

PAPER

OPEN ACCESS

RECEIVED
9 June 2025REVISED
25 November 2025ACCEPTED FOR PUBLICATION
10 December 2025PUBLISHED
6 January 2026

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

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



Modality-agnostic, patient-specific digital twins modeling temporally varying digestive motion

Jorge Tapias Gomez^{1,2,*} , Nishant Nadkarni² , Lando S Bosma³ , Jue Jiang² , Ergys D Subashi⁴ , William P Segars⁵ , James M Balter⁶ , Mert R Sabuncu¹ , Neelam Tyagi^{2,7} and Harini Veeraraghavan^{2,7}

¹ Computer and Information Science, Cornell University, New York, NY, United States of America

² Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY, United States of America

³ University Medical Center Utrecht, Utrecht, The Netherlands

⁴ Department of Radiation Physics, University of Texas MD Anderson Cancer Center, Houston, TX, United States of America

⁵ Carl E. Ravin Advanced Imaging Laboratories and Center for Virtual Imaging Trials, Duke University Medical Center, Durham, NC, United States of America

⁶ Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, United States of America

⁷ Co-Senior Author.

* Author to whom any correspondence should be addressed.

E-mail: tapiasj@mskcc.org

Keywords: digital twins, gastrointestinal motion modeling, deformable image registration validation, dose accumulation validation for MR guided adaptive radiotherapy

Supplementary material for this article is available [online](#)

Abstract

Objective. Clinical implementation of deformable image registration (DIR) requires voxel-based spatial accuracy metrics such as manually identified landmarks, which are challenging to implement for highly mobile gastrointestinal (GI) organs. To address this, patient-specific digital twins (DTs) modeling temporally varying motion were created to assess the accuracy of DIR methods. **Approach.** A total of 21 motion phases simulating digestive GI motion as 4D image sequences were generated from static 3D patient scans using published analytical GI motion models through a multi-step semi-automated pipeline. Eleven datasets, including six T2-weighted FSE MRI (T2w MRI), two T1-weighted 4D golden-angle stack-of-stars, and three contrast-enhanced computed tomography scans were analyzed. The motion amplitudes of the DTs were assessed against real patient stomach motion amplitudes extracted from independent 4D MRI datasets using hierarchical motion reconstruction. The patient-specific DTs were then used to assess six different DIR methods using target registration error, Dice similarity coefficient (DSC), and the 95th percentile Hausdorff distance using summary metrics and voxel-level granular visualizations. Finally, for a subset of T2w MRI scans collected from patients treated with magnetic resonance-guided radiation therapy, dose distributions were warped and accumulated to assess dose warping errors (DWEs), including evaluations of DIR performance in both low- and high-dose regions for patient-specific error estimation. **Main results.** Our proposed pipeline synthesized patient-specific DTs modeling realistic GI motion, achieving mean and maximum motion amplitudes and a mean log Jacobian determinant within 0.8 mm and 0.01, respectively, similar to published real-patient gastric motion data. It also enables the extraction of detailed quantitative DIR performance metrics and supports rigorous validation of dose mapping accuracy prior to clinical implementation. **Significance.** The developed pipeline enables rigorously testing DIR tools for dynamic, anatomically complex regions facilitating granular spatial and dosimetric accuracies.

1. Introduction

Deformable image registration (DIR) is a critical component of both computed tomography (CT) and magnetic resonance (MR) image-guided adaptive radiotherapy (Rigaud *et al* 2019, Lowther *et al* 2022,

Meyer *et al* 2025). It enables automated propagation of organ-at-risk (OAR) segmentations across treatment fractions and facilitates accurate estimation of the radiation dose delivered to both OARs and target structures throughout the treatment course. A key obstacle towards clinical implementation of DIR, particularly for luminal gastrointestinal (GI) organs where motion can range from a few millimeters up to 3 cm (Mostafaei *et al* 2018, Liu *et al* 2021, Zhang *et al* 2021), is the lack of methods to assess DIR accuracy at a patient-specific and granular level.

Target registration error (TRE), which measures the displacement errors of known key points as outlined in TG-132 (Brock *et al* 2017), provides a granular and unbiased assessment of DIR accuracy. In contrast, segmentation overlap metrics are influenced by the size of evaluated structures, potentially limiting their sensitivity and interpretability. However, TRE requires manually or semi-automatically placed visually discernible keypoints, which are difficult to identify reliably in GI organs due to their large and often unpredictable motion across and within radiation treatment fractions. Keypoints are typically placed in high-contrast regions with large intensity gradients, areas where DIR algorithms tend to perform well, potentially biasing the evaluation. Manually placed keypoints may not reflect registration accuracy in low-contrast soft-tissue regions, leading to an overestimation of DIR performance. Previous studies have highlighted the importance of using digital phantoms to address this limitation and evaluate DIR methods for dose accumulation (Papachristou *et al* 2024) and even generated ground truths to assess the performance of DIR methods for prostate cancer (Katsoulakis *et al* 2024). The goal of this study is to develop a patient-specific digital twin (DT) (Bosma *et al* 2021, 2024) framework to serve as known ground truth for evaluating the accuracy of DIR and accumulated dose estimates in the GI tract.

Population-level digital phantoms simulating respiratory and GI motions have been previously developed (Segars *et al* 2013, Subashi *et al* 2023). Prior work by Subashi *et al* implemented various GI motion patterns including peristalsis, rhythmic segmentation in the stomach, small and large bowel, high amplitude propagating contractions (HAPCs) in the large bowel, and tonic contractions in the GI sphincters using published analytical models within an XCAT computational phantom (2023). Prior studies have also focused on extracting individual motions such as breathing versus stomach contractile motions to extract the amplitude and time scale of such motions from real patients (Johansson *et al* 2021, Zhang *et al* 2021). However, while population-based anatomical models are an important tool for advancing research in medical imaging and radiation therapy (Segars *et al* 1999, Caon 2004, Johnson *et al* 2009, Cassola *et al* 2011, Ding *et al* 2012, Bosca and Jackson 2016), they fall short in capturing many of the complexities of individual patient anatomy.

First, population-level models such as XCAT can generate a wide range of physiologically realistic motions; however, these motions are simulated on generic male and female anatomies. In contrast, patients exhibit nuanced anatomical variations that cannot be fully captured by population-level phantoms. Second, XCAT-based MR phantoms rely on predefined signal intensity values to represent body composition, which cannot be used to evaluate deep learning (DL) DIR methods, as they are sensitive to intensity distribution shifts from training patient scans.

To address the aforementioned limitations of XCAT phantoms, we developed a framework to generate patient-specific DTs simulating temporally varying GI motion. Our DT framework aligns with the definition of DTs set forth by the ecosystem DTs in health (EDITH) Coordination and Support Action funded by the European Commission (EDITH, 2022). EDITH defines DTs for health as computer simulations, incorporating both knowledge-driven and data-driven models to predict clinically relevant quantities otherwise challenging to experimentally measure. In accordance with EDITH, our framework predicts feasible anatomic variations starting from 3D scans for MR guided adaptive treatment planning, where acquiring 4D MRIs are not clinically performed during treatment. In addition, our models used published amplitude and time-scale data derived from real patient imaging (Zhang *et al* 2021) and apply this motion directly to each patient's scan.

Our pipeline starts from individual patient imaging data, extracting the various organs using a semi-automated manner, and creating a digitized representation as non-uniform rational B-spline (NURBS) surfaces. A motion pattern is then applied to this model to generate a patient-specific sequence of deformations simulating organ motion. We validate these simulated motion patterns by comparing them to real patient motion data from the literature.

Our key contributions include: (a) development of a patient-specific DT framework that simulates temporally varying GI motion and provides known ground truth for DIR evaluation, (b) demonstration of this framework across multiple imaging modality, including contrast-enhanced CT (CECT), T2-weighted fast spin echo MRI (T2w MRI), and T1-weighted 4D golden-angle stack-of-stars MRI (T1w gaSOS), (c) the first framework, to our knowledge, enabling granular, patient-specific assessment of registration and dosimetric accuracy in the GI tract, and (d) a modality and algorithm independent evaluation framework, demonstrated on both iterative and DL-based DIR methods.

We envision our approach to be useful for performing day-to-day and patient-specific assessments of a preferred DIR method used in a clinic for the MR-Linac (Fallone 2014, Keall *et al* 2014, Lagendijk *et al* 2014, Mutic and Dempsey 2014) enabled adaptive radiation treatment or assessing quality of treatment planning simulations.

2. Materials and methods

2.1. DT pipeline for GI motion simulation

We developed a semi-automated pipeline to generate patient-specific DTs that simulate temporally varying GI motion. The pipeline produces 4D image sequences by deforming static 3D patient scans using analytical GI motion models (Subashi *et al* 2023) applied to organ-specific surface representations. Twenty-one motion phases were synthesized to match the temporal resolution of stomach contractions reported by Johanssen *et al* (2021). The pipeline comprises the following steps (figure 1):

- (a) Organ segmentation: A published AI model called the self-distilled masked image transformer (Jiang *et al* 2022), was used to segment the liver, stomach, duodenum, spleen, kidneys, small and large bowel from a given patient image set (MR or CT). Segmentations were manually verified by an expert (figure 1. Step 1, row 1).
- (b) Organ-specific skeleton graph extraction: Morphological thinning of binary 3D segmentation masks is performed to obtain a minimally connected sequence of points by making successive passes of the image and removing the identified border pixel, while preventing a break in connectivity of points spanning the organ (Lee *et al* 1994). This skeleton may contain branches which require further processing (figure 1. Step 1, row 2).
- (c) Organ centerline extraction using the longest medial axis: Starting and ending points are manually selected from each extracted skeleton. From the starting point we build a graph by connecting all directly neighboring points until all points have been explored. Then we apply a breadth-first search algorithm (Lee 1961) to determine the longest connected path graph between the two manually selected points. This path is retained as the organ centerline while all branching structures are discarded (figure 1. Step 1, row 3).
- (d) NURBS surface extraction: For each organ, we sample points along the extracted centerline to create a set of equidistant center points. For each of these points we cast radial vectors perpendicular to the direction of the centerline. These vectors are expanded outward from the centerpoint until they intersect with the boundary of the mask surface. These boundary points on the mask surface served as control points to define the NURBS surface to define the organ's geometry. Each cross-section along the centerline defined a sectional curve (figure 1. Step 2) For the surface. These NURBS surfaces were then used to synthesize GI motion by applying a motion model at the surface points.
- (e) Multi-phase motion synthesis: Analytical motion models developed by Subashi *et al* (2023) were applied to the NURBS surfaces to simulate realistic GI motion patterns using traveling sinusoidal waves. These include HAPCs, tonic contractions, and peristalsis (figure 1. Step 3).

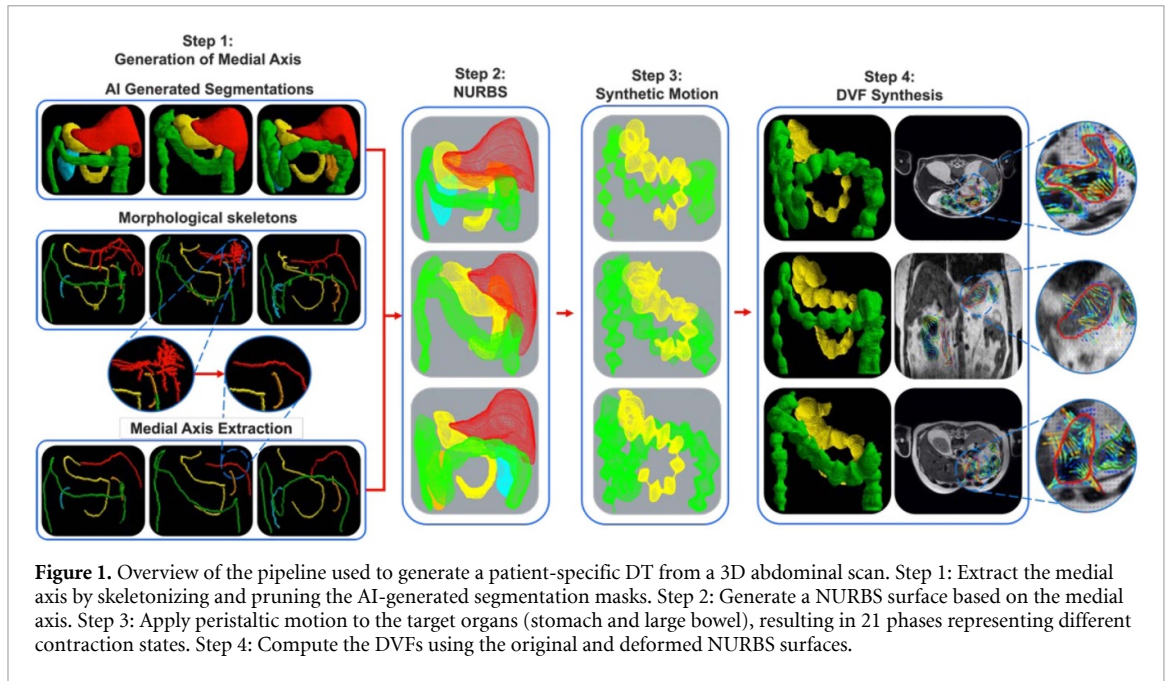
The overall modeling framework is given by the following equation:

$$P'_{i,j} = P_{i,j} + F(\vec{x}_i, t) \cdot D_s \cdot D_t \cdot \vec{d}$$

where $P_{i,j}$ represents the control point j at the sectional curve i of a given NURBS surface. $F(\vec{x}_i, t)$ is a function modeling the non-dispersive component of the wave, D_s and D_t represent functions modeling dispersions of the wave in the spatial coordinate ($\vec{x}_i$) and temporal (t) domain respectively and $\vec{d}$ represents a directional vector to apply a consistent expansion/contraction in the radial direction from the center for each sectional curve.

Using this framework, we iterate through each sectional curve of the given NURBS surface and compute the wave and dispersion magnitudes as functions of space and time. In this work, we specifically focused on modeling the peristaltic motion ($F_{PS}(\vec{x}_i, t)$), within the stomach and large bowel. The wave function was modulated via $F_{PS}(\vec{x}_i, t)$ and an exponential dispersion function $D_z(u)$. Of note, the DT framework is flexible and can be adapted to incorporate any GI motion patterns and dispersion functions. The peristaltic wave and exponential dispersion function are given by:

$$F_{PS}(\vec{x}_i, t) = \frac{1}{\sqrt{3}} \cdot A \cdot \sin\left(2\pi \frac{L_i - c \cdot t}{\lambda}\right)$$



$$D_z(u) = e^{-\alpha \cdot u}$$

where A , c , λ , and α are user-defined parameters representing the amplitude, speed, and wavelength of the peristaltic wave, and the attenuation of amplitude in the spatial or temporal domain, respectively. L_i is the position of the sectional curve along the organ's length, and u is a generalized variable that corresponds to either $\vec{x}_i$ or t , and z is an index identifying the spatial (s) or temporal (t) domain.

As illustrated in figure 2, for each motion phase we iterate through each sectional curve (figure 2. Line 3) and each control point in the NURBS (figure 2. Line 8) calculating and applying the deformation to each control point (figure 2. Line 9–15). Each full iteration over the centerline of the NURBS generates one phase. Applying the motion models to the NURBS surfaces yields a sequence of deformed surfaces, which serve as the basis for computing the corresponding deformation vector fields (DVFs).

- (f) (f) Extraction of DVFs: From the sequence of deformed NURBS surfaces, we can derive a 3D DVF that captures the deformation for each NURBS by:

Computing inner-shell vectors using radial interpolation. To capture internal surface deformation within each GI organ, intermediate 'inner shell' surfaces were generated between the NURBS outer surface and the organ's centerline. These shells were parameterized by a radial factor $p \in [0, 1]$, with $p = 0$ at the outer surface and $p = 1$ at the centerline. For each shell defined by p_k , a new NURBS surface was computed by interpolating between the outer and centerline surfaces. Sampling each shell over a uniform (u, v) grid yielded corresponding 3D points in both the original and deformed configurations. These parameters are standard in NURBS surface definitions, where u and v define positions within the 2D parameter space of the surface.

Deformation vectors: Deformation vectors were then calculated pointwise as the difference between the deformed and original positions: $V_{\text{surf}}(u_i v_j, p_k) = X'(u_i v_j, p_k) - X(u_i v_j, p_k)$

Voxelization and smoothing of the DVF: The computed deformation vectors across all (u, v, p) samples were mapped into a fixed 3D voxel grid by assigning each vector to a voxel based on its original undeformed position. When multiple vectors mapped into the same voxel, they were averaged. To improve continuity, a spatial smoothing step was applied wherein zero-motion voxels were iteratively updated by averaging the vectors of neighboring non-zero voxels until the average change in each voxel was less than 0.001 mm.

This will generate a DVF that can be applied to the patient's static scan to deform it in a way that reflects both surface and internal anatomical displacements (figure 1, Step 4).

- (g) Creation of 4D MRI/ 4D CT sequences: By repeating the previous step across the full sequence of deformed NURBS surfaces, we synthesize patient-specific motion that is agnostic to any imaging modality. Applying the corresponding DVFs to the original patient scans produces a 4D sequence of

Algorithm 1: NURBS Deformation Algorithm per PhaseNURBS model Pw **Input:** Dispersion parameters $(\alpha_{space}, \alpha_{time})$ Wave equation parameters (A, c, λ) **Output:** Deformed NURBS surface

```

1 for  $p \leftarrow$  each phase do
2    $m \leftarrow$  total sectional curves in NURBS
3   for  $i \leftarrow 0$  to  $m - 1$  do
4      $n \leftarrow$  total control points in  $m$ 
5      $disp_{space} \leftarrow e^{-\alpha_{space} \times s}$  // Spatial dispersion
6      $disp_{time} \leftarrow e^{-\alpha_{time} \times t}$  // Temporal dispersion
7      $dispersion \leftarrow disp_{space} \cdot disp_{time}$ 
8     for  $j \leftarrow 0$  to  $n - 1$  do
9        $L_i = i \times \frac{organ\_length}{m}$ 
10       $wave = (\frac{1}{\sqrt{3}}) \times A(2\pi \frac{L_i \times -c \times t}{\lambda})$ 
11       $\Delta \leftarrow wave \times dispersion;$ 
12       $ux, uy, uz \leftarrow$  Compute unit vector
13       $Pw[j][i].x \leftarrow Pw[j][i].x + \Delta \cdot ux$  // Update control point x
14       $Pw[j][i].y \leftarrow Pw[j][i].y + \Delta \cdot uy$  // Update control point y
15       $Pw[j][i].z \leftarrow Pw[j][i].z + \Delta \cdot uz$  // Update control point z

```

Figure 2. Pseudocode illustrating the implementation of the motion model applied to the NURBS surfaces of the stomach and large bowel, as detailed in the methods section part (e).

GI motion while preserving the internal contents of the luminal organs (e.g. air, fluid). Sample videos of synthesized motion sequences are included in supplemental material for representative cases.

For additional details on steps (a) through (d), please refer to the pseudocode provided in the appendix document. Additionally, the code is available upon request to the authors.

2.2. Datasets and motion parameters

Under institutional review board–approved protocols (IRB 21-129 and HUM00260112), all retrospective patient data were analyzed in accordance with ethical standards. In total, 11 datasets were analyzed, comprising six T2w MRI, two T1w gaSOS, and three CECT scans. Additionally, three independent T1w gaSOS datasets were used solely for quality assurance, with motion models tailored to the motion patterns observed in those specific cases. To the stomach, we applied a traveling wave equation with parameters $A = 16$ mm, $\lambda = 55$ mm, and $c = 5$ mm s^{−1}. In addition, for the five datasets that included the large bowel, a second traveling wave equation $A = 16$ mm, $\lambda = 40$ mm, $c = 8$ mm s^{−1} was applied to simulate large bowel motion. Large bowel motion was not analyzed for a subset of patients either because of a smaller field of view, as in the cases of patients with liver cancer (patients 1 and 2), or due to insufficient scan lengths that did not capture the full extent of the large bowel (patients 3–6), which prevented the reliable generation of corresponding NURBS surfaces. For all synthesized motions used in the DIR evaluation, $\alpha = 0$ was chosen to emulate high-magnitude, non-dispersive GI motion. The amplitude and wavelength parameters were chosen to align with the peristaltic motion measured from 10 patients treated on a 1.5 T MR-linac and reported by Subashi *et al* (2023).

2.3. Patient-specific motion quality assurance

To evaluate motion realism, we compared synthetic stomach motion from the DT framework to motion derived from three T1w gaSOS MRI datasets with known stomach contractions. Specifically, we used the previously published known stomach deformation motions, which includes detailed recordings of stomach contractions over 21 distinct phases. Previous studies have used hierarchical MRI reconstructions to model stomach contractions and slow drifts with and without respiratory motion (Zhang *et al* 2021). Employing the ground truth DVFs derived from three scans with hierarchical motion models as baseline references, we fit traveling sinusoidal waves tailored specifically to each contraction phase. This allowed the DT pipeline to generate synthetic stomach motions closely aligned with the experimentally measured motions. To validate the synthesized motion, we computed the mean and maximum displacements of the ground-truth DVFs and the log of the Jacobian which captures the degree of local non-rigid, non-affine deformations.

2.4. Motion analysis

For all 21 synthetically generated motion phases across the 11 patient scans, the maximum and mean displacement magnitudes of the DVFs generated by the pipeline were computed for the full body as well as two gastric organs: the stomach and large bowel.

2.5. Evaluation of DIR methods using DTs

The goal of the experiment was to assess the feasibility of using the DTs to evaluate a range of DIR methods using the synthesized GI motions produced with various imaging modalities for multiple patients. Of note, the goal was not to compare individual methods with respect to one another; all methods were used as is without performing any hyperparameter or other optimization for the evaluated DT datasets.

Six different DIR algorithms were assessed including 4 variational and 2 DL methods. The variational methods included Horn–Schunck optical flow (HSOF), EVolution multimodality (EVO), Elastix with mutual information, and the Iterative Demons algorithm, all of which used iterative optimization using image intensities. Key differences between the methods include the use of brightness consistency and smoothness regularization for motion fields used in HSOF (Horn and Schunck 1981), normalized intensity gradients with smoothness regularization used in EVO (Denis de Senneville *et al* 2016), B-spline parameterization of the transformation used in Elastix (Klein *et al* 2010, Leibfarth *et al* 2013) and optical flow with Gaussian smoothing in diffeomorphic Demons (Thirion 1998, Vercauteren *et al* 2009). The two DL DIR methods were VoxelMorph (Balakrishnan *et al* 2019) and an enhancement of VoxelMorph called progressively refined registration-segmentation (ProRSeg) that uses convolutional long short-term memory networks in the encoder (Jiang *et al* 2023). Testing datasets used for evaluation were never used for training VoxelMorph and ProRSeg to avoid data leaks. Both VoxelMorph and ProRSeg were trained on different sets of real patient T2w datasets used in Jiang *et al* (2023) and were only applied to DTs corresponding to the same MR scanning sequence.

2.6. Evaluation metrics

Registration was performed between the original scan and the phase exhibiting the maximum deformation. Geometric accuracy was computed by measuring organ segmentation accuracy using the DSC and Hausdorff distance at 95th percentile as well as TRE using the surface and inner shell points of a given NURBS surface. Differences in dose accumulation across methods were assessed by simulating dose accumulation using both the ground truth DVF and the DVFs produced by each registration method. Dose deformation was performed using direct dose mapping. DWE comparing the accumulated dose distributions was calculated as:

$$DWE = \sum_{i=1}^N \frac{\text{Accum } D_{i,DIR} - \text{Accum } D_{i,GT}}{\text{Accum } D_{i,GT}}.$$

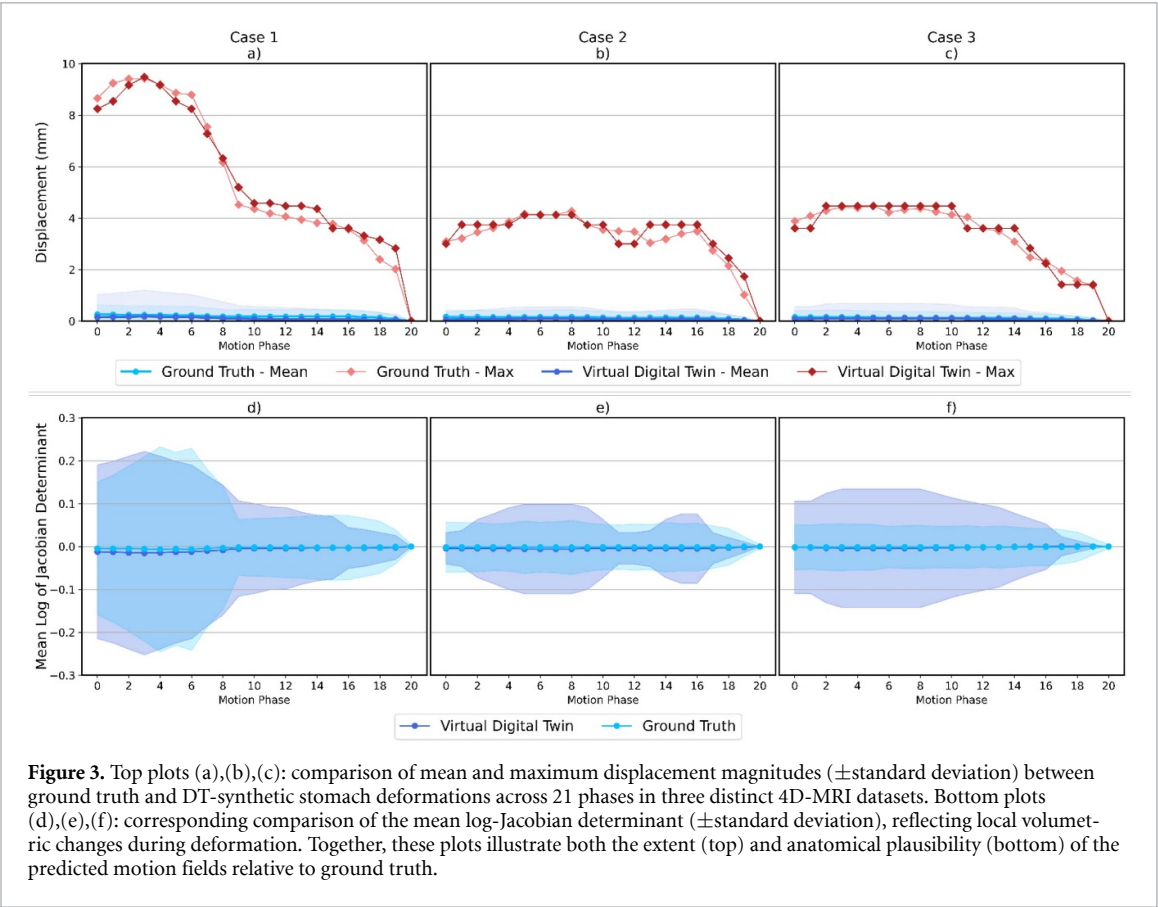
Granular assessment of dosimetric and displacement errors were computed at the anatomic voxel level and quantified using root mean squared error (RMSE), which then were also visualized as heat-maps for a more detailed evaluation of model accuracy in regions subject to large deformations or high-radiation dose exposure.

3. Results

Our semi-automated pipeline was applied to generate motion for 11 different patients from various anatomic imaging modalities. The framework takes 5–15 min to synthesize peristaltic motion from a 3D

Table 1. Details of the image volumes and voxel dimensions used for motion synthesis across 11 patient datasets.

Patient	MR sequence details	Volume size (voxels)	Voxel size (mm)
1	T1w gaSOS	224 × 224 × 96	2.188 × 2.188 × 3.5
2	T1w gaSOS	224 × 224 × 96	2.188 × 2.188 × 3.5
3	CECT	512 × 512 × 201	0.9766 × 0.9766 × 2
4	CECT	512 × 512 × 201	0.9766 × 0.9766 × 2
5	CECT	512 × 512 × 201	0.9766 × 0.9766 × 2
6	T2w MRI	512 × 512 × 50	0.7813 × 0.7813 × 5
7	T2w MRI	448 × 448 × 125	1 × 1 × 2
8	T2w MRI	448 × 448 × 125	1 × 1 × 2
9	T2w MRI	448 × 448 × 125	1 × 1 × 2
10	T2w MRI	448 × 448 × 125	1 × 1 × 2
11	T2w MRI	448 × 448 × 125	1 × 1 × 2



input image, given fixed motion parameters, NURBS surfaces, with the synthesis time varying due to the number of set motion phases. The framework allows to interactively change the motion parameter configurations as well as NURBS surfaces to also visually match known peristaltic patterns. The framework was capable of generating motion for image volumes and voxels with varying sizes, as shown in table 1. The method was applicable to axial and coronal image reconstructions.

3.1. Assessing patient-specific motion quality assurance

The top row of figure 3 shows a comparison between DT-generated and ‘ground truth’ stomach motions derived from Zhang *et al* using hierarchical motion modeling (2021), showing both mean and maximum displacement magnitudes. Across all motion phases, the DT-generated motion exhibited mean and maximum displacement differences within 0.8 mm of ground truth. Furthermore, as shown in the bottom row of figure 3, the DT-simulated deformations yielded consistent log-Jacobian means across all motion phases, remaining within 0.01 of the ground truth mean over all 21 phases. This indicates strong agreement in the overall deformation characteristics.

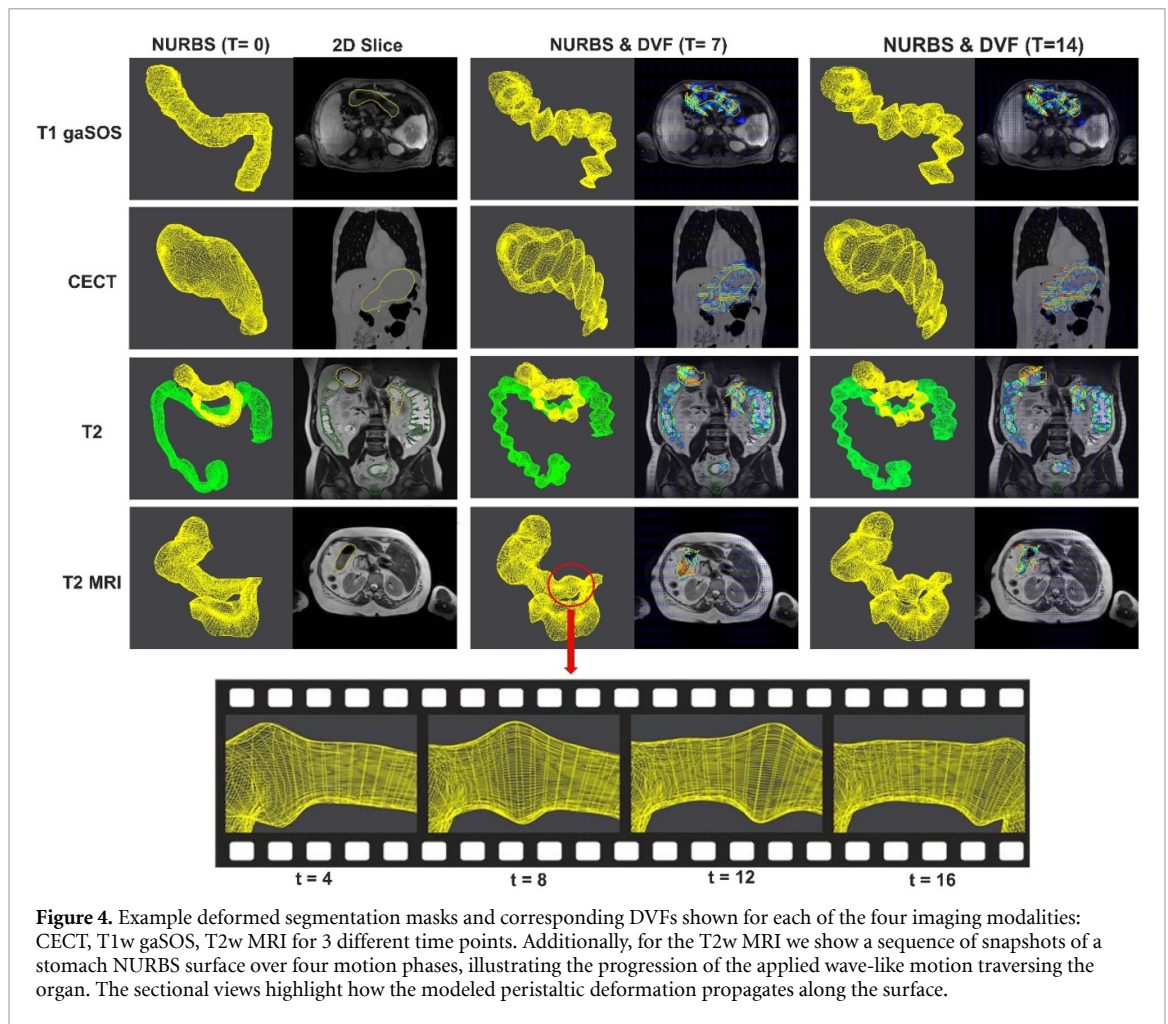


Figure 4. Example deformed segmentation masks and corresponding DVFs shown for each of the four imaging modalities: CECT, T1w gaSOS, T2w MRI for 3 different time points. Additionally, for the T2w MRI we show a sequence of snapshots of a stomach NURBS surface over four motion phases, illustrating the progression of the applied wave-like motion traversing the organ. The sectional views highlight how the modeled peristaltic deformation propagates along the surface.

Table 2. Maximum motion magnitudes for 2 gastric organs (stomach and large bowel) computed from the DVFs generated across the 21 phases of motion for all 11 patients with synthesized motion. Patients 1–2 lacked large bowel data due to a limited field of view, and patients 3–6 were excluded because of inaccurate large bowel segmentations that prevented reliable NURBS surface generation.

Patient	Full body	Motion magnitude (mm)	
		Stomach	Large bowel
1	13.2	13.2	—
2	14.2	13.9	—
3	13.2	13.2	—
4	13.5	13.1	—
5	13.5	13.4	—
6	14.3	14.3	—
7	8.7	8.7	8.6
8	8.6	6.6	8.6
9	8.6	8.6	8.6
10	8.7	8.7	7.7
11	8.6	8.6	8.6

3.2. Modality agnostic temporally varying motion synthesis

Figure 4 depicts the synthesized motions with the corresponding DVFs for all three analyzed imaging modalities, CECT, T1w gaSOS, and T2w MRI. The deformed masks and 2D image slices taken at different time points along the traveling wave sequence propagating through the stomach are also shown. The maximum stomach motion shown in table 2 ranged from 8.56 mm to 14.34 mm, while the large bowel exhibited motion magnitudes between 7.69 mm and 8.64 mm, indicating the ability to simulate motion at varying amplitudes. The mean and standard deviation in the motion magnitudes for the gastric organs and the whole body are summarized in table 3, which showed stomach mean displacements ranging from 0.92 mm to 3.69 mm, and the large bowel from 0.77 mm to 2.96 mm.

Table 3. Means and standard deviations of motion magnitudes for 2 gastric organs (stomach and large bowel) computed from the DVFs generated across the 21 phases of motion for all 11 patients with synthesized motion. Patients 1–2 lacked large bowel data due to a limited field of view, and patients 3–6 were excluded because of inaccurate large bowel segmentations that prevented reliable NURBS surface generation.

Patient	Motion magnitude (mm)		
	Full body	Stomach	Large bowel
1	0.04 ± 0.01	3.05 ± 0.83	—
2	0.03 ± 0.01	2.85 ± 0.9	—
3	0.04 ± 0.02	2.22 ± 0.72	—
4	0.01 ± 0.01	2.32 ± 0.65	—
5	0.06 ± 0.02	3.27 ± 0.96	—
6	0.04 ± 0.02	3.24 ± 0.82	—
7	0.06 ± 0.00	0.92 ± 0.03	0.77 ± 0.03
8	0.08 ± 0.00	2.57 ± 0.01	2.74 ± 0.04
9	0.08 ± 0.00	3.65 ± 0.05	2.96 ± 0.01
10	0.07 ± 0.00	3.57 ± 0.03	2.78 ± 0.05
11	0.09 ± 0.00	3.69 ± 0.01	2.96 ± 0.03

Table 4. Segmentation accuracy of various variational registration methods applied to T1wI gaSOS, and CECT. Stomach & Duo: duodenum was included in the segmentation mask of the stomach.

Metric	DIR Algorithm	T1w gaSOS		CECT		
		Patient 1 Stomach & Duo	Patient 2 Stomach & Duo	Patient 3 Stomach	Patient 4 Stomach	Patient 5 Stomach
TRE ↓ (mm)	HSOF	2.14 ± 1.31	2.70 ± 1.94	3.06 ± 1.62	3.33 ± 1.93	3.11 ± 2.01
	EVO	3.02 ± 1.70	3.64 ± 2.12	3.88 ± 2.30	3.33 ± 1.90	3.44 ± 2.27
	Elastix	3.40 ± 2.19	2.58 ± 1.45	2.48 ± 1.33	3.48 ± 2.12	3.48 ± 2.41
	Demons	3.26 ± 1.93	3.08 ± 1.94	3.27 ± 1.68	3.60 ± 2.00	3.46 ± 2.05
DSC ↑	HSOF	0.81	0.78	0.99	0.79	0.77
	EVO	0.77	0.81	0.98	0.82	0.81
	Elastix	0.68	0.77	0.98	0.76	0.71
	Demons	0.74	0.8	0.99	0.79	0.77
HD95 ↓ (mm)	HSOF	6.00	7.00	1.00	7.48	8.31
	EVO	8.06	8.77	1.73	6.71	8.06
	Elastix	8.06	7.35	2.24	7.18	8.66
	Demons	8.06	6.16	1.41	7.34	8.83

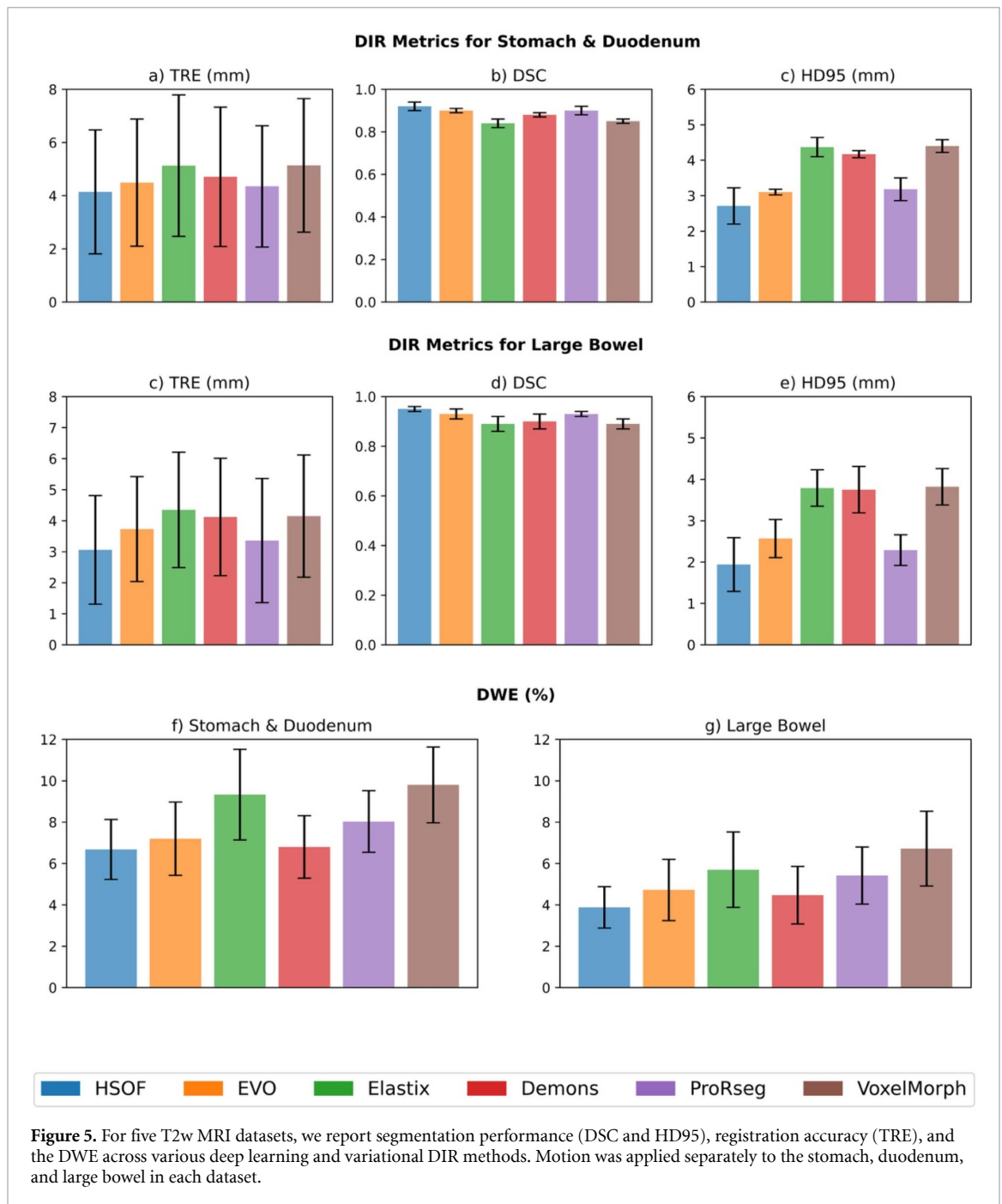
3.3. Evaluating DIR performance using the DT framework

DIR was computed using various methods between the phase 0 (static input patient 3D scan) and the phase with the largest amplitude displacement with respect to phase 0. Variational DIR methods were applied to the CECT, T1w gaSOS, and T2w MRI, whereas the DL-DIRs trained with T2w MRI were only applied to T2w MRI datasets.

Table 4 shows the accuracy metrics computed for the variational methods applied to the CECT and T1w gaSOS scans. The DIR methods were similarly accurate for T1w gaSOS compared to CECT using TRE (2.14 ± 1.31 mm to 3.64 ± 2.12 mm versus 2.48 ± 1.33 mm to 3.60 ± 2.00 mm), HD95 (6.00 mm to 8.77 mm versus 1 mm to 8.83 mm), and DSC (0.68–0.81 versus 0.71–0.99). HD95 and DSC were computed for two CECT patients and three T1w gaSOS cases; however, standard deviation values are not reported due to the limited sample size.

In the case of T2w MRI from 5 patients with pancreatic cancer, the DIR methods showed slightly higher TRE 4.14 ± 2.23 mm to 5.14 ± 2.51 mm for the stomach and duodenum compared to T1wI gaSOS and CECT images, a higher DSC from 0.84 ± 0.02 – 0.92 ± 0.02 , and lower HD95 ranging from 2.71 ± 0.51 mm to 4.4 ± 0.18 mm. The registration accuracies for the large bowel were slightly lower than the stomach and duodenum with TRE ranging from 3.06 ± 0.48 mm to 4.35 ± 0.29 mm, DSC between 0.89 ± 0.02 – 0.95 ± 0.01 , and HD95 between 1.94 ± 0.65 mm and 3.82 ± 0.44 mm, respectively.

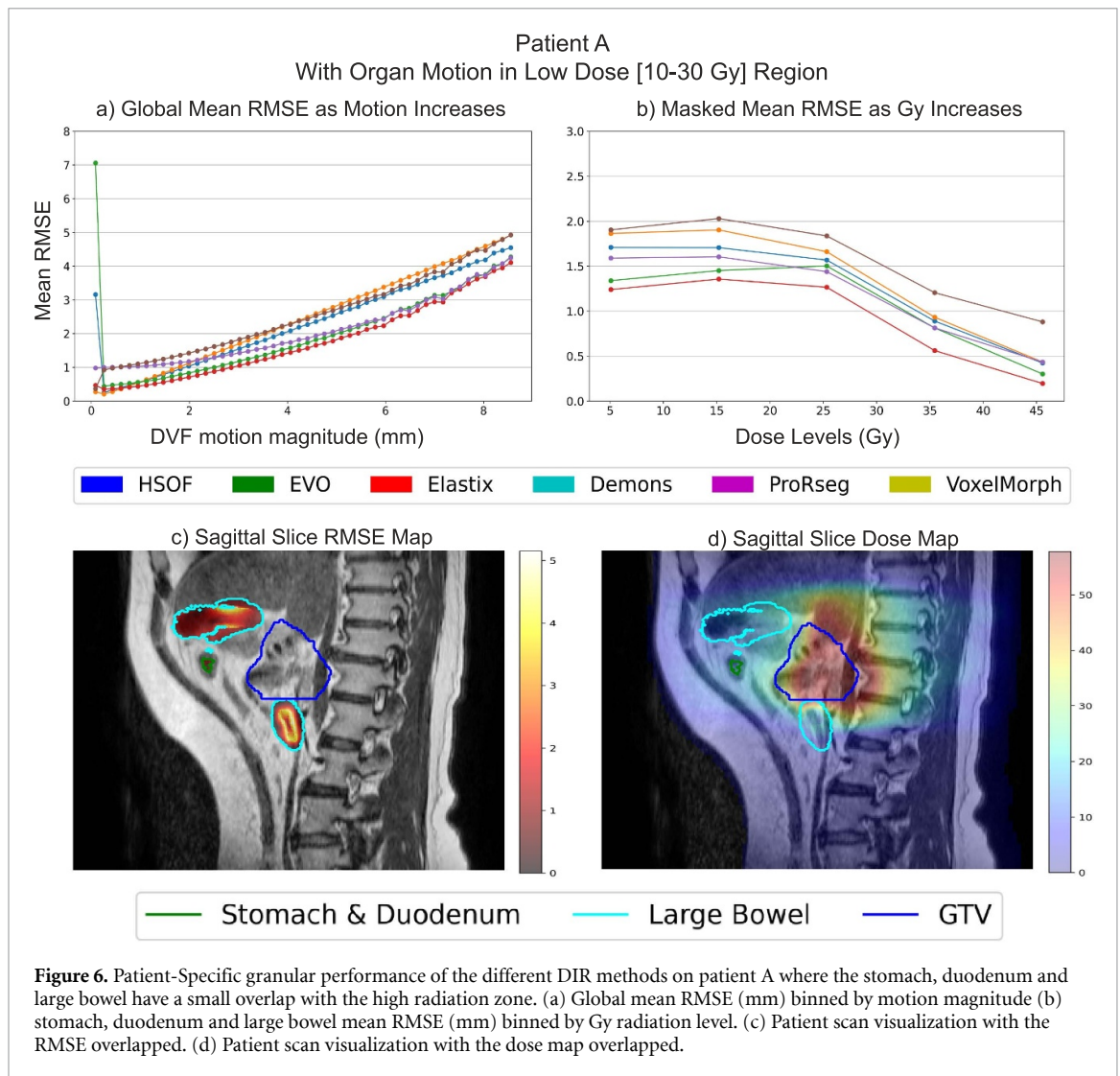
In addition, the same 5 patients also had radiation treatment dose maps, which were used to calculate the DWE for the same three organs. As shown in figure 5, the mean DWE ranged from $6.68 \pm 1.45\%$ to $9.80 \pm 1.83\%$ for stomach and duodenum and a smaller error of $3.88 \pm 1.00\%$ to $6.72 \pm 1.81\%$ for the large bowel. The results also show variation in the performance of the different



methods across the different datasets and modalities, thus providing an approach to assess the relative merits of the various methods for segmentation, registration, and dose warping on an individual patient level.

3.4. Patient-specific granular errors

Capability to assess granular (or voxel-level) error visualization on a patient level is demonstrated for two representative patients, patient A (figure 6) where organs undergoing motion occurred in the low-radiation dose regions (10–30 Gy) (see figure 6(d)) and patient B (figure 7) with the same organs located in the high radiation dose regions (exceeding 30 Gy) (see figure 7(d)). Global RMSE is visualized with respect to increasing motion magnitudes to assess errors as a function of motion (figures 6(a) and 7(a)). Motion magnitude was binned at fixed intervals ranging from a minimum of 0 mm to a maximum of 8.65 mm. Patient specific analysis showed that the mean RMSE ranged from 0.7 mm in low-motion regions (0–1 mm) to 5 mm in high-motion regions (>8 mm) for the two patients. RMSE was also computed within the GI organs undergoing motion to assess impact of motion on the accuracy with respect to the radiation dose delivered to the organ, which showed a range from 0.10 mm to



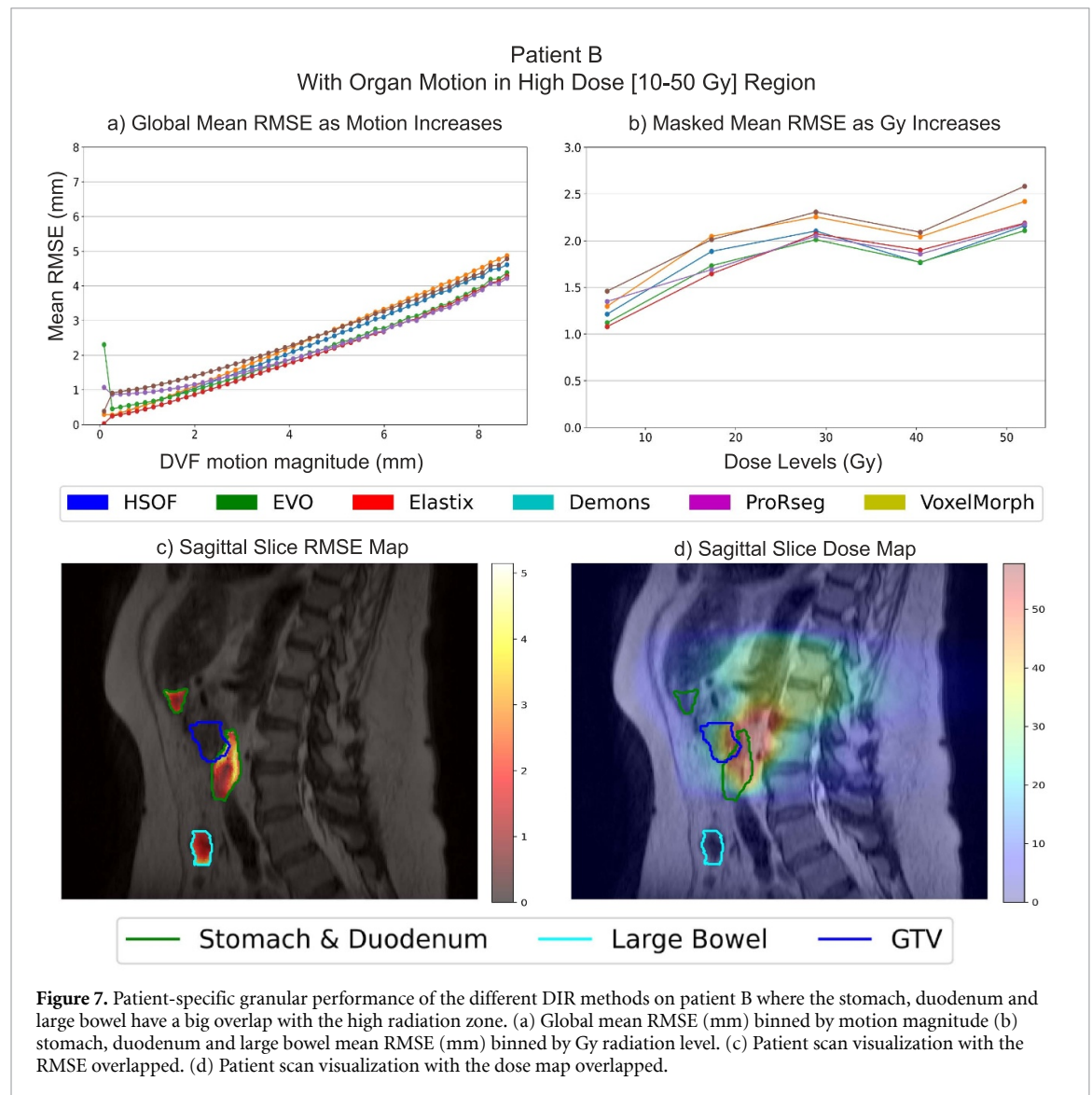
2.00 mm, with the highest errors occurring mostly in low-dose regions for Patient A, but a higher error of 1.00 mm to 2.5 mm in the high-dose region for Patient B (figures 6(b) and 7(b)).

Figures 6(c), 7(c), 6(d) and 7(d) show a visualization of voxel-wise RMSE within the organs undergoing motion and the dose maps, respectively to provide a visual representation of the errors for individual patients.

4. Discussion

In this work, we extended the concept of population-level digital phantoms to develop patient-specific DTs that model temporally varying GI motion. Our semi-automated pipeline starts from AI automated organ segmentations, which then are used to generate peristaltic motion of varying amplitudes and time scales for luminal organs such as the stomach, duodenum and the large bowel. Our analysis indicates that the synthesized stomach motions can closely match known reference values, both in displacement magnitude and deformation characteristics producing motion magnitudes consistent with those reported by Zhang *et al* (2021). Thus, the framework showed initial feasibility to generate realistic motion for three different GI organs such as stomach, duodenum, and large bowel, with differing geometries across multiple common radiological imaging modalities,

Our framework is applicable to multiple anatomic imaging modalities and demonstrated feasibility to evaluate multiple variational and two different DL DIR methods. In addition to evaluating registration, our framework can be easily extended to evaluate dose warping accuracy summarizing errors for individual organs as well as on a granular level to assess accuracy variations on voxel-level. As a result, our DT framework enables voxel-wise visualization of registration errors, facilitates analysis of error patterns



across motion regimes (e.g. low vs high motion), and supports individualized assessment of how registration inaccuracies can affect radiation dose. To our knowledge, this is the first comprehensive simulation and evaluation of DIR using patient-specific DTs across multiple imaging modalities for GI luminal organs.

Whereas, previous efforts focused on modeling respiratory and cardiac motion, ours focused on modeling the digestive motion (Segars *et al* 1999, 2013, 2018, 2019). One prior work by Subashi *et al* demonstrated the ability to generate MR-like digital phantoms incorporating a range of GI motion types including peristalsis, slow and fast gastric contractions, and HAPCs (Subashi *et al* 2023). However, all aforementioned prior works synthesized digital phantoms modeling population-level anatomy built on the generic adult male and female XCAT models. A limitation with modeling population-level motion is that it does not represent individual patient anatomy variations. Our work, for the first time, addresses the key issue of modeling patient-specific variations by creating patient-specific DTs of gastric motion.

Another limitation of population level modeling using XCAT requires simulation of MR images using fixed signal intensities for each organ, that can create a domain shift for assessing DL registration methods. In our work, synthesis starts from the original MRI, the synthesized motions are also created on MRI, hence allowing to evaluate DL DIR methods.

Finally, GI organs undergo substantial and arbitrary motion that varies from patient to patient despite common motion mitigation strategies such as pneumatic compression belts (Mostafaei *et al* 2018, Liu *et al* 2021, Zhang *et al* 2021, Alam *et al* 2022). Our approach allows to vary and create a variety of motion amplitudes and rigorously evaluate DIR methods under various GI motion amplitudes.

Limitations of our current framework include the lack of support for modeling respiratory motion as the focus of this work was isolated GI motion. We also excluded small bowel motion simulation

because it is generally segmented in the clinic as a ‘bowel-bag’, making the extraction of NURBS to model motion along the tubular region difficult. Nevertheless, our approach can potentially be extended to additional organs and other types of motions involving such organs. Finally, GPU acceleration is not currently implemented, and manual steps such as parameter selection increase computation time. This design prioritizes flexibility, allowing the generation of motions with user-defined amplitudes. While the current pipeline is intended for offline analysis, incorporating GPU acceleration and predefined parameters could reduce processing time and enable near real-time patient-specific DT generation.

5. Conclusion

We developed a semi-automated DT pipeline to generate realistic GI temporally varying motion in the stomach, duodenum and large bowel from multiple anatomic imaging modalities. Our framework showed capability to generate motions within ranges seen in real patients, indicating feasibility to evaluate multiple DIR methods. Our framework enables evaluating dose warping and registration errors in a granular voxel-wise manner for individualized patient-level analysis, suitable for rigorous analysis required for clinical deployment.

Data Availability Statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.5281/zenodo.17702490> (Tapias Gomez *et al* 2025).

Appendix available at <https://doi.org/10.1088/1361-6560/ae2b46/data1>.

T2 MRI Synthetic Motion available at <https://doi.org/10.1088/1361-6560/ae2b46/data2>.

T2 MRI Synthetic Motion w/ DVF available at <https://doi.org/10.1088/1361-6560/ae2b46/data3>.

Acknowledgments

This research was supported by NIBIB R01EB032825 and partially supported by the NIH/NCI Cancer Center Support Grant/Core Grant (P30 CA008748), MSK IMRAS and Cornell Tech.

ORCID iDs

Jorge Tapias Gomez  0009-0008-8692-1105

Lando S Bosma  0000-0002-8206-271X

Jue Jiang  0000-0001-9642-3515

Ergys D Subashi  0000-0001-5168-6928

William P Segars  0000-0003-3687-5733

James M Balter  0000-0003-1513-4936

Mert R Sabuncu  0000-0002-7068-719X

Neelam Tyagi  0000-0002-6256-7817

Harini Veeraraghavan  0000-0002-4054-5529

References

- Alam S, Veeraraghavan H, Tringale K, Amooteng E, Subashi E, Wu A J, Crane C H and Tyagi N 2022 Inter- and intrafraction motion assessment and accumulated dose quantification of upper gastrointestinal organs during magnetic resonance-guided ablative radiation therapy of pancreas patients *Phys. Imaging Radiat. Oncol.* **21** 54–61
- Balakrishnan G, Zhao A, Sabuncu M R, Gutttag J and Dalca A V 2019 VoxelMorph: a learning framework for deformable medical image registration *IEEE Trans. Med. Imaging* **38** 1788–800
- Bosca R J and Jackson E F 2016 Creating an anthropomorphic digital MR phantom—an extensible tool for comparing and evaluating quantitative imaging algorithms *Phys. Med. Biol.* **61** 974
- Bosma L S, Hussein M, Jameson M G, Asghar S, Brock K K, McClelland J R, Poeta S, Yuen J, Zachiu C and Yeo A U 2024 Tools and recommendations for commissioning and quality assurance of deformable image registration in radiotherapy *Phys. Imaging Radiat. Oncol.* **32** 100647
- Bosma L S, Zachiu C, Ries M, Denis de Senneville B and Raaymakers B W 2021 Quantitative investigation of dose accumulation errors from intra-fraction motion in MRgRT for prostate cancer *Phys. Med. Biol.* **66** 065002
- Brock K K, Mutic S, McNutt T R, Li H and Kessler M L 2017 Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy committee task group No. 132 *Med. Phys.* **44** e43–e76
- Caon M 2004 Voxelbased computational models of real human anatomy: a review *Radiat. Environ. Biophys.* **42** 229–35
- Cassola V F, Milian F M, Kramer R, de Oliveira Lira C A B and Khoury H J 2011 Standing adult human phantoms based on 10th, 50th and 90th mass and height percentiles of male and female Caucasian populations *Phys. Med. Biol.* **56** 3749
- Denis de Senneville B, Zachiu C, Ries M and Moonen C 2016 EVolution: an edge-based variational method for non-rigid multi-modal image registration *Phys. Med. Biol.* **61** 7377

- Digital Europe program of the European Commission 2022 *EDITH: European Virtual Human Twin* (available at: www.edith-csa.eu/) (Accessed 1 June 2025)
- Ding A, Mille M M, Liu T, Caracappa P F and Xu X G 2012 Extension of RPI-adult male and female computational phantoms to obese patients and a Monte Carlo study of the effect on CT imaging dose *Phys. Med. Biol.* **57** 2441
- Fallone B G 2014 The rotating biplanar linac–magnetic resonance imaging system *Magn. Reson. Imaging Radiat. Oncol.* **24** 200–2
- Horn B K P and Schunck B G 1981 Determining optical flow *Artif. Intell.* **17** 185–203
- Jiang J, Hong J, Tringale K, Reyngold M, Crane C, Tyagi N and Veeraraghavan H 2023 Progressively refined deep joint registration segmentation (ProRSeg) of gastrointestinal organs at risk: application to MRI and conebeam CT *Med. Phys.* **50** 4758–74
- Jiang J et al 2022 Self-supervised 3D anatomy segmentation using Self-distilled Masked Image Transformer (SMIT) *Medical Image Computing and Computer Assisted Intervention—MICCAI 2022* ed L Wang et al (Springer) pp 556–66
- Johansson A, Balter J M and Cao Y 2021 Gastrointestinal 4D MRI with respiratory motion correction *Med. Phys.* **48** 2521–7
- Johnson P B, Whalen S R, Wayson M, Juneja B, Lee C and Bolch W E 2009 Hybrid patient-dependent phantoms covering statistical distributions of body morphometry in the U.S. adult and pediatric population *Proc. IEEE* **97** 2060–75
- Katsoulakis E et al 2024 Digital twins for health: a scoping review *npj Digit. Med.* **7** 77
- Keall P J, Barton M and Crozier S 2014 The Australian magnetic resonance imaging–linac program *Magn. Reson. Imaging Radiat. Oncol.* **24** 203–6
- Klein S, Staring M, Murphy K, Viergever M A and Pluim J 2010 Elastix: a toolbox for intensity-based medical image registration *IEEE Trans. Med. Imaging* **29** 196–205
- Legendijk J J W, Raaymakers B W and Vulpen V 2014 The magnetic resonance imaging–linac system *Magn. Reson. Imaging Radiat. Oncol.* **24** 207–9
- Lee C Y 1961 An algorithm for path connections and its applications *IRE Trans. Electron. Comput.* **EC10** 346–65
- Lee T C, Kashyap R L and Chu C N 1994 Building skeleton models via 3D medial surface axis thinning algorithms *CVGIP, Graph. Models Image Process.* **56** 462–78
- Leibfarth S, Mönnich D, Welz S, Siegel C, Schwenzer N, Schmidt H, Zips D and Thorwarth D 2013 A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning *Acta Oncol.* **52** 1353–9
- Liu L, Johansson A, Cao Y, Kashani R, Lawrence T S and Balter J M 2021 Modeling intra-fractional abdominal configuration changes using breathing motion-corrected radial MRI *Phys. Med. Biol.* **66** 085002
- Lowther N, Louwe R, Yuen J, Hardcastle N, Yeo A and Jameson M 2022 MIRSIG position paper: the use of image registration and fusion algorithms in radiotherapy *Phys. Eng. Sci. Med.* **45** 421–8
- Meyer S, Hu Y-C, Rimner A, Mechalakos J, Cerviño L and Zhang P 2025 Deformable image registration uncertainty-encompassing dose accumulation for adaptive radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **122** 818–26
- Mostafaei F, Tai A, Omari E, Song Y, Christian J, Paulson E, Hall W, Erickson B and Li X A 2018 Variations of MRI assessed peristaltic motions during radiation therapy *PLoS One* **13** e0205917
- Mutic S and Dempsey J F 2014 The ViewRay system: magnetic resonance–guided and controlled radiotherapy *Magn. Reson. Imaging Radiat. Oncol.* **24** 196–9
- Papachristou K, Katsakiori P F, Papadimitroulas P, Strigari L and Kagadis G C 2024 Digital twins’ advancements and applications in healthcare, towards precision medicine *J. Pers. Med.* **14** 1101
- Rigaud B et al 2019 Deformable image registration for radiation therapy: principle, methods, applications and evaluation *Acta Oncol.* **58** 1225–37
- Segars W P et al 2013 Population of anatomically variable 4D XCAT adult phantoms for imaging research and optimization *Med. Phys.* **40** 043701
- Segars W P, Lalush D S and Tsui W 1999 A realistic spline-based dynamic heart phantom *IEEE Trans. Nucl. Sci.* **46** 503–6
- Segars W P, Tsui B M W, Cai J, Yin F-F, Fung G S K and Samei E 2018 Application of the 4-D XCAT phantoms in biomedical imaging and beyond *IEEE Trans. Med. Imaging* **37** 680–92
- Segars W P, Veress A I, Sturgeon G M and Samei E 2019 Incorporation of the living heart model into the 4D XCAT phantom for cardiac imaging research *IEEE Trans. Radiat. Plasma Med. Sci.* **3** 54–60
- Subashi E, Segars P, Veeraraghavan H, Deasy J and Tyagi N 2023 A model for gastrointestinal tract motility in a 4D imaging phantom of human anatomy *Med. Phys.* **50** 3066–75
- Tapias Gomez J, Nadkarni N, Bosma L, Jiang J, Subashi E, Segars P, Balter J, Sabuncu M, Tyagi N and Veeraraghavan H 2025 Patient-specific digital twin gastrointestinal motion dataset (available at: <https://doi.org/10.5281/zenodo.17702490>)
- Thirion J-P 1998 Image matching as a diffusion process: an analogy with Maxwell’s Demons *Med. Image Anal.* **2** 243–60
- Vercauteren T et al 2009 Diffeomorphic demons: efficient nonparametric image registration *Math. Brain Imaging* **45** S61–S72
- Zhang Y, Kashani R, Cao Y, Lawrence T S, Johansson A and Balter J M 2021 A hierarchical model of abdominal configuration changes extracted from golden angle radial magnetic resonance imaging *Phys. Med. Biol.* **66** 045018