise distances between the pS and the IMS to the pointwise distances between the SES and the IMS.

2.2.3 | Evaluation of dosimetric impact

Dosimetric evaluation was based on a retrospective analysis of clinical IMRT plans for locally advanced prostate cancer, including treatment to the prostate, seminal vesicles, and the pelvic lymph nodes. Hypofractionated RT was prescribed in 25 fractions simultaneously delivering fraction doses of 2.7 Gy to the prostate clinical target volume, 2.4 Gy to prostate and seminal vesicles and 2.0 Gy to a larger target also including the pelvic lymph nodes, see Hysing[1] for details. For all patients, the structures pS, IMS, and SES were imported into Varian Eclipse for dosimetric analysis. We do not have available the true accumulated dose over all fractions. As a substitute, we used the dose-volume histogram (DVH) based on the IMS as a representation of the ground truth. This removes systematic errors, but does not take into account the random variations that occur from fraction to fraction.

For dosimetric comparison, the population average DVH and its 95% confidence interval was computed separately for the pS, SES, and IMS. In addition, two parameters were extracted from the DVHs: The equivalent uniform dose (EUD) with a volume factor of 12 and the minimum dose to the hottest 5% volume (D5%). These parameters have been shown to correlate with late rectal toxicity by Söhn et al.[32] and Thor et al,[33] respectively. The average differences between these parameters from the PS to the IMS and from the SES to the IMS were computed.

2.2.4 | Statistical tests

All tests and calculations were performed using Matlab v R2020b with the statistics and machine learning toolbox. All significance levels were set at $\alpha = 0.05$.

The 95% CI of all the geometric similarity and dosimetric metrics was calculated by the bootstrap method

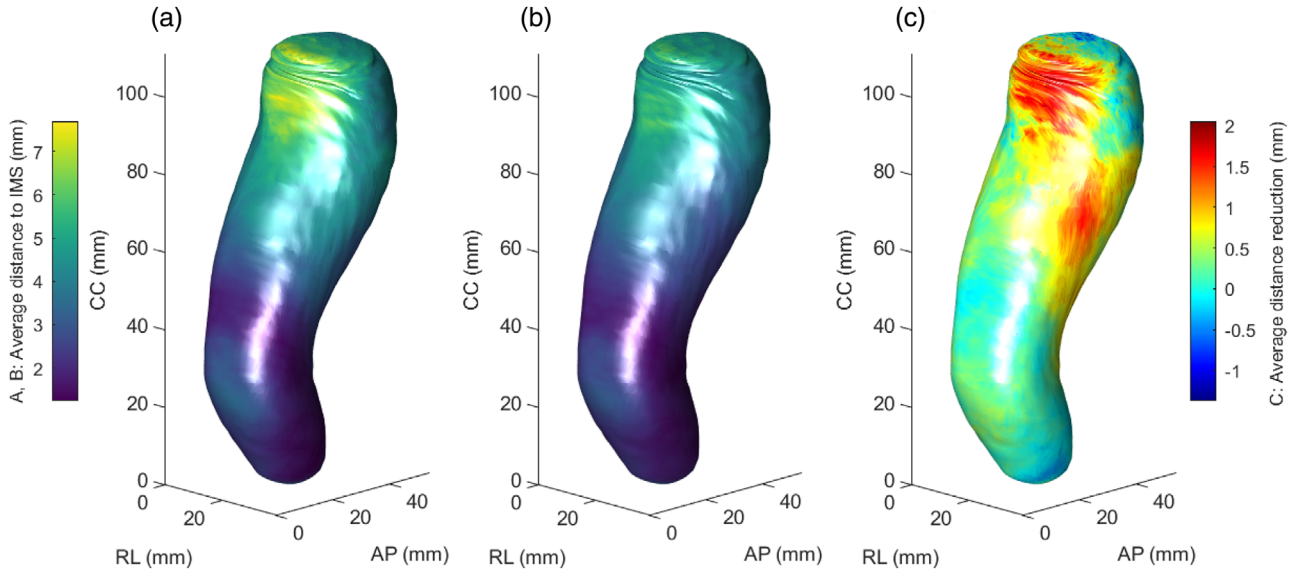

FIGURE 2 Population average systematic error over the rectum (represented by the population mean shape [PMS] rectum), for the pS versus the shrinkage estimated shape (SES). (a) Local distance (bidirectional local distance [BLD]) between the pS and the individual mean shape (IMS). (b) Local distance (BLD) between the SES and the IMS. (c) Difference between (a) and (b), that is, improvement when using SES

TABLE 1 Results for the geometric comparison metrics: Dice similarity index (DSI), median and mean bidirectional local distance (BLD), and Hausdorff distance (HD)

	pS to IMS		SES to IMS		Improvement			
	μ	σ	μ	σ	μ	*CI*	*p-value*	*% +*
DSI	0.74	0.07	0.78	0.06	0.04 (15%)	0.03–0.06	5.8 e-6	89%
Median BLD (mm)	2.9	1.0	2.4	0.7	0.5 (17%)	0.3–0.7	2.0 e-5	84%
Mean BLD (mm)	3.6	1.1	3.1	1.0	0.5 (13%)	0.3–0.7	2.7 e-5	84%
HD (mm)	15.6	5.8	13.4	4.9	2.2 (14%)	1.2–3.2	1.8 e-4	76%

Note: For DSI, the average improvement in percent was calculated by dividing the absolute improvement (0.04) by the maximum achievable improvement (1–0.74 = 0.26).
Abbreviations: CI, confidence interval; IMS, individual mean shape; SES, shrinkage estimated shape.
The column symbols are μ: mean; σ: standard deviation; % +: percentage of of patients with improvement.

with one million samples. The Wilcoxon signed rank test was used to test for difference in the median values in DSI, since the DSI is bounded between 0 and 1 and therefore not normally distributed. For the other values, the paired *t*-test was used, after testing for evidence against normality with the Anderson-Darling test.

3 | RESULTS

The PMS shape in Figure 2 shows the spatial distribution of systematic errors being heterogeneously, distributed over the rectum. The population average pointwise distance from the pS to the IMS was largest in the cranial end, with values in the range of 5–7 mm (Figure 2a). The average distance decreased steadily toward the caudal part of the rectum to below 2 mm.

The minimum MSE shrinkage factor for the whole 37-patient dataset was 0.37. Under CV, where one patient was held out of the training data for each validation, the shrinkage factors ranged from 0.34 to 0.38. Using these factors to estimate the SES for each patient decreased the average distance between the SES and the IMS across the population for most parts of the rectum (Figure 2c). The population average improvement was greatest, an error reduction of 2 mm, at the cranial-anterior part of the rectum.

The results for the geometric similarity metrics are shown in Table 1. All metrics showed moderate (13%–17%) but statistically significant improvement. The individual results for the DSI and the median BLD are shown in Figure 3. Although the average improvement is moderate, a very high percentage of patients did show improvement – 33/37 had improvement in DSI, 31/37 in median and mean BLD, and 28/37 in HD.

The rectum shapes (pS, SES, and IMS) of three example patients are shown in Figure 4. These patients represent the 10th, 50th, and 90th percentiles in terms of DSI improvement and illustrate the SES model for different

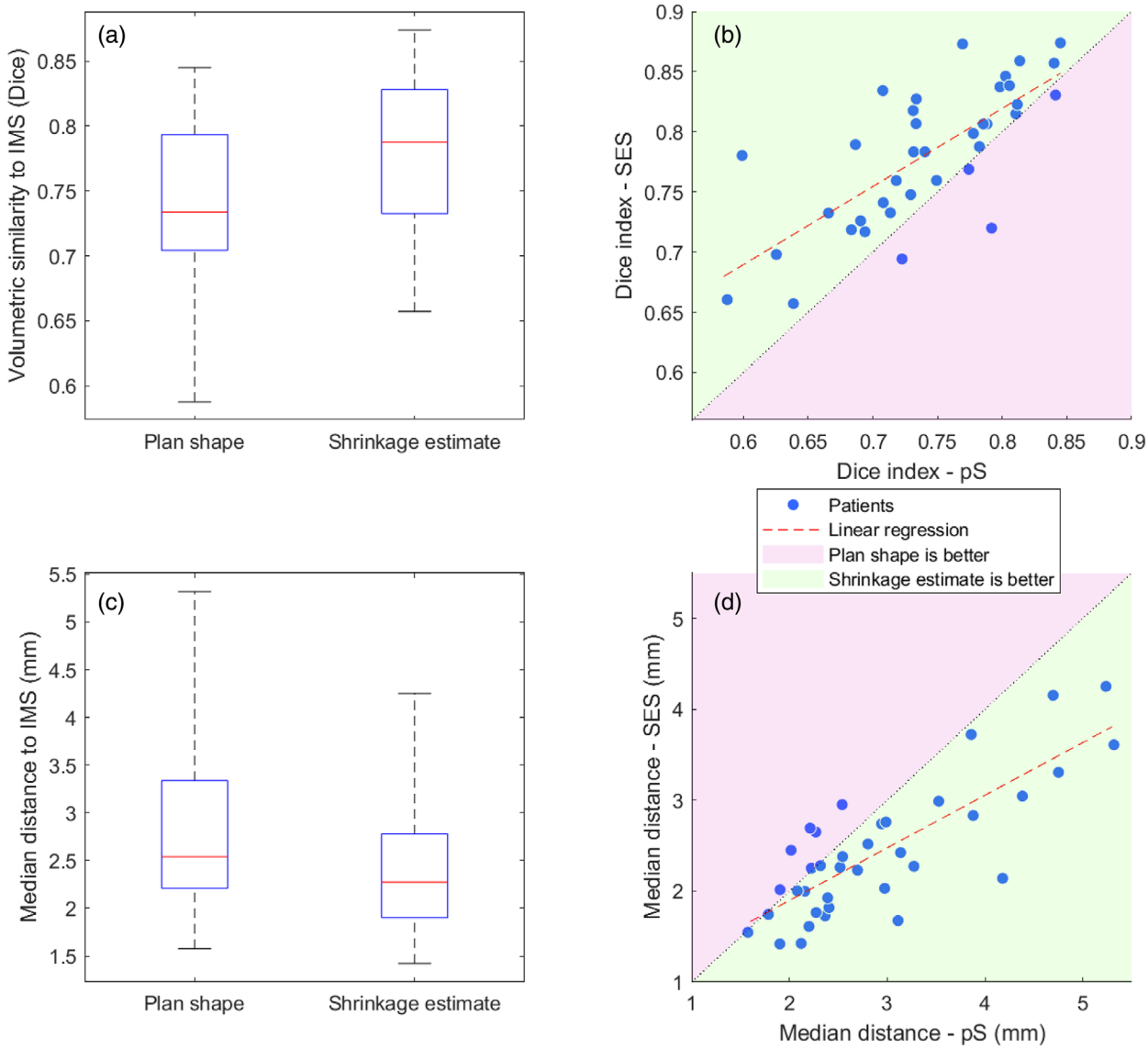

FIGURE 3 Similarity to the individual patient mean for the shrinkage estimated recti and the planning CT recti in each patient. (a and c) Box plots of the similarity to the patient mean shape of the plan shape and the shrinkage estimate as median (red line) with 25th and 75th percentiles and maximum and minimum values (whiskers). (b and c) scatter plots of the similarity to the patient's mean shape for the plan shape versus the shrinkage estimate. In (a) and (b) (Dice similarity), higher is better, while in (c) and (d) (distance), lower is better

geometries. In the 90th percentile patient, the pS shows more bending in the lower (anorectal) flexure than the average, but the shrinkage estimate has overcompensated this feature. In the 50th percentile patient where the mean rectum volume was reduced compared to the pS, the shrinkage estimate reduced the volume further, and thereby achieved better conformity to the IMS. In the 10th percentile patient, the mean rectum was larger, and the shrinkage method compensated for this.

Population average DVHs derived from the pS, SES, and IMS structures are shown in Figure 5. In the low dose region, the pS and SES show similar bias. In the higher dose region, the SES is closer to the IMS. The differences were significant ($p < 0.05$) for doses in the range of 52.5–65.0 Gy. In this range, the bias was, on average, reduced from −1.3 percentage points to −0.2 percentage points.

The results for the dose metrics EUD and D5% are shown in Table 2. For both metrics, the SES gave a better dose estimate on average, but the improvement was not significant. This can be anticipated by investigating Figure 6, which shows the structures with dose for an example patient. The higher doses are restricted to the anterior-middle part of the rectum. The high dose (D5%) region is where the similarity between the SES and PS to the IMS is the greatest.

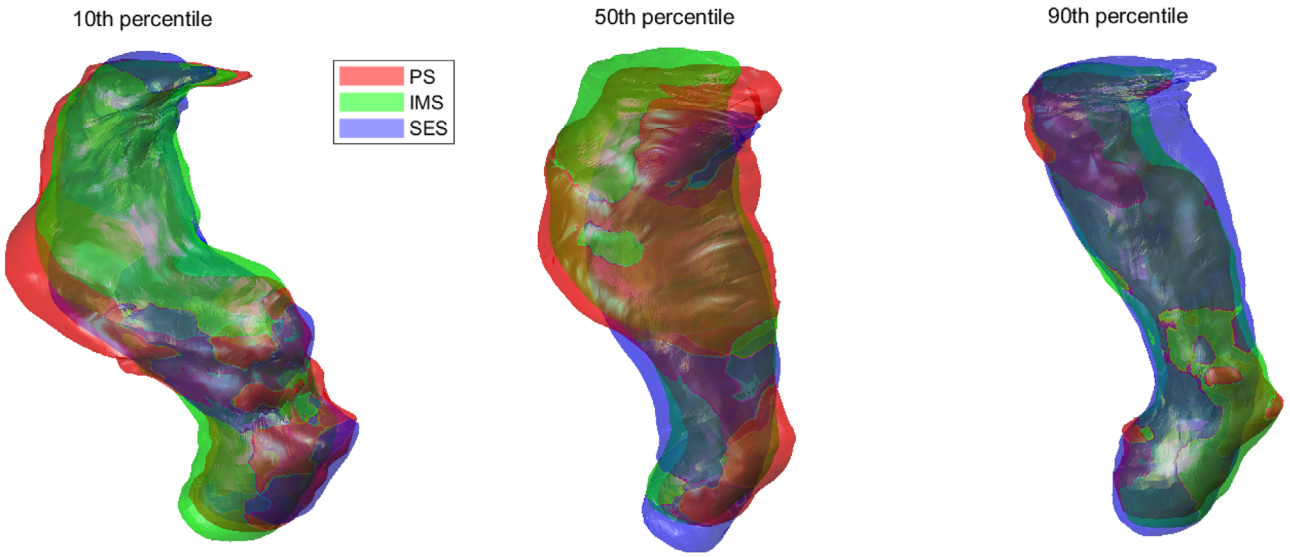

FIGURE 4 Example recti in three different patients with the planning structure (red), the shrinkage estimated rectum (green), and the individual mean shape (blue). The patients were chosen based on their percentiles in terms of improvement in Dice index to the mean rectum, ranging from poorer (left) to better (right)

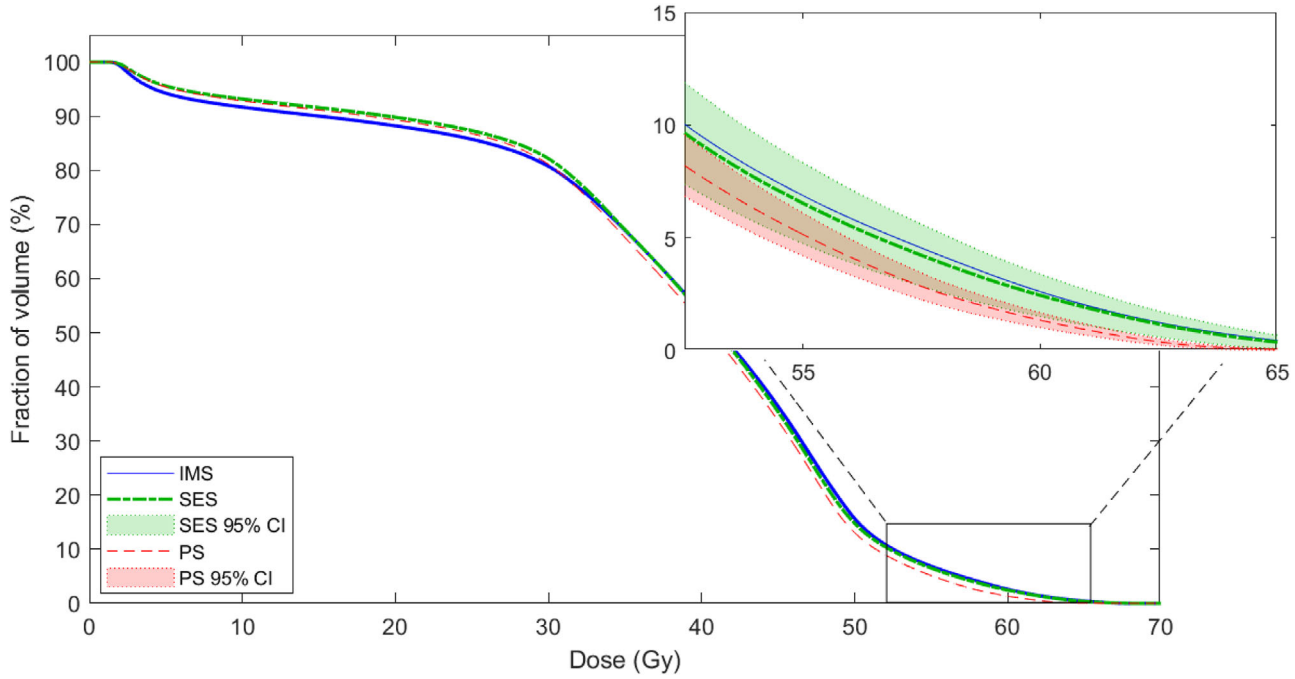

FIGURE 5 Population average DVHs of the three rectum structures individual mean shape (IMS) (blue), planning shape (pS) (red) and shrinkage-estimated shape (SES) (green). The magnification shows the region where the SES showed significantly less bias than the IMS, that is, the dose range from 52.5 to 65.0 Gy

TABLE 2 Comparison of dosimetric parameters between the pS and the SES. The column symbols are μ: mean; σ: standard deviation

	IMS		pS error		SES error		Improvement		
	μ	σ	μ	σ	μ	σ	μ	*CI*	*p-value*
D5% (Gy)	55.8	3.6	−1.4	3.3	−0.6	3.4	0.8	0.0–1.7	0.070
EUD (Gy)	49.0	2.5	−1.1	2.1	−0.6	2.4	0.5	0.0–1.2	0.11

Abbreviations: CI, confidence interval; EUD, equivalent uniform dose; IMS, individual mean shape; SES, shrinkage estimated shape.

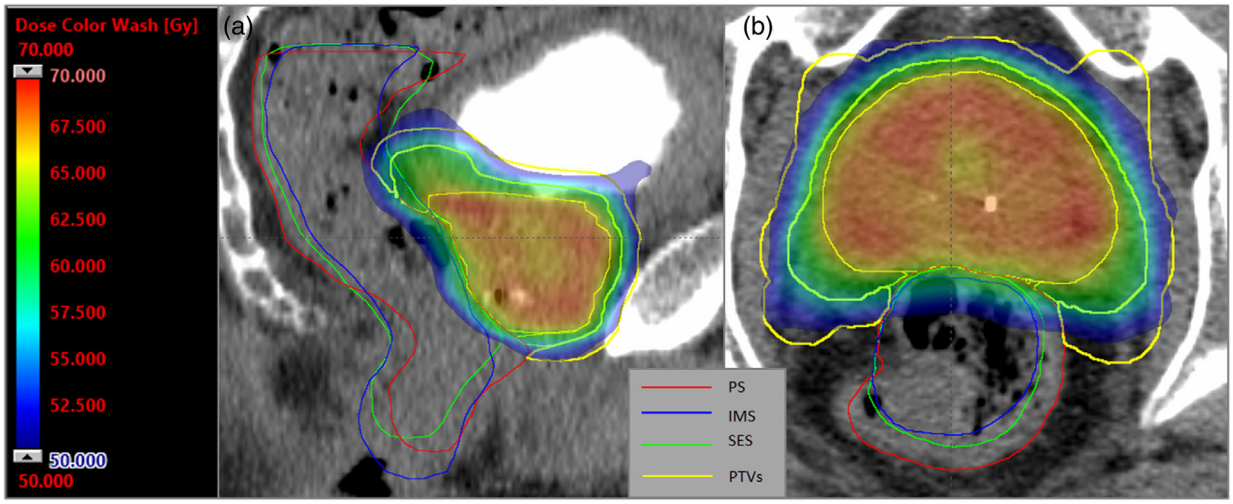

FIGURE 6 2D view of rectum and PTV structures with dose color wash for an example patient. The PTVs, from outermost to innermost, are PTV50, PTV60, and PTV67.5. (a) sagittal view, (b) transversal view. The dose threshold is 50 Gy, identical to the prescription to the PTV50 and to the lowest dose level used to assess rectal exposure in our planning procedure

4 | DISCUSSION

The proposed method showed significantly improved geometric similarity to the patient mean rectum. The average improvement is about 15% in terms of both distance and volume. However, this is just the average; several patients had much greater improvement (Figure 3). A positive trait is that the patients that have the greatest systematic error in the pCT are those that typically see the greatest improvement. Also, the method is low-risk, as few patients see worse systematic errors.

Whether this improvement is important will depend on the application, the dose distribution, and other factors. In our dataset, the improvements in geometric similarity did not translate into significant improvement of estimated accumulated dose to the rectum. This may be due to the procedures of image-guidance that we used. All treatment CTs were rigidly aligned to gold markers in the prostate, following our clinical procedures of RT for these patients. This means that all movements are relative to the prostate, and there are no fixed points in the rectum. The volume of the rectum near the prostate is therefore rather stable, which contributes to the relatively low improvement in the anterior region where the highest doses are located (Figure 2c). For the example patient seen in Figure 6, the SES is closer to the IMS almost everywhere except in the high-dose region, where the SES, the pS, and the IMS are all very similar. Because the patients were part of a trial in hypofractionation, extra care was taken to avoid rectal toxicity by cropping the planning target volume (PTV) where it would otherwise overlap with the rectum, also seen in Figure 6. Clinical results have proven this fear unfounded, and the practice has since been changed into cropping the rectum in such cases.[1]

Systematic shape errors due to rectal deformation have previously been studied by Hoogeman et al.[21] and Haekal et al.[22] There, the treatment CTs were not registered to gold markers, and, contrary to our results, both studies found the largest systematic errors in the anterior region. As such, the current results are a testament to the success of fiducial markers. If the higher doses were delivered to volumes where the systematic shape errors are greater, we would likely have seen higher impact from the method also in the dosimetric analysis. We can therefore expect to see different results if the method is applied to different patient groups, for example, bladder or cervical cancer, or different OARs.

The amount of dose degradation in EUD and D5% that is acceptable is not known for the rectum. This is likely dependent on several factors additional to organ motion, for example, dose distribution and delivery modality, nonuniform radio-sensitivity of the rectum.[13,23,24] In our clinical procedure, the patients in the present study were planned for a maximum of 60 Gy (about 62GyEDQ2) delivered to 10 cc of the rectal volume and with restrictions to the volume receiving 50 Gy.[1] This resulted in very little rectal toxicity, indirectly indicating that the obtained changes to the EUD from pCT in these patients were acceptable.[25]

Since dose accumulation across fractions is difficult, we used the dose to the IMS as a substitute for the true dose when comparing dose-volume parameters. Under the assumption that the rectum is in a dose region with a constant gradient, the dose to the IMS will be the same as the expected accumulated dose to the deforming organ. This is because each voxel moves both into higher and lower dose regions, which, in the mean, cancel each other out.[a] Although this assumption is false,

[a] Formally, if x is the (random) position of a voxel, with distribution $p(x)$, and $D(x) = Gx$ is the dose at position x, with constant gradient G, then $E[D(x)] = \int_{-\infty}^{\infty} D(x)p(x)dx = G\int_{-\infty}^{\infty} xp(x)dx = GE[x] = D(E[x])$.

the smoothness of the gradient outside the PTV justifies the use of the dose to the IMS as an approximation of the true accumulated dose.

We have chosen to present both volumetric (DSI) and distance-based (DS and BLD) similarity metrics for the evaluation of our proposed method. Distance-based metrics are easier to interpret and are more relevant for radiotherapy (RT), for example, to assess the influence of deformation in relation to a dose gradient. The DSI, on the other hand, includes the whole structure as opposed to a single point and is a common metric to evaluate the performance of deform registration.

The deformable registration will introduce some geometric distortion. The distortion is greater for inter-patient registration than the intra-patient, since there is no "true" transformation between patients. However, the improvement in DSI is itself a proof-of-concept of the inter-patient registration. Our registrations were optimized to yield high forward accuracy and backward consistency. The average forward accuracy (see Osorio et al.[18] for definition) was measured at 0.23 mm for the intra-patient registrations and 0.22 mm for the second iteration of inter-patient registrations (0.56 mm in the first iteration). In addition, the obtained vector fields were checked visually. Our results thus reflect a rather meticulous verification of the registrations, which may be more difficult to achieve in commercial registration software.

The choice to use the 50% most caudal points in the rigid alignment between the pS and the PMS was relatively arbitrary; we saw the need to exclude the cranial part of the rectum and at the same time include the points near the prostate, but apart from that, the exact percentage of points chosen did not make a great difference. The method can potentially be improved by optimizing the rigid shift w.r.t geometry or dose using training data.

Hoogeman et al.[26] have previously investigated the possibility of predicting systematic changes to the rectum and prostate based on the pCT rectal volume alone. For the rectum, they achieved best results in the upper anterior part, with a 30% improvement in one coordinate (AP). Their results are, however, not directly comparable to ours, as their images were aligned to the bony anatomy instead of gold markers.

In this paper, we have looked at the possibility of measuring dose on the SES to estimate the accumulated dose to the rectum. In clinical applications, discrepancies between the SES and the pS — either geometrically overall, locally at key regions or in calculated dose — could be used for screening of patients for further imaging, or for additional verification based on in-treatment CBCTs, which may again indicate adaptive RT.

Other applications include motion as well as normal-tissue complication probability modeling. Many deformable organ motion models do not account for systematic error or require multiple CTs to do so.[19,27–31] Such models can be improved or simplified by replacing the pS or the multiple-CT average by the SES. For voxel-based normal tissue complication probability (NTCP) models, the method can be used to reduce the bias in the pCT.

It is also possible to use the method in plan optimization: A robust optimization algorithm may for example consider the pS and the SES as separate scenarios. One complicating factor is the need to take into account the correlation between target motion and the motion of OARs; for example, the SES rectum may overlap with the target. However, the method does not need to be restricted to a single organ. A potential solution is therefore to include multiple organs (in this case, at least the rectum and the prostate) in the shrinkage method. Still, inclusion of the target is not straightforward, as tumor shrinkage must be taken into account. The accuracy of the predicted anatomy should be verified through daily CBCT scans.

5 | CONCLUSIONS

The shrinkage method was successfully applied to improve estimates of the individual mean rectum shapes of prostate cancer patients when only the planning CT is available. The method reduced the systematic shape changes, especially for patients where the similarity between the planning CT and the individucal mean shape was poor. The method could be applied to more complex motion modeling, plan optimization as well as toxicity assessments.

ACKNOWLEDGMENTS

We gratefully thank Prof. Heijmen and Prof. Hoogeman and PhD Zolnay at Erasmus MC Cancer Institute, Rotterdam for valuable discussions and advice on motion modeling as well as sharing the deformable registration software. We also thank Maria Thor and Lise Fjellsbø for rectum contouring. Thanks to Helge Egil Seime Pettersen for invaluable help in importing structures to Eclipse. The work was funded by Trond Mohn Foundation (grant number: BFS2017TMT07).

CONFLICT OF INTEREST

The authors have no relevant conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

The data representing the population mean rectum shape are available in the supplementary material of this article. The raw data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical restrictions.

REFERENCES

1. Hysing LB, Ekanger C, Zolnay Á, et al. Statistical motion modelling for robust evaluation of clinically delivered accumulated dose distributions after curative radiotherapy of locally advanced prostate cancer. *Radiother Oncol*. 2018;128(2):327-335.

2. Unkelbach J, Alber M, Bangert M, et al. Robust radiotherapy planning. *Phys Med Biol*. 2018;63(22):22TR02.
3. Muren LP, Ekerold R, Kvinnsland Y, Karlsdottir À, Dahl O. On the use of margins for geometrical uncertainties around the rectum in radiotherapy planning. *Radiother Oncol*. 2004;70(1):11-19.
4. de Boer HC, Heijmen BJ. A protocol for the reduction of systematic patient setup errors with minimal portal imaging workload. *Int J Radiat Oncol Biol Phys*. 2001;50(5):1350-1365.
5. Söhn M, Birkner M, Yan D, Alber M. Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation. *Phys Med Biol*. 2005;50(24):5893-5908.
6. Scaife JE, Thomas SJ, Harrison K, et al. Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer. *Br J Radiol*. 2015;88(1054):20150243.
7. Thörnqvist S, Petersen JBB, Høyer M, Bentzen LN, Muren LP. Propagation of target and organ at risk contours in radiotherapy of prostate cancer using deformable image registration. *Acta Oncol*. 2010;49(7):1023-1032.
8. Kleijnen J-PJE, van Asselen B, Burbach JPM, et al. Evolution of motion uncertainty in rectal cancer: implications for adaptive radiotherapy. *Phys Med Biol*. 2015;61(1):1.
9. Liu L, Johansson A, Cao Y, Kashani R, Lawrence TS, Balter JM. Modeling intra-fractional abdominal configuration changes using breathing motion-corrected radial MRI. *Phys Med Biol*. 2021;66(8):085002.
10. Jackson A, Marks LB, Bentzen SM, et al. The lessons of QUANTEC: recommendations for reporting and gathering data on dose-volume dependencies of treatment outcome. *Int J Radiat Oncol Biol Phys*. 2010;76:S155-S160.
11. Michalski JM, Gay H, Jackson A, Tucker SL, Deasy JO. Radiation dose-volume effects in radiation-induced rectal injury. *Int J Radiat Oncol Biol Phys*. 2010;76:S123-S129.
12. Söhn M, Alber M, Yan D. Principal component analysis-based pattern analysis of dose-volume histograms and influence on rectal toxicity. *Int J Radiat Oncol Biol Phys*. 2007;69(1):230-239.
13. Casares-Magaz O, Bülow S, Pettersson NJ, et al. High accumulated doses to the inferior rectum are associated with late gastrointestinal toxicity in a case-control study of prostate cancer patients treated with radiotherapy. *Acta Oncol*. 2019;58(10):1543-1546.
14. Shelley LEA, Scaife JE, Romanchikova M, et al. Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. *Radiother Oncol*. 2017;123(3):466-471.
15. Thor M, Bentzen L, Hysing LB, et al. Prediction of rectum and bladder morbidity following radiotherapy of prostate cancer based on motion-inclusive dose distributions. *Radiother Oncol*. 2013;107(2):147-152.
16. Fourdrinier D, Strawderman WE, Wells MT. Decision theory preliminaries. In: Fourdrinier D, Strawderman WE, Wells MT, eds. *Shrinkage Estimation*. Springer International Publishing; 2018:1-28.
17. Chui H, Rangarajan A. A new point matching algorithm for non-rigid registration. *Comput Vision Image Understanding*. 2003;89(2):114-141.
18. Osorio EMV, Hoogeman MS, Bondar L, Levendag PC, Heijmen BJM. A novel flexible framework with automatic feature correspondence optimization for nonrigid registration in radiotherapy. *Medical Physics*. 2009;36(7):2848-2859.
19. Budiarto E, Keijzer M, Storchi PR, et al. A population-based model to describe geometrical uncertainties in radiotherapy: applied to prostate cases. *Phys Med Biol*. 2011;56(4):1045-1061.
20. Kim HS, Park SB, Lo SS, Monroe JI, Sohn JW. Bidirectional local distance measure for comparing segmentations. *Med Phys*. 2012;39(11):6779-6790.
21. Hoogeman MS, van Herk M, de Bois J, Muller-Timmermans P, Koper PCM, Lebesque JV. Quantification of local rectal wall displacements by virtual rectum unfolding. *Radiother Oncol*. 2004;70(1):21-30.
22. Haekal M, Arimura H, Hirose T, et al. Computational analysis of interfractional anisotropic shape variations of the rectum in prostate cancer radiation therapy. *Physica Medica*. 2018;46:168-179.
23. Troeller A, Yan D, Marina O, et al. Comparison and limitations of DVH-based NTCP models derived from 3D-CRT and IMRT data for prediction of gastrointestinal toxicities in prostate cancer patients by using propensity score matched pair analysis. *Int J Radiat Oncol Biol Phys*. 2015;91(2):435-443.
24. Pedersen J, Flampouri S, Bryant C, et al. Cross-modality applicability of rectal normal tissue complication probability models from photon- to proton-based radiotherapy. *Radiother Oncol*. 2020;142:253-260.
25. Ekanger C, Helle SI, Heinrich D, et al. Ten-year results from a phase II study on image guided, intensity modulated radiation therapy with simultaneous integrated boost in high-risk prostate cancer. *Adv Radiat Oncol*. 2020;5(3):396-403.
26. Hoogeman MS, van Herk M, de Bois J, Lebesque JV. Strategies to reduce the systematic error due to tumor and rectum motion in radiotherapy of prostate cancer. *Radiother Oncol*. 2005;74(2):177-185.
27. Bondar L, Intven M, Burbach JPM, et al. Statistical modeling of CTV motion and deformation for IMRT of early-stage rectal cancer. *Int J Radiation Oncol Biol Phys*. 2014;90(3):664-672.
28. Fontenla E, Pelizzari CA, Roeske JC, Chen GTY. Using serial imaging data to model variabilities in organ position and shape during radiotherapy. *Phys Med Biol*. 2001;46(9):2317-2336.
29. Hoogeman MS, van Herk M, Yan D, Boersma LJ, Koper PCM, Lebesque JV. A model to simulate day-to-day variations in rectum shape. *Int J Radiat Oncol Biol Phys*. 2002;54(2):615-625.
30. Magallon-Baro A, Loi M, Milder MTW, et al. Modeling daily changes in organ-at-risk anatomy in a cohort of pancreatic cancer patients. *Radiother Oncol*. 2019;134:127-134.
31. Rios R, De Crevoisier R, Ospina JD, et al. Population model of bladder motion and deformation based on dominant eigenmodes and mixed-effects models in prostate cancer radiotherapy. *Medical Image Analysis*. 2017;38:133-149.
32. Söhn M, Yan D, Liang J, Meldolesi E, Vargas C, Alber M. Incidence of late rectal bleeding in high-dose conformal radiotherapy of prostate cancer using equivalent uniform dose–based and dose–volume–based normal tissue complication probability models. *International Journal of Radiation Oncology*Biology*Physics*. 2007;67:(4):1066–1073. http://doi.org/10.1016/j.ijrobp.2006.10.014
33. Thor M, Deasy JO, Paulus R, et al. Tolerance doses for late adverse events after hypofractionated radiotherapy for prostate cancer on trial NRG Oncology/RTOG 0415. *Radiotherapy and Oncology*. 2019;135:19–24. http://doi.org/10.1016/j.radonc.2019.02.014

SUPPORTING INFORMATION

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

How to cite this article: Rørtveit ØL, Hysing LB, Stordal AS, Pilskog S. Reducing systematic errors due to deformation of organs at risk in radiotherapy. *Med Phys.* 2021;48:6578–6587. https://doi.org/10.1002/mp.15262