Physics in Medicine &
Biology

IPEM
Institute of Physics and Engineering in Medicine

PAPER • OPEN ACCESS

Evaluation of volumetric breathing motion prediction of the stomach using dynamic golden-angle radial MRI

To cite this article: Robert Jones *et al* 2026 *Phys. Med. Biol.* **71** 175017

View the article online for updates and enhancements.

You may also like

- Volumetric prediction of breathing and slow drifting motion in the abdomen using radial MRI and multi-temporal resolution modeling
Lianli Liu, Adam Johansson, Yue Cao et al.
- Real-time prediction of stomach motions based upon gastric contraction and breathing models
Yuhang Zhang, Yue Cao, Rojano Kashani et al.
- Real-time prediction of respiratory motion using a cascade structure of an extended Kalman filter and support vector regression
S-M Hong and W Bukhari

This content was downloaded from IP address 141.213.15.172 on 02/09/2026 at 19:11

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

Check for updates

OPEN ACCESS

RECEIVED
13 March 2026

REVISED
9 July 2026

ACCEPTED FOR PUBLICATION
27 July 2026

PUBLISHED
2 September 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.

PAPER

Evaluation of volumetric breathing motion prediction of the stomach using dynamic golden-angle radial MRI

Robert Jones[1,2,*], **Lianli Liu**[3], **Daekeun You**[1], **Jeffrey A Fessler**[2] and **James M Balter**[1]

[1] Department of Radiation Oncology, University of Michigan, Ann Arbor, MI 48109, United States of America
[2] Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109, United States of America
[3] Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA 94305, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: rjjones@umich.edu

Keywords: MRI-guided radiotherapy, multi-temporal resolution imaging, breathing motion prediction, golden-angle radial MRI, abdominal organ motion, dynamic MRI

Supplementary material for this article is available online

Abstract

Objective. The motion of abdominal organs complicates accurate radiotherapy (RT) planning and delivery. Accurate short-term prediction of respiratory motion is essential for image-guided adaptive RT in the abdomen. This study evaluates the performance of a kernel ridge regression-based breathing motion prediction framework (KRR-BMP) for abdominal organs, with particular emphasis on the stomach and its centerline. *Approach*. Dynamic MRI datasets from 27 scans acquired in 16 patients were processed to generate low-rank breathing motion models based on principal component analysis (PCA) of deformation vector fields (DVFs). KRR-BMP was used to predict future PCA coefficients of breathing motion, enabling reconstruction of volumetric breathing motion deformation fields at specified prediction horizons. The accuracy of predicted breathing DVFs was evaluated across multiple organs and anatomical structures, training set sizes, PCA ranks, and prediction horizons. *Main results*. At a 340 ms prediction horizon, KRR-BMP achieved sub-voxel accuracy across all evaluated organs, with population-mean endpoint errors below 1.3 mm and 95th percentile errors consistently below 3 mm. Error distributions were relatively narrow across subjects and training sets, demonstrating robustness and generalizability. Reductions in training data led to predictable increases in error, particularly for gastric structures, although population-level upper bound errors remained within clinically relevant ranges. A PCA rank of two provided the best balance between accuracy and robustness, while higher ranks increased inter-subject variability. *Significance*. KRR-BMP provides accurate and robust short-term respiratory motion prediction for abdominal organs in MR-guided RT. The demonstrated performance for stomach subregions establishes a foundation for future integration of respiratory motion prediction with real-time gastric motility modeling and motion-aware adaptive treatment workflows.

1. Introduction

The geometry and local configurations of abdominal organs continuously change due to physiological processes such as breathing, antral contractions, and gastrointestinal (GI) peristalsis. The resulting organ motion produces complex displacements and deformations of abdominal organs and tumors, complicating accurate radiotherapy (RT) planning and delivery (Keall *et al* 2006, Wysocka *et al* 2010). Intra-fraction motion is a major source of geometric uncertainties in RT (Fast *et al* 2024), and if disregarded can lead to under-dosage of target volumes or unintended irradiation of nearby organs-at-risk (OARs) (Balter *et al* 1996, Jayachandran *et al* 2010). Effective management of intra-fractional motion is therefore crucial to fully realize the benefits of precision RT (Keall *et al* 2006, Abbas *et al* 2014, Molitoris *et al* 2018, Burton *et al* 2022, Fast *et al* 2024).

© 2026 The Author(s). Published on behalf of Institute of Physics and Engineering in Medicine by IOP Publishing Ltd

Magnetic resonance-guided radiotherapy (MRgRT) systems enable continuous real-time acquisition of MRI data during treatment delivery (Ng *et al* 2023). MRI guidance provides excellent soft tissue contrast for organ visualization without ionizing radiation as compared to conventional x-ray-based image guidance. In an online adaptive RT workflow, MR-guidance can be used for real-time monitoring, motion management, beam adaptation, and target tracking (Fast *et al* 2024, Lombardo *et al* 2024). However, the implementation of motion monitoring has largely been limited to acquiring and analyzing one or more orthogonal cine MRI planes (Seregni *et al* 2016, Keiper *et al* 2020, Kielbasa *et al* 2020, Jassar *et al* 2022, 2023, Shaffer *et al* 2025, Sung *et al* 2025). To account for the composite of breathing and other physiologic motions, volumetric monitoring is necessary. Several investigators and commercial solutions have employed golden-angle radial (GAR) sampling schemes to reconstruct breathing-resolved MRI volumes of different respiratory phases (Zhang *et al* 2021, Huttinga *et al* 2022, Feng 2023, Siddiq *et al* 2024). Such sampling schemes offer control over the trade-off between temporal and spatial resolution during 4D MRI reconstruction, which is a key advantage for motion monitoring and has formed the basis for prior investigations of hierarchical motion models in the abdomen (Zhang *et al* 2021).

Breathing is the dominant source of abdominal motion in RT and has been extensively studied in the context of motion monitoring, tracking and dose perturbation (Alam *et al* 2022, Keall *et al* 2006, Fast *et al* 2014, Stemkens *et al* 2016, Bourque *et al* 2018, Borman *et al* 2019, Liu *et al* 2022, Feng 2023, Grimbergen *et al* 2023, Zhang *et al* 2023a). In MRgRT workflows, motion prediction is necessary to compensate for system latencies introduced by image acquisition, reconstruction, motion estimation and machine control (Borman *et al* 2018), which have been reported to take between 200 and 500 ms (Kim *et al* 2020, 2021, Hu *et al* 2022). These latencies can cause inaccurate dose delivery if ignored. By forecasting respiratory motion over short horizons, motion prediction enables latency-compensated gating, tracking, and adaptation strategies in real-time MRgRT.

Various studies have investigated breathing motion prediction (BMP) (Ali Mirzapour *et al* 2019, Lin *et al* 2019, Pham *et al* 2019, 2020a, Romaguera *et al* 2020b, Lombardo *et al* 2022, 2023a, 2023b, Li *et al* 2023, Numakura *et al* 2024, Renner *et al* 2024, Weng *et al* 2024) using classical methods, such as linear regression and Kalman filters, and deep learning (DL) models, such as transformers, long-short term memory (LSTM) and implicit neural representation networks. Regardless of the prediction model, real-time MRgRT requires accurate forecasting with minimal computational latency. Although DL and artificial intelligence approaches have recently shown promising performance for respiratory motion modeling and prediction, developing models that generalize across patients typically requires large, diverse training datasets and computationally intensive optimization (Salari *et al* 2024). In real-time MRgRT, however, strict workflow constraints limit motion model training immediately prior to treatment to only a few minutes of patient-specific imaging, making extensive model optimization impractical. As such, efficient machine learning methods that can be trained rapidly from limited data while maintaining low-latency inference remain relevant.

Among established machine learning techniques, principal component analysis (PCA) has been widely adopted for respiratory motion modeling (Zhang *et al* 2007, Fayad *et al* 2010, 2012, Li *et al* 2011a, Li *et al* 2011b, Klinder and Lorenz 2012, McClelland *et al* 2013) because respiratory deformations fields can often be accurately described by a small number of dominant spatial modes. This dimensionality reduction enables prediction algorithms to forecast only a few PCA coefficients rather than dense deformation fields, substantially reducing both computational cost and model complexity. Building on this representation, Liu *et al* (2021) proposed a patient-specific framework that combines PCA with kernel ridge regression (KRR), in which KRR predicts future PCA coefficients to reconstruct volumetric breathing motion deformation vector fields (DVFs). KRR provides an efficient means of modeling nonlinear relationships through closed-form solutions for training and inference, avoiding the iterative optimization typically required by many DL approaches. By coupling a compact motion representation with nonlinear regression, the PCA-KRR framework provides an efficient approach for patient-specific motion prediction, with low-latency inference characteristics that are desirable for real-time MRgRT applications.

Beyond breathing, other modes of abdominal motion have significant impacts on organ positions. A recent study reported GI antral contractions impart motions proximal to the stomach with similar magnitudes as breathing (Zhang *et al* 2021). Although gastric motility and stomach motion have received less attention compared to breathing, several studies have worked on modeling and predicting such motions (Wysocka *et al* 2014, Lu *et al* 2022, Driever *et al* 2023, Zhang *et al* 2023b). Recent approaches have described contractions as traveling waves along the length of the luminal organs of the digestive tract (Subashi *et al* 2023, van Harten *et al* 2025). These works analyzed stomach motions along its medial axis (MA) (Blum 1967), an imaginary centerline that traces the length of the stomach. The MA facilitates compact modeling of stomach motion and provides a simplified frame of reference to both describe stomach contractile motion and composite it with other motions such as breathing.

During free-breathing MR acquisitions, respiratory-induced motion can complicate the study of stomach contractile motion. Mostafaei *et al* (2018) found that composite respiratory and peristaltic motion of the stomach centroid exceeded 10 mm, and Umezawa *et al* (2021) reported respiratory-induced motions of up to 9.8 mm for the duodenum and 11.4 mm for the stomach. Zhang *et al* (2021) studied respiratory and antral contraction motions of the bowel, stomach and duodenum in a population of 20 scans. For all three organs, they found mean and maximal respiratory motion magnitudes exceeded 7.5 mm and 17 mm, respectively, while mean and maximal contractile motions were about 1.5 mm and 7.5 mm, respectively. Given the magnitude and frequency of respiration, accurate real-time characterization of breathing motion is an important first step towards modeling stomach contractile motion in MRgRT.

In this study, we aim to characterize stomach motion in patients during radiation therapy treatment. Based on previous observations demonstrating the separability of breathing and contractile motions of the stomach (Zhang *et al* 2021), we posit that real-time extraction of breathing motion of the stomach centerline and outer wall (OW) would simplify the modeling of stomach contractile motion in free-breathing patients. To this end, we build upon the patient-specific PCA-KRR framework of Liu *et al* (2021) for BMP from GAR MRI data. In that work, PCA was used to decompose breathing motion DVFs. KRR was then applied to predict future PCA coefficients, which were used to perform rank-1 reconstructions (using only the leading principal component) of future breathing motion DVFs. However, the analysis in Liu *et al* (2021) focused primarily on tracking centroid positions of tumors and OARs, rather than full organ or anatomical region motion, and explored only a limited set of experimental parameters. These limitations motivate a more comprehensive evaluation of BMP performance in the stomach across anatomical regions (*e.g.* MA, OW) and experimental conditions.

The focus of this work is two-fold: first, we characterize the accuracy of BMP of the stomach and its centerline; and second, we provide a detailed analysis of the KRR breathing prediction model under different experimental conditions. We investigate the effect of the amount of training data, the number of predicted PCA coefficients and the future prediction horizon on BMP performance. Together, our findings offer a quantitative benchmark for the precision of respiratory motion prediction via KRR across various abdominal structures and regions, serving as a reference for future motion management strategies in real-time MRgRT.

2. Methods

2.1. Theory: KRR

KRR is a nonlinear regression method performed by kernelizing linear ridge regression (Hofmann *et al* 2008, Schölkopf and Smola 2002). It first applies a nonlinear transformation to the input features, then performs linear ridge regression in the transformed feature space.

Let $\{(\mathbf{x}_1, y_1),\ \ldots,\ (\mathbf{x}_n, y_n)\}$ denote a set of n paired training samples, with features $\mathbf{x}_i \in \mathbb{R}^d$ and responses $y_i \in \mathbb{R}$. Let $\mathbf{y} = [y_1 \cdots y_n]^\mathsf{T}$. KRR selects a kernel k associated with a (possibly infinite-dimensional) nonlinear feature map $\mathbf{\Phi} : \mathbb{R}^d \to \mathbb{R}^D$ such that $k(\mathbf{x}, \mathbf{x}') = \langle \mathbf{\Phi}(\mathbf{x}), \mathbf{\Phi}(\mathbf{x}') \rangle$. Training seeks a weight vector $\mathbf{w} \in \mathbb{R}^d$ and offset $b \in \mathbb{R}$ that minimize the objective function:

$$\sum_{i=1}^{n} \left(y_i - \mathbf{w}^\mathsf{T} \Phi(\boldsymbol{x}_i) - b \right)^2 + \lambda \|\mathbf{w}\|_2^2$$

where λ is the ridge regularization parameter. Using the representer theorem, the prediction for a test sample $\mathbf{x}$ can be written (Nataraj *et al* 2018) as:

$$\hat{f}(\mathbf{x}) = \mathbf{y}^\mathsf{T} \left(\frac{1}{n} \mathbf{1}_n + \mathbf{M}(\mathbf{MKM} + n\lambda \mathbf{I})^{-1} \left(\mathbf{k}(\mathbf{x}) - \frac{1}{n} \mathbf{K} \mathbf{1}_n \right) \right) \tag{1}$$

where $\mathbf{K} \in \mathbb{R}^{n \times n}$ is the kernel matrix with entries $k(\mathbf{x}_i, \mathbf{x}_j)$ for $i, j \in \{1, \ldots, n\}$; the kernel vector $\mathbf{k}(\mathbf{x}) \in \mathbb{R}^n$ has entries $k(\mathbf{x}, \mathbf{x}_i)$; $\mathbf{M} = \mathbf{I} - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^\mathsf{T} \in \mathbb{R}^{n \times n}$ is a de-meaning matrix; $\mathbf{I} \in \mathbb{R}^{n \times n}$ is the identity matrix; and $\mathbf{1}_n \in \mathbb{R}^n$ is a vector of ones. Centering the kernel matrix absorbs the offset term b into the feature space representation. Additional derivation details are described in Nataraj *et al* (2018).

In this work, we used a Laplace (or Laplacian) kernel, defined as:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2}{h} \right) \tag{2}$$

where h is the kernel bandwidth. Preliminary comparisons with Gaussian kernels (not shown) indicated that the Laplace kernel produced more accurate BMPs. We selected the two KRR hyperparameters, the kernel bandwidth h and ridge regularization parameter λ, via 2-fold cross-validation on the training data.

2.2. Data acquisition and preprocessing

Data previously collected from dynamic contrast-enhanced (DCE) MRI examinations performed on patients undergoing RT treatment for hepatocellular cancer under an institution review board approved protocol (The University of Michigan Institutional Review Board, HUM00068061) were analyzed. Patients were asked not to eat for 2 h before scanning. Free-breathing DCE-MRI data covering the upper abdomen were acquired using a spoiled gradient echo sequence with fat suppression (radial VIBE) and GAR stack-of-spokes (GAR-SOS) sampling (Block *et al* 2014, Chandarana *et al* 2014, Feng *et al* 2016, Johansson *et al* 2021). The sequence parameters included repetition times between 2.72–3.96 ms, echo times between 1.16–1.21 ms, and flip angles between 12°–14°. Contrast was injected intravenously at the start of each examination, using either gadoxetic acid (Eovist ®) or gadobenic acid (Multihance ®). MRI data were acquired on a 3 T scanner (Skyra, Siemens Healthineers, Erlangen, Germany) with an 18-channel flexible surface coil (Body Matrix) placed anteriorly and one or more sets of posterior 4-channel coil segments of a 32-segment array (Spine Matrix). A stack of radial k-space samples (termed sample or spoke) was acquired every 170 ms. Further details on the acquisition and sequence parameters can be found in previously published work (Johansson *et al* 2021, Liu *et al* 2021, Zhang *et al* 2021).

A total of 27 DCE-MRI scans acquired from 16 different subjects were analyzed, 18 of which were acquired with Multihance contrast agent and 9 with Eovist. The total number of samples acquired for Multihance and Eovist examinations was 2000 and 7000 samples, corresponding to total scan times of ∼5 min 40 s and ∼19 min 50 s respectively. Subject information and MR acquisition parameters are given in the supplementary material (supplementary tables S1 and S2). The dataset IDs are numbered by contrast agent and patient, with E and M denoting scans with Eovist and Multihance contrast agent, respectively. Appended letters indicate multiple scans from the same subject acquired on separate days.

Preprocessing of the raw k-space data, which included gradient delay correction (Block and Uecker 2011), partial Fourier filling (Cuppen and van Est 1987), distortion correction, noise whitening (Martini *et al* 2010), coil sensitivity estimation and coil combination (Brey and Narayana 1988), was performed as previously described (Johansson *et al* 2018a).

2.3. Image reconstruction

A view-sharing (VS) filter method was used to reconstruct image volumes from GAR-SOS k-space data. First, the preprocessed radial k-space samples were gridded to form complex back-projection volumes at each spoke (Beatty *et al* 2005). A Gaussian VS filter technique was then applied to reconstruct 4D MRIs (Barger *et al* 2002, Bernstein *et al* 2004, Johansson *et al* 2018b). The extent of VS is defined by the parameters $\sigma_{\min}$ and $\sigma_{\max}$ that control the degree of data sharing for low- and high-frequency k-space samples, respectively, and thus define the temporal footprint of the reconstructed images. The temporal resolution of the reconstructed images was defined by the VS filter spacing, or stride. Final images were obtained by taking the magnitude of the complex-valued VS reconstructions. All images were reconstructed to a matrix size of $224 \times 224 \times 96$. Voxel sizes varied by subject depending on the imaging field-of-view (supplementary table S2), with in-plane spatial resolutions ranging from 2.01–2.34 mm^2 and slice thicknesses ranging from 3–4 mm.

2.4. Breathing motion modeling

Dynamic MRI datasets were processed to generate hierarchical motion models using a previously described process (Zhang *et al* 2021) which decomposes motion into three orthogonal components—breathing, GI contractions, and slow drift—represented using DVFs. Breathing motion was assumed to be quasi-cyclic with state characterized by liver position. A surrogate respiratory signal was estimated from high-temporal-resolution (HT-res) images reconstructed using a VS filter with $\sigma_{\min} = 1$, $\sigma_{\max} = 3$ samples, and a temporal resolution of one sample (170 ms) (figure 1(a)). Each HT-res image volume was automatically rigidly aligned to a vendor-reconstructed reference image using a region-limited approach restricted to liver voxels (Johansson *et al* 2016, 2018a) defined within a liver mask extracted from the reference volume (Johansson *et al* 2018b). The inferior-superior translational component of the resulting rigid transformations was used as the surrogate respiratory signal. The acquired k-space samples were sorted according to this signal, and high-spatial-resolution VS reconstructions were performed along the breathing dimension to generate 21 canonical breathing motion state images (figure 1(b)). For

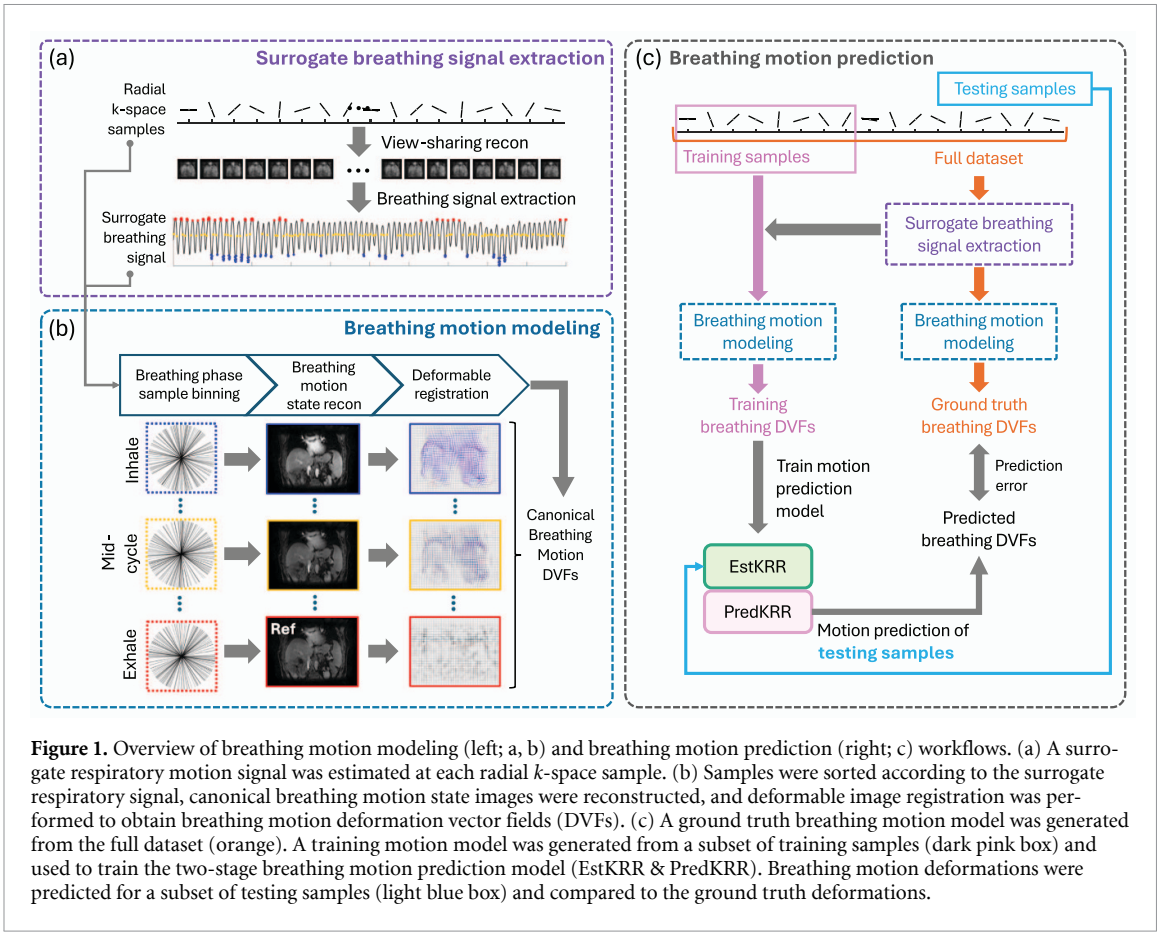

Figure 1. Overview of breathing motion modeling (left; a, b) and breathing motion prediction (right; c) workflows. (a) A surrogate respiratory motion signal was estimated at each radial *k*-space sample. (b) Samples were sorted according to the surrogate respiratory signal, canonical breathing motion state images were reconstructed, and deformable image registration was performed to obtain breathing motion deformation vector fields (DVFs). (c) A ground truth breathing motion model was generated from the full dataset (orange). A training motion model was generated from a subset of training samples (dark pink box) and used to train the two-stage breathing motion prediction model (EstKRR & PredKRR). Breathing motion deformations were predicted for a subset of testing samples (light blue box) and compared to the ground truth deformations.

Table 1. View-sharing filter parameters used for canonical breathing motion state reconstructions, in units of radial samples.

Dataset	$\sigma_{\min}$	$\sigma_{\max}$	Stride
Full (Eovist, 7000 samples)	350	700	350
Full (Multihance, 2000 samples)	100	200	100
T1000 (1000 samples)	50	100	50
T500 (500 samples)	25	50	25
T350 (350 samples)	18	36	18

canonical motion state reconstructions, the VS parameters used were $\sigma_{\min} = 350$, $\sigma_{\max} = 700$ for Eovist scans, and $\sigma_{\min} = 100$, $\sigma_{\max} = 200$ for Multihance scans (table 1).

The canonical breathing motion states were deformably aligned to the end-exhale state using cubic B-spline-based deformable image registration (DIR) implemented in NiftyReg (Modat *et al* 2010), with a normalized mutual information cost function and parameters matching previous studies (Johansson *et al* 2018a, Liu *et al* 2021). This yielded a set of 21 canonical breathing DVFs, denoted $\mathbf{D}_i \in \mathbb{R}^{224\times224\times96\times3}$, where the end-exhale DVF, $\mathbf{D}_{21}$, was an identity transform corresponding to zero displacement. Sample-wise breathing DVFs were obtained by min-max normalizing the surrogate respiratory signal to a continuous breathing state between 1 and 21 and linearly interpolating the canonical DVFs accordingly (Zhang *et al* 2021).

To obtain compact representations of breathing motion, PCA (Hotelling 1933) was performed on the canonical breathing DVFs after conversion to displacement vector fields by subtracting the identity DVF. A binary mask excluding background voxels was generated on the reference motion state using intensity thresholding and morphological operations. The displacement fields were masked to exclude background voxels and converted to vectorized representations, $\mathbf{B}_i$. The vectorized fields were centered by subtracting the mean DVF vector $\mathbf{b}_0$ and concatenated into a 2D matrix $\widetilde{\mathbf{B}}$. PCA was performed on the centered data using an economy-size singular value decomposition:

$$\widetilde{\mathbf{B}} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}}$$

where $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_{21}]$ contains the left singular vectors, whose columns form the PCA basis vectors, $\mathbf{S}$ is a diagonal matrix of singular values, and $\mathbf{V}$ contains the right singular vectors.

PCA coefficients provide a low-dimensional representation of breathing motion that can be combined with the PCA basis vectors to recover volumetric breathing deformations (Liu *et al* 2021). The PCA coefficients, $\boldsymbol{\alpha}_{\text{canon}}$, of the canonical breathing motion states were calculated as
Each column of $\boldsymbol{\alpha}_{\text{canon}}$ corresponds to one canonical breathing motion state, and each row corresponds to one PCA mode. The ith row, $\boldsymbol{\alpha}^{(i)}_{\text{canon}}$, represents the coefficient trajectory of the ith PCA mode across the canonical breathing motion states. Sample-wise PCA coefficients were obtained by linearly interpolating each canonical coefficient trajectory $\boldsymbol{\alpha}^{(i)}_{\text{canon}}$ according to the surrogate respiratory signal (Liu *et al* 2021), yielding scalar coefficients $\alpha^{(i)}_t \in \mathbb{R}$ at each sample t. These sample-wise PCA coefficients, computed from the ground truth data, are referred to as the true PCA coefficients and were used to represent breathing motion for prediction.

Breathing DVFs at sample t were reconstructed using a rank-r PCA model as:

$$\widehat{\mathbf{d}}_t = \mathbf{b}_0 + \sum_{k=1}^{r} \alpha^{(k)}_t \mathbf{u}_k \tag{3}$$

where using $r < 21$ yields low-rank approximations and $r = 21$ recovers the full breathing DVFs. The PCA-reconstructed displacement fields were unmasked and reshaped into their original volumetric form $\widehat{\mathbf{d}}_t \in \mathbb{R}^{224\times224\times96\times3}$.

2.5. Training and ground truth breathing motion models

Breathing motion modeling was performed on each dataset using all acquired data (*i.e.* 2000 or 7000 GAR-SOS samples), which we refer to as the *full* breathing motion models. Breathing motion DVFs derived from these full motion models served as the ground truth for evaluating motion prediction performance.

To simulate a pre-treatment MRI acquisition used for training a motion prediction model, training data subsets containing n_{train} GAR-SOS samples were extracted from each full dynamic MRI scan and processed to generate corresponding training breathing motion models. Training datasets containing $n_{\text{train}} = 350$, 500 and 1000 samples were investigated, corresponding to acquisition times of approximately $t_{\text{train}} = 59.5$, 85, and 170 s respectively. These training sets are denoted as T350, T500, and T1000.

Each training dataset was processed independently to generate a training breathing motion model (figure 1(c)). The modeling procedure was identical to that used for the full dataset, except that canonical breathing motion states were reconstructed using *only* the training samples. The VS parameters used for each training dataset to reconstruct the 21 canonical breathing motion state images are listed in table 1 and were scaled according to the number of available radial samples. Coronal planes sampled through canonical breathing motion state images reconstructed from one example patient are shown in figure 2. DVFs derived from the training motion models were processed using PCA to obtain PCA basis vectors $\mathbf{U}$, sample-wise PCA coefficients $\alpha^{(i)}_t$, and the mean DVF vector $\mathbf{b}_0$. Ground truth breathing motion deformations were derived from the full motion models, whereas the training motion models were used for BMP.

2.6. KRR-BMP model

We implemented a KRR-based BMP model, termed KRR-BMP, following the approach introduced by Liu *et al* (2021). The model consists of two stages (figure 3): in the first stage (EstKRR), breathing motion at the current time is estimated, and in the second stage (PredKRR), breathing motion at a future time is predicted. The model exploits the low-rank structure of breathing motion to represent deformations via PCA coefficients, reducing computational cost. We use the term *motion estimation* to refer to estimates of current motion, and *motion prediction* to refer to estimates of future motion (McClelland *et al* 2013).

2.6.1. EstKRR stage

The first stage, EstKRR, estimated the leading PCA coefficient of breathing motion at the current time (*i.e.* without future prediction) using features extracted from an image volume reconstructed at that time. For each GAR-SOS sample k_i acquired at time i, an image volume $\mathbf{H}_i \in \mathbb{R}^{224\times224\times96}$ was reconstructed using a truncated HT-res (tHT-res) VS filter centered at sample k_{i-1} with a width of three samples (*i.e.* using samples k_{i-2}, k_{i-1} and k_i). The tHT-res reconstructions were identical to the previously

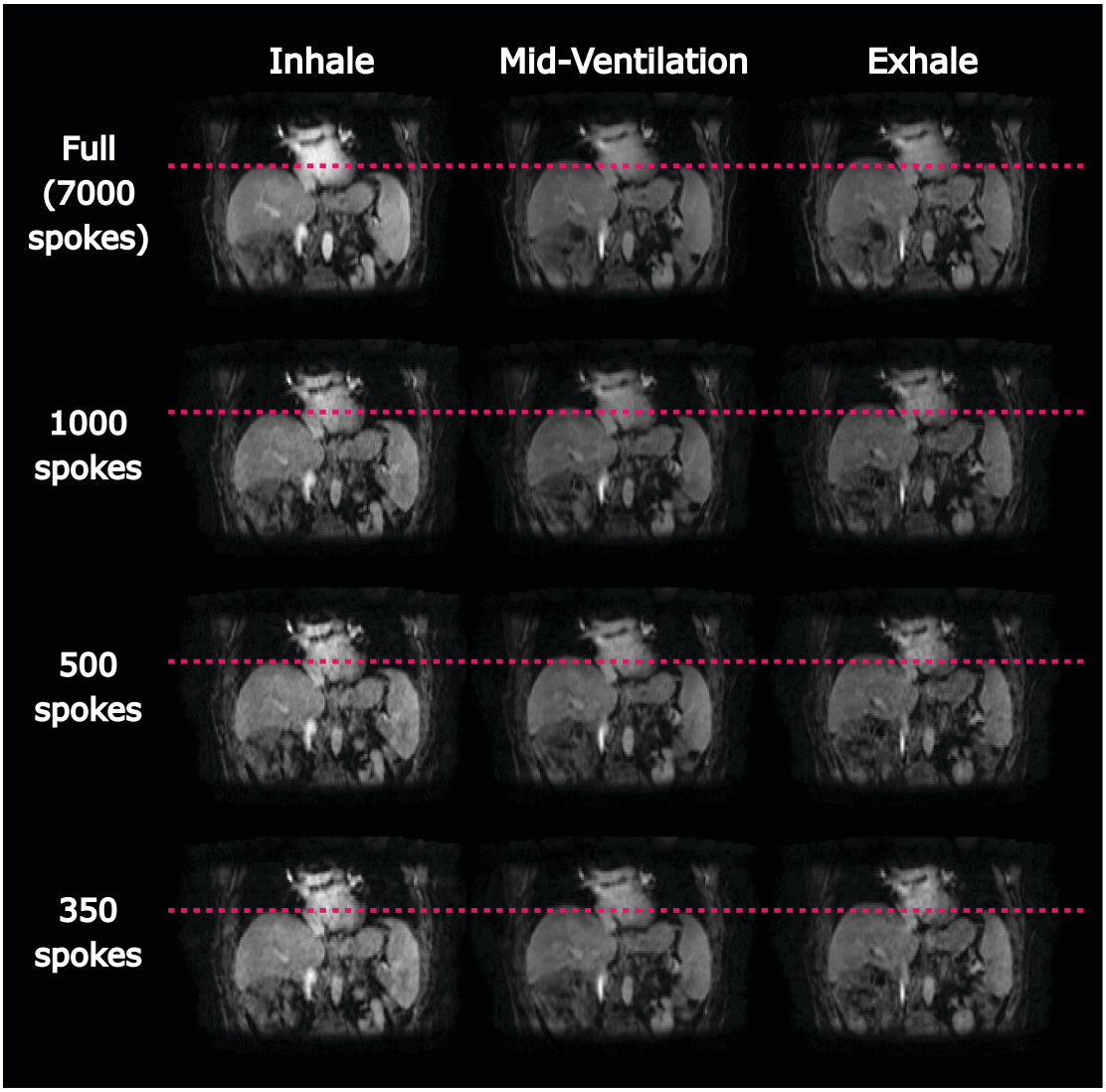

Figure 2. Comparison of canonical breathing motion state view-sharing reconstructions from an example patient. Breathing motion states at inhale (left column), mid-ventilation (middle column) and exhale (right column) are shown, reconstructed from datasets with 7000 samples (top row), 1000 samples (second row), 500 samples (third row) and 350 samples (bottom row). Red dashed lines indicate liver position in the inhale motion states.

described HT-res reconstructions (which used $\sigma_{\min} = 1$ and $\sigma_{\max} = 3$), except that the VS filter width was truncated to strictly three samples.

A liver contour was automatically generated on the reference breathing motion (end-exhale) state and cropped to include only the superior-most one-third of liver voxels. Voxel intensities within the cropped liver mask were extracted from image $\mathbf{H}_i$ and organized into a feature vector, which was subsequently scaled to unit norm. The number of cropped liver voxels, n_{liver}, ranged from approximately 27 000–87 000 across datasets.

Letting n denote the number of EstKRR training samples, the training feature vectors were collected in the set $\{\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_n\}$. The corresponding scalar training responses—*i.e.* the first PCA coefficients of breathing motion derived from the training DVFs—were collected into the vector $\mathbf{y}_{\text{est}} = \left[\alpha_1^{(1)}, \alpha_2^{(1)}, \ldots, \alpha_n^{(1)}\right]^{\mathsf{T}}$. During EstKRR training, the feature vectors were used to construct the kernel matrix $\mathbf{K}_{\text{est}} \in \mathbb{R}^{n\times n}$, with entries $[\mathbf{K}_{\text{est}}]_{i,j} := k_{h_{\text{est}}}(\mathbf{L}_i, \mathbf{L}_j)$, $i,j = 1,\ldots,n$, where $k_{h_{\text{est}}}(\cdot,\cdot)$ denotes the Laplacian kernel defined in equation (2) with bandwidth h_{est}.

For a test sample t with feature vector $\mathbf{L}_t$, the kernel vector $\mathbf{k}_{\text{est}}(\mathbf{L}_t)$ was computed as $\mathbf{k}_{\text{est}}(\mathbf{L}_t) := [k_{h_{\text{est}}}(\mathbf{L}_t, \mathbf{L}_1), \ldots, k_{h_{\text{est}}}(\mathbf{L}_t, \mathbf{L}_n)]^{\mathsf{T}}$, and the first PCA coefficient response was estimated using equation (1):

$$\hat{\alpha}_t^{(1)} = (\mathbf{y}_{\text{est}})^{\mathsf{T}} \left(\frac{1}{n}\mathbf{1}_n + \mathbf{M}(\mathbf{M}\mathbf{K}_{\text{est}}\mathbf{M} + n\lambda_{\text{est}}\mathbf{I})^{-1} \left(\mathbf{k}_{\text{est}}(\mathbf{L}_t) - \frac{1}{n}\mathbf{K}_{\text{est}}\mathbf{1}_n \right) \right). \tag{4}$$

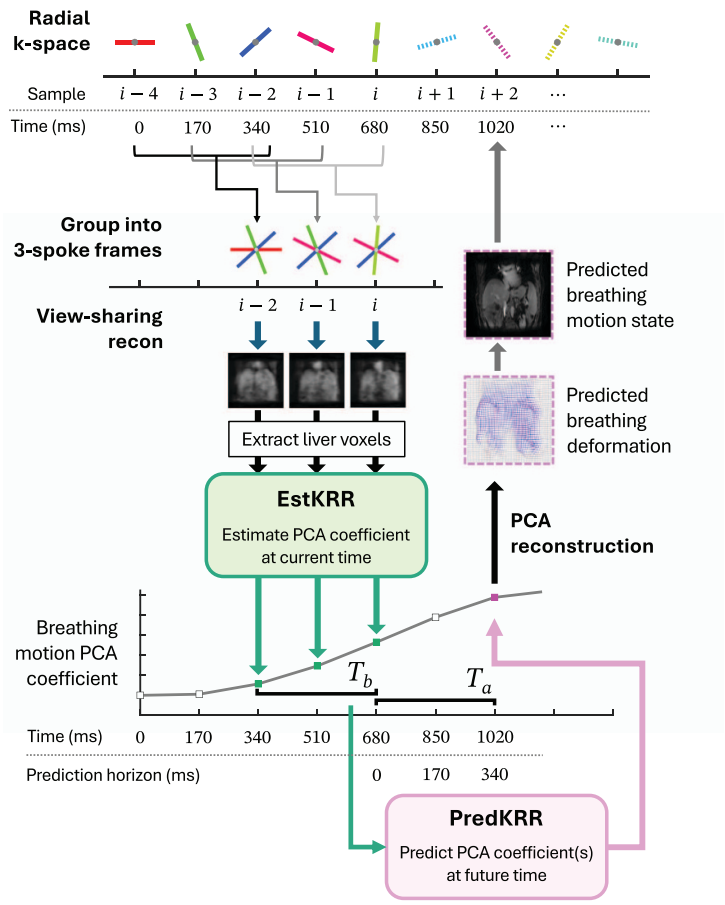

Figure 3. Outline of kernel ridge regression-based breathing motion prediction (KRR-BMP). Golden-angle radial stack-of-spokes k-space data are grouped into 3-sample frames and reconstructed using truncated high temporal-resolution view-sharing (tHT-res VS). The EstKRR stage estimates the first PCA coefficient at the latest frame time from liver voxel intensities. The PredKRR stage predicts the first r PCA coefficients at a future time T_a ms ahead using EstKRR-estimated coefficients from the past T_b ms. The predicted coefficients and training PCA basis vectors are used to perform a rank-r reconstruction of the breathing deformation vector field and corresponding breathing motion state.

The optimal hyperparameters h_{est} and λ_{est} were selected via 2-fold cross-validation on the training dataset. Because all experiments were conducted retrospectively, the EstKRR model was applied to all test samples simultaneously, enabling efficient computation while yielding results identical to those obtained from sequential processing.

2.6.2. PredKRR stage

The second stage, PredKRR, predicted PCA coefficients at a future time using EstKRR-estimated coefficients as input features. This stage introduced three additional experimental parameters: $n_b \in \mathbb{Z}^+$, the number of previously estimated PCA coefficients used as features (*i.e.* the feature vector dimension); $n_a \in \mathbb{Z}^+$, the number of samples ahead to predict (*i.e.* the prediction horizon); and $r \in \mathbb{Z}^+$, the number of leading PCA coefficients predicted at each sample. The parameters n_b and n_a were expressed in units of samples but could be converted to time by multiplying by the inter-sample interval of 170 ms, yielding a look-back window duration $T_b = 170n_b$ and a prediction horizon $T_a = 170n_a$, both in milliseconds. All prediction experiments presented here used $n_b = 5$ samples ($T_b = 850$ ms).

During PredKRR training, feature vectors were formed using the leading ground truth PCA coefficients (derived from motion modeling of the full dataset) from the n_b most recent samples, $\mathbf{x}_i = \left[\alpha^{(1)}_{i-n_b+1}, \ldots, \alpha^{(1)}_i\right]^{\mathsf{T}} \in \mathbb{R}^{n_b}$. The corresponding responses were the r leading PCA coefficients at the future sample, $\mathbf{y}_i = \left[\alpha^{(1)}_{i+n_a}, \ldots, \alpha^{(r)}_{i+n_a}\right]^{\mathsf{T}} \in \mathbb{R}^r$. The training responses were collected into the matrix $\mathbf{y}_{\text{pred}} = \left[\mathbf{y}_1, \ldots, \mathbf{y}_{n_f}\right]^{\mathsf{T}} \in \mathbb{R}^{n_f \times r}$, where $n_f = n - n_b - n_a + 1$ is the number of available training samples. The training feature vectors were used to construct the kernel matrix $\mathbf{K}_{\text{pred}} \in \mathbb{R}^{n_f \times n_f}$, with entries $[\mathbf{K}_{\text{pred}}]_{i,j} := k_{h_{\text{pred}}}(\mathbf{x}_i, \mathbf{x}_j)$, $i, j = 1, \ldots, n_f$, where $k_{h_{\text{pred}}}(\cdot,\cdot)$ denotes the Laplacian kernel defined in equation (2) with bandwidth h_{pred}.

For a test sample acquired at time index t, the PredKRR feature vector $\widehat{\mathbf{x}}_t$ was constructed using the Est-KRR estimated PCA coefficients (equation (4)) from the n_b preceding samples, $\widehat{\mathbf{x}}_t = \left[\hat{\alpha}^{(1)}_{t-n_b+1}, \ldots, \hat{\alpha}^{(1)}_{t-1}, \hat{\alpha}^{(1)}_t\right]^{\mathsf{T}} \in \mathbb{R}^{n_b}$. The kernel vector was computed as: $\mathbf{k}_{\text{pred}}(\widehat{\mathbf{x}}_t) := \left[k_{h_{\text{pred}}}(\widehat{\mathbf{x}}_t, \mathbf{x}_1), \ldots, k_{h_{\text{pred}}}(\widehat{\mathbf{x}}_t, \mathbf{x}_{n_f})\right]^{\mathsf{T}}$. The r leading PCA coefficients predicted at sample $t+n_a$, denoted $\widehat{\widehat{\boldsymbol{\alpha}}}_{t+n_a} \in \mathbb{R}^r$, were then computed following equation (1):

$$\widehat{\widehat{\boldsymbol{\alpha}}}_{t+n_a} = \left(\mathbf{y}_{\text{pred}}\right)^{\mathsf{T}} \left(\frac{1}{n_f}\mathbf{1}_{n_f} + \mathbf{M}\left(\mathbf{M}\mathbf{K}_{\text{pred}}\mathbf{M} + n_f\lambda_{\text{pred}}\mathbf{I}\right)^{-1}\left(\mathbf{k}_{\text{pred}}(\widehat{\mathbf{x}}_t) - \frac{1}{n_f}\mathbf{K}_{\text{pred}}\mathbf{1}_{n_f}\right)\right). \tag{5}$$

The optimal hyperparameters h_{pred} and λ_{pred} were selected via 2-fold cross-validation on the training dataset. The double hat notation $\widehat{\widehat{\boldsymbol{\alpha}}}$ is used to emphasize that the PredKRR predictions were obtained using input features that were themselves estimates from EstKRR. As with EstKRR, the PredKRR model was applied to all test samples simultaneously.

2.6.3. Reconstruction of breathing motion DVFs

Following PredKRR, the predicted breathing motion DVFs were reconstructed at each test sample by substituting the predicted PCA coefficients, computed using equation (5), into equation (3):

$$\widehat{\mathbf{d}}_t = \mathbf{b}_0 + \sum_{k=1}^{r} \widehat{\widehat{\alpha}}^{(k)}_t \mathbf{u}_k, \tag{6}$$

where $\widehat{\widehat{\alpha}}^{(k)}_t$ is the predicted kth PCA coefficient at sample t, $\mathbf{u}_k$ is the kth PCA basis vector, $\mathbf{b}_0$ is the mean DVF vector, and r is the number of predicted PCA coefficients used for reconstruction (*i.e.* the PCA rank).

Breathing motion DVFs were also reconstructed using two PCA-based reference models. First, DVFs were reconstructed using the true PCA coefficients obtained from the training data in place of the predicted coefficients, which we refer to as TrainPCA. These DVFs were computed by substituting $\alpha^{(k)}_t$ in place of in equation (6) and represent the scenario of perfect PCA coefficient prediction by KRR-BMP. Second, DVFs were reconstructed using PCA basis vectors and coefficients derived from the full dataset, rather than those from the training data, which we refer to as FullPCA. The FullPCA DVFs represent the best possible rank-r approximation of the ground truth breathing motion DVFs in the least-squares (PCA) sense. The difference between TrainPCA and FullPCA isolates the error arising from training-set-limited motion modeling. Together, PredKRR, TrainPCA, and FullPCA form a hierarchy that isolates errors due to coefficient prediction, training-set PCA modeling, and rank truncation.

2.6.4. Cross-validation for hyperparameter tuning

For each training set, the hyperparameters h and λ for EstKRR and PredKRR were determined by 2-fold cross-validation on the training data. The training samples were randomly partitioned into two equally sized folds, and a grid search was performed over the following log-spaced hyperparameter values: $h \in \left\{10^1, 10^2, \cdots, 10^7, 10^8\right\}$, $\lambda \in \left\{0, 10^{-15}, 10^{-13}, \cdots, 10^{-3}, 10^{-1}\right\}$. The parameter combination yielding the lowest average normalized root-mean-square error (NRMSE) in the first PCA coefficient across the two cross-validation folds was selected for testing. The NRMSE was obtained by dividing the RMSE by the L2 norm of the ground truth signal after mean subtraction. Cross-validation was performed separately for the EstKRR and PredKRR stages, and for PredKRR it was also performed independently for each future prediction time.

2.7. Evaluation of prediction accuracy

The accuracy of BMPs was evaluated by comparing the predicted breathing motion DVFs with ground truth DVFs. The ground truth DVFs corresponded to the sample-wise breathing motion DVFs obtained from motion modeling using all samples in the full dataset. These full DVFs were chosen as ground truth because they incorporated all data from the scan session, providing the most comprehensive representation of the motions occurring during the scan.

For each test sample acquired at time index t, the predicted breathing DVFs were reconstructed using equation (6), and the prediction residual error was calculated as the difference between the predicted and ground truth DVFs. The end-point error (EPE) (Baker *et al* 2011) of the predicted DVFs was computed as the Euclidean norm (magnitude) of the residual error vector at each voxel location r:

$$\text{EPE}_t(r) = \left\|\widehat{\mathbf{d}}_t(r) - \mathbf{d}_t(r)\right\|_2.$$

The EPEs quantify the absolute error magnitude of BMPs with respect to the ground truth and are expressed in millimeters.

2.7.1. Stomach- and organ-specific prediction evaluation

For each dataset, binary masks of the body, liver, liver surface, stomach, stomach OW, and stomach MA were generated on the breathing reference (exhale) state image. The body mask excluded background voxels and was automatically extracted based on intensity thresholding and morphological operations. Contours of the liver and stomach surfaces were automatically generated by an in-house auto-contouring tool based on the MONAI U-Net framework (Cardoso *et al* 2022) and subsequently manually refined by an expert. Liver and stomach masks were defined as the voxels within the refined contours. Liver surface and stomach OW masks were defined as the periphery voxels of the respective organs. The stomach MA was computed from the refined stomach segmentation by manually specifying the esophageal and duodenal endpoints and applying the *Extract Centerline* module from the SlicerVMTK extension in 3D Slicer (Antiga *et al* 2008, Hähn 2009, Piccinelli *et al* 2009). These masks defined the spatial regions of interest (ROIs) over which the prediction errors were computed.

One of the primary objectives of this study was to assess the accuracy of BMPs in the stomach, with an additional focus on the stomach MA and OW. For comparison, prediction errors were also computed across the entire body, liver, and liver surface. These additional regions provided baseline references, facilitating context for the stomach-specific results as well as further evaluation of the prediction model's generalizability across abdominal organs and substructures.

Summary statistics of the prediction EPEs were computed within each anatomical ROI described above and across different test samples. For each ROI analyzed, the mean, standard deviation, median, 95th percentile (P95), and 99th percentile (P99) of the prediction EPEs were calculated, both cumulatively across all test samples (cumulative statistics) and separately for each test sample (sample-level statistics).

2.8. BMP experiments

We systematically evaluated the performance of the KRR-BMP model at the population-level by examining three key experimental parameters that influence motion prediction accuracy: training dataset size, prediction horizon, and PCA rank. For each dynamic MRI scan, the first 500 samples (85 s) were excluded due to signal intensity variations caused by contrast agent arrival. Disjoint training and testing sets were constructed from samples 501–2000.

Training sets containing 1000 (T1000), 500 (T500) and 350 (T350) samples were evaluated, corresponding to acquisition times of 170 s, 85 s, and 59.5 s, respectively, to simulate realistic clinical scenarios in which training data are acquired immediately before treatment. The training datasets were processed as described in section 2.5 to obtain training breathing motion models. Two different test sets were used: test set 1 contained samples 1501–2000, and test set 2 contained samples 501–1000. For test set 1, the T1000 training set used samples 501–1500, and for test set 2, it used samples 1001–2000. The T500 training set used samples 1001–1500, and the T350 training set used samples 1151–1500 for both test sets.

For each training/testing set combination, experiments were performed using prediction horizons $T_a = 170$, 340, 510, 680, 1020, and 2040 ms (corresponding to 1, 2, 3, 4, 6, and 12 samples), and reconstructed using PCA ranks $r = 1$, 2, 3, 4, and 5. The latter represents the number of PCA coefficients predicted by the KRR-BMP model.

Additionally, experiments were performed to investigate the effect of intra-scan variability on training and prediction using a subset of nine Eovist datasets containing 7000 samples (corresponding to 20 min scan times). For each dataset, the full 7000 sample acquisition was partitioned into seven non-overlapping 1000 sample (170 s) subsets. Each 1000 sample subset was used for training, and testing was performed on the remaining 6000 samples (figure 8(a)). We refer to each training set as a *block*; for example, block 1 was trained on samples 1–1000 and tested on samples 1001–7000. The EPEs of the predicted DVFs were compared with those from the TrainPCA and FullPCA baselines (described in section 2.6.3). These experiments used a prediction horizon of $T_a = 340$ ms and a PCA rank of 1.

2.9. Computational inference time

To evaluate the computational feasibility of real-time deployment, inference times were measured using an optimized single-sample implementation of the proposed KRR-BMP framework (described in sections 2.6.1 and 2.6.2), in which all training-dependent quantities were precomputed and reused across test samples. Timing measurements were obtained for the EstKRR inference, PredKRR inference and PCA-based breathing DVF reconstruction steps.

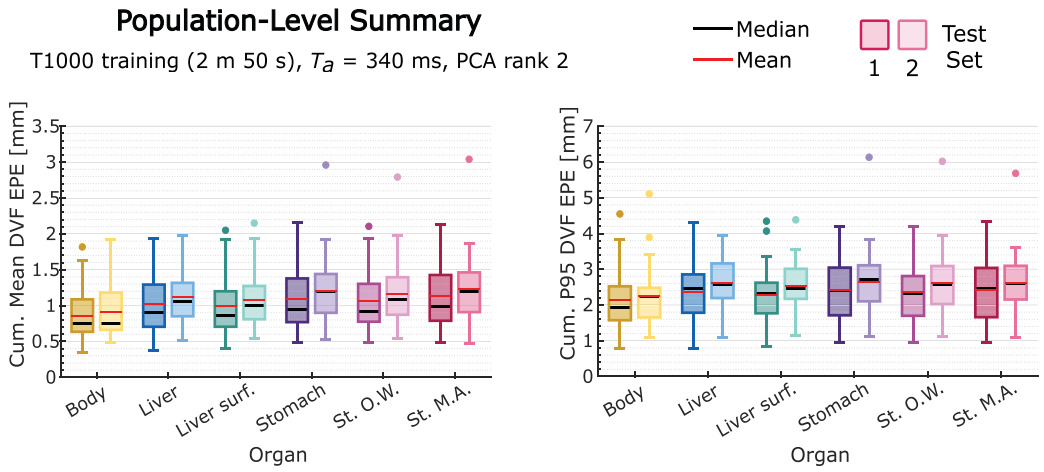

Figure 4. Boxplots of cumulative mean (left) and 95th percentile (P95; right) DVF end-point error (EPE) across 27 datasets for experiments using a prediction horizon of $T_a = 340$ ms, PCA rank of 2, and T1000 training data. Colors denote organs; darker and lighter shades correspond to test sets 1 and 2, respectively. Black and red lines show population median and mean, and outliers are identified as circles.

All timing experiments were performed using MATLAB R2024b (version 24.2.0.2833386) running on Red Hat Enterprise Linux 8.10 (Ootpa) on a workstation equipped with two 18-core Intel Xeon Gold 6354 processors (36 physical CPU cores total, 3.00 GHz) and 791 GB of system memory. GPU acceleration and explicit MATLAB parallel processing were not used. MATLAB was configured with a maximum of 36 computational threads (maxNumCompThreads = 36).

Reported timing measurements corresponded to the online inference stage only, assuming that the training feature vectors and responses and the test feature vector were already loaded and stored in memory. Therefore, data loading, image reconstruction, feature extraction, preprocessing steps and motion modeling were excluded from the reported computation times.

Inference time experiments were performed using six representative datasets spanning a range of cropped liver and body voxel counts. These quantities determine the dimensions of the EstKRR feature vectors and the rank-r PCA basis vectors used for DVF reconstruction, respectively, and therefore directly influence the computational cost. For each dataset, test set 1 was used for evaluation. The EstKRR inference, PredKRR inference, and rank-2 PCA-based breathing DVF reconstruction steps were timed independently for each of the 500 test samples. The reported inference times correspond to the mean computation time per test sample for each processing stage, averaged across all 500 test samples within each dataset.

3. Results

Figure 4 and table 2 show group summaries of cumulative prediction error statistics from experiments using T1000 training data, $T_a = 340$ ms and PCA rank 2. Figure 4 presents boxplots of the cumulative mean (left) and 95th percentile (right) DVF EPEs across all 27 patient datasets. Each box represents the population interquartile range (25th–75th percentile), with horizontal black and red lines marking the population median and mean, respectively. Whiskers indicate 1.5 times the interquartile range, with outliers falling outside of the whiskers denoted with dots. Boxplots for the body (yellow), liver (blue), liver surface (green), stomach (purple), stomach OW (St. OW; pink), and stomach MA (St. MA; red) organs are shown along the x-axes. For each organ, results from test datasets 1 and 2 are shown adjacently in the darker and lighter shade boxplots, respectively.

Table 2 reports corresponding population summary statistics. Across all organs, the population average of the mean EPE values was below 1.24 mm, indicating sub-voxel prediction accuracy. The body exhibited the lowest errors (population-average mean EPE ≈0.9 mm, P95 ≈2.2 mm), while the stomach and its subregions showed slightly higher errors (population-average mean EPE ≈1.2 mm, P95 ≈2.5 mm). Results were similar across test sets 1 and 2, with test set 2 showing a slight, consistent increase in both mean and 95th percentile EPEs (∼0.1 mm and ∼0.25 mm, respectively), suggesting stable model generalizability. The modest inter-subject variability, reflected by tight interquartile ranges and limited error spread (figure 4), further indicates robust performance across subjects, test sets and organs.

Table 2. Group-level statistics of cumulative breathing motion prediction errors from experiments with $T_a = 340$ ms, PCA rank 2, and T1000 training set.

	Mean EPE (mm)				95th Percentile (P95) EPE (mm)			
	Population mean ± std		Population min—max		Population mean ± std		Population min—max	
Organ	Test set 1	Test set 2	Test set 1	Test set 2	Test set 1	Test set 2	Test set 1	Test set 2
Body	0.86 ± 0.35	0.91 ± 0.36	0.35–1.82	0.48–1.93	2.13 ± 0.84	2.26 ± 0.88	0.79–4.55	1.08–5.11
Liver	1.02 ± 0.41	1.11 ± 0.39	0.37–1.93	0.51–1.98	2.38 ± 0.86	2.61 ± 0.75	0.78–4.31	1.10–3.93
Liver surface	0.99 ± 0.40	1.08 ± 039	0.41–2.05	0.54–2.15	2.29 ± 0.81	2.52 ± 0.70	0.85–4.35	1.15–4.38
Stomach	1.09 ± 0.41	1.20 ± 0.50	0.48–2.16	0.53–2.96	2.42 ± 0.84	2.66 ± 0.97	0.96–4.20	1.11–6.13
Stomach outer wall	1.06 ± 0.40	1.16 ± 0.48	0.48–2.11	0.54–2.79	2.37 ± 0.83	2.61 ± 0.97	0.95–4.20	1.12–6.02
Stomach medial axis	1.13 ± 0.44	1.24 ± 0.51	0.48–2.13	0.47–3.04	2.43 ± 0.93	2.62 ± 0.90	0.94–4.33	1.10–5.68

Figure 5 compares prediction accuracy across training set sizes (T1000, T500 and T350) for predictions of test set 1 at $T_a = 340$ ms and PCA rank 2. Separate subplots are shown for the mean (a, top left), median (b, top right), 95th percentile (c, bottom left), and 99th percentile (d, bottom right) EPEs. Overall, the T1000 model achieved the lowest mean and upper bound (P95, P99) errors across the population. As expected, reducing the number of training samples led to a gradual increase in cumulative DVF EPE across all organs, with noticeable differences observed in the stomach and its subregions. For the stomach, the population-average mean and 95th percentile EPEs increased from approximately 1.09 mm and 2.42 mm (T1000) to 1.28 mm and 2.84 mm (T500), and further to about 1.32 mm and 2.93 mm (T350), respectively. For mean and median EPE, a steady increase in the population averages (red lines), 75th percentiles (tops of boxes) and upper bounds (top whiskers) was consistently observed in all organs, with the T350 exhibiting a noticeable increase in the spread of errors compared to T1000 and T500. The error spreads for P95 and P99 EPE exhibited less variability across training set sizes and organs, with all population-average P95 EPEs less than 3 mm.

Figure 6 compares prediction accuracy across prediction horizons for experiments using test set 1 with T1000 training data and PCA rank 2. Surprisingly, some datasets exhibited decreases in prediction errors at the farthest horizon (2040 ms, right-most boxplots in each group) compared to shorter horizons. However, in general, prediction errors increased with prediction horizon. Error spreads at the population level increased in the stomach ROIs at the farthest horizon, but otherwise remained consistent across prediction horizons for both mean and 95th percentile EPE.

Figure 7 compares prediction accuracy across PCA ranks for experiments using test set 1 with T1000 training data and a prediction horizon of $T_a = 340$ ms. Group-level mean EPE (red lines) decreased when using a PCA rank of 2, but errors began to increase when the rank exceeded 2. Increasing the PCA rank also widened the group-level error distribution: the population 25th and 50th percentile errors (bottom and middle edges of the boxplots) decreased, whereas the upper bound errors (75th percentiles, top edges) increased, indicating larger variability across subjects.

Bar plots summarizing BMP performance for PredKRR, with a prediction horizon of $T_a = 340$ ms, and the two PCA-based reference reconstructions (TrainPCA and FullPCA) across multiple 1000-sample training sets are shown in figure 8(b). Corresponding statistics of the cumulative mean DVF EPE of the stomach MA are reported in table 3. FullPCA represents the best possible rank-1 approximation of the ground truth breathing motion and as expected, exhibited the lowest errors, with a population average of 0.43 mm (table 3). Unlike TrainPCA and PredKRR, which rely on motion models derived solely from the training block to reconstruct test-sample DVFs, FullPCA uses a motion model derived from all available samples and therefore serves as a full-scan oracle. PredKRR errors averaged across training blocks did not exceed 1.8 mm, with a population-average error of 1.36 mm (table 3). Across the nine datasets analyzed, the average difference between PredKRR and TrainPCA errors ranged from 0.06 mm to 0.53 mm, whereas the gap between TrainPCA and FullPCA errors ranged from 0.31 mm to 0.94 mm (table 3). One dataset (E6) exhibited substantially larger TrainPCA and PredKRR errors for one training block, which were approximately twice those observed in the remaining blocks (figure 8(b)).

Computational inference times are summarized in table 4. Timing experiments were performed on six representative datasets spanning a range of cropped liver and body voxel counts to evaluate the effect of feature and reconstruction dimensions on computational cost. Across the six datasets, the mean EstKRR, PredKRR, and DVF reconstruction times were 5.17 ms, 0.43 ms, and 7.43 ms per sample, respectively, corresponding to a mean total online prediction time of 13.02 ms per sample.

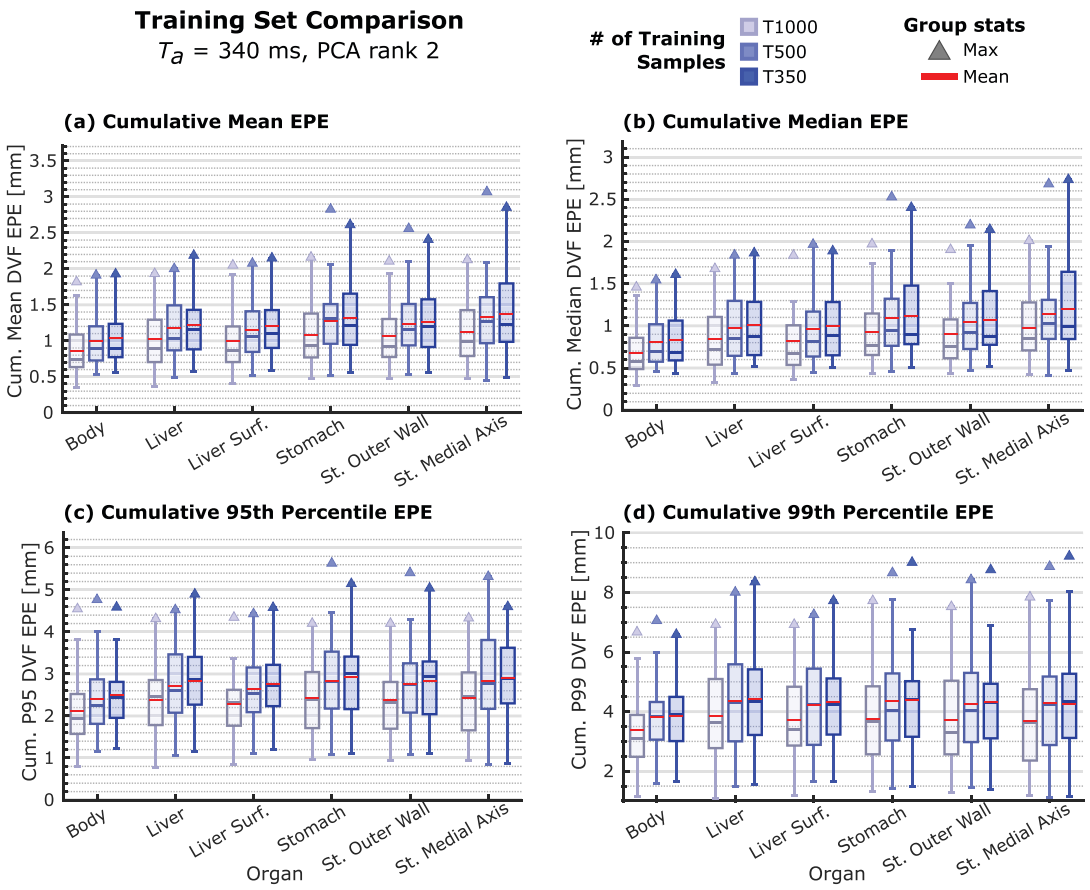

Figure 5. Effect of training set size on cumulative DVF end-point errors (EPEs) across organs for test set 2, using a prediction horizon of $T_a = 340$ ms and a PCA rank of 2. Boxplots show cumulative mean (a, top left), median (b, top right), 95th percentile (P95; c, bottom left), and 99th percentile (P99; d, bottom right) DVF EPEs. Box colors correspond to training dataset size given in the legend (top right corner). Red lines denote group means and triangles indicate group maxima.

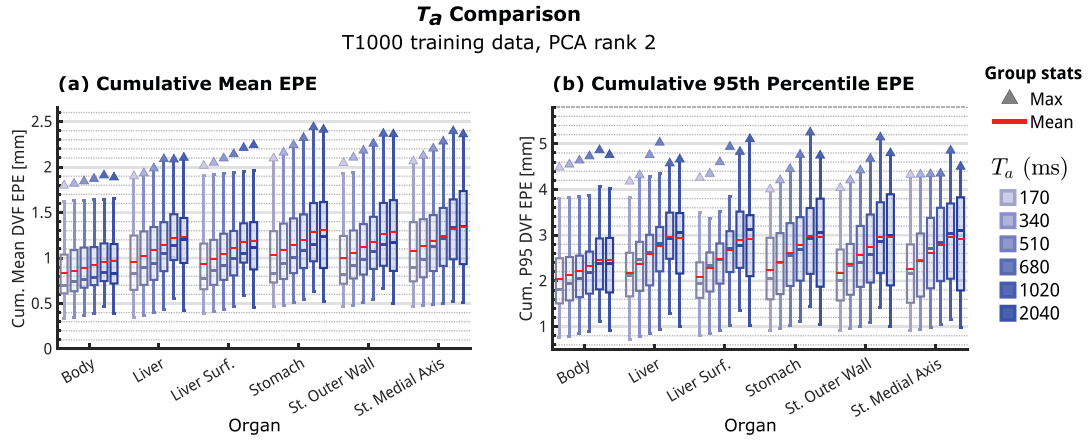

Figure 6. Boxplots comparing the effect of prediction horizon on cumulative mean end-point error (EPE) (a, left) and cumulative 95th percentile EPE (b, right) across patient datasets. Box colors correspond to prediction horizons given in the legend (far right). Results are shown for test set 1, using T1000 training data and PCA rank 2.

4. Discussion

Here, we assessed BMP of the stomach, including its centerline and OW, as a foundational step toward real-time gastric motility modeling in MRgRT. By comparing prediction performance across multiple forecasting horizons and training set sizes, we characterized anatomy-specific prediction accuracy under a range of realistic experimental conditions. Our results offer insight into the amount of training data

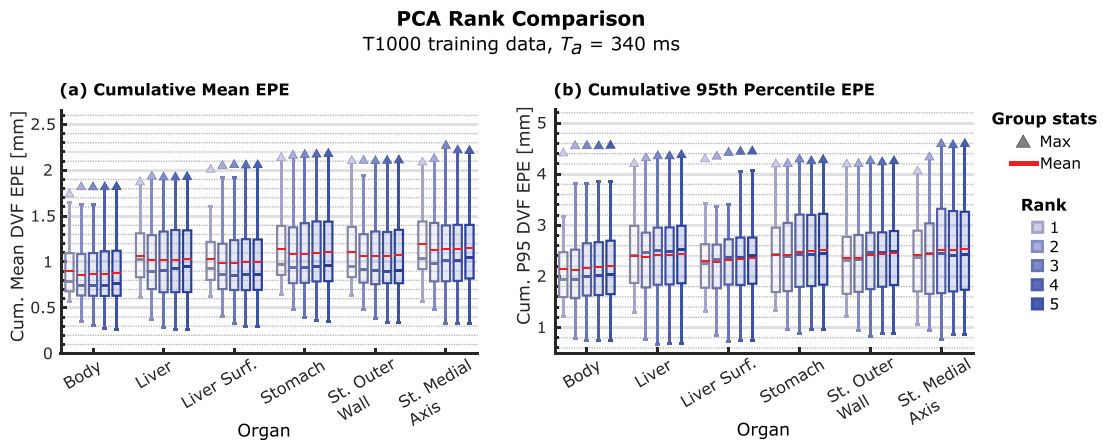

Figure 7. Boxplots comparing the effect of PCA rank on cumulative mean end-point error (EPE) (a, left) and cumulative 95th percentile EPE (b, right) across patient datasets. Box colors correspond to PCA ranks given in the legend (far right). Results are shown for test set 1, using T1000 training data and prediction horizon $T_a = 340$ ms.

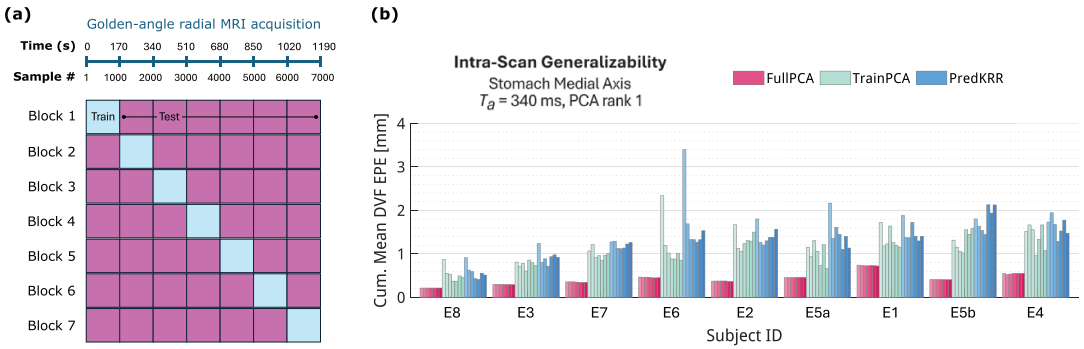

Figure 8. Analysis of intra-scan variability on training and breathing motion prediction. (a) The full 7000 sample (20 min) golden-angle radial MRI datasets were partitioned into blocks of 1000 sample (∼3 min) subsets used for training (pink), with the remaining data used for testing (blue). (b) Bar plots of the cumulative mean prediction end-point error (EPE) of the stomach medial axis. Bar colors denote reconstruction model (FullPCA, red; TrainPCA, green; PredKRR, blue). For each model, seven bars are shown, corresponding to blocks 1–7 (from left to right). All models used PCA rank 1, and PredKRR experiments used a prediction horizon $T_a = 340$ ms.

Table 3. Summary of intra-scan variability experiments. Mean and standard deviation (across blocks 1–7) of the mean DVF EPE of the stomach medial axis for different reconstruction models.

	Mean DVF EPE (mm) of stomach medial axis, Mean ± std. dev. across blocks		
Dataset ID	FullPCA	TrainPCA	PredKRR
E8	0.22 ± 0.0004	0.52 ± 0.17	0.58 ± 0.17
E3	0.30 ± 0.004	0.76 ± 0.08	0.92 ± 0.16
E7	0.35 ± 0.006	1.00 ± 0.12	1.21 ± 0.08
E6	0.46 ± 0.008	1.17 ± 0.53	1.70 ± 0.77
E2	0.37 ± 0.004	1.31 ± 0.21	1.42 ± 0.21
E5a	0.46 ± 0.003	1.01 ± 0.24	1.46 ± 0.36
E1	0.73 ± 0.007	1.34 ± 0.24	1.50 ± 0.22
E5b	0.41 ± 0.002	1.31 ± 0.23	1.80 ± 0.28
E4	0.55 ± 0.01	1.40 ± 0.28	1.63 ± 0.22
Group average ± std. dev.	0.43 ± 0.15	1.09 ± 0.30	1.36 ± 0.39

Table 4. Mean online inference times per test sample for six representative datasets spanning a range of cropped liver and body voxel counts. Timing experiments were performed using test set 1, with inference times measured independently for each of the 500 test samples per dataset. Reported values correspond to the mean computation time per sample for the EstKRR inference, PredKRR inference, rank-2 PCA DVF reconstruction, and their combined total online inference time. The final row reports the average across datasets.

Subj. ID	Cropped liver voxels	Body voxels	EstKRR (ms)	PredKRR (ms)	DVF recon (ms)	Total online inference (ms)
E1	51 581	1258 322	5.16	0.58	8.37	14.11
M3b	27 827	1128 501	4.22	0.38	7.67	12.27
E4	87 404	1108 765	7.53	0.40	7.82	15.75
E6	42 892	1087 606	4.46	0.38	7.27	12.11
E7	65 546	1425 597	5.86	0.39	9.66	15.91
E8	33 586	650 432	3.76	0.43	3.79	7.98
Average across datasets	51 473	1109 871	5.17	0.43	7.43	13.02

required to achieve accurate BMPs in abdominal structures of interest at different prediction horizons. Together, these results provide a quantitative benchmark for stomach-specific BMP performance and inform future investigations into real-time gastric motility management in MRgRT.

Building on prior work (Liu *et al* 2021), we employed a KRR-BMP framework that constructs patient- and fraction-specific motion models and forecasts future breathing-induced deformations via PCA coefficient prediction. While breathing DVFs in this study were derived using a specific hierarchical motion modeling framework (Zhang *et al* 2021), the KRR-BMP approach is model-agnostic and can be readily integrated with other breathing motion models. Using less than 90 s of training data and a prediction horizon of 340 ms, consistent with reported MRgRT system latencies, we achieved sub-voxel accuracy and mean prediction errors of under 1.5 mm within stomach-related structures in most datasets.

Compared to prior work that relied on about 5 min of training data and rank-1 PCA reconstructions (Liu *et al* 2021), we demonstrated that accurate predictions can be obtained with substantially shorter training stages. Training datasets of 1–3 min yielded comparable prediction accuracy while significantly reducing computational burden, particularly during hyperparameter tuning and kernel matrix construction. These improvements make the prediction workflow more amenable to the time constraints that accompany clinical applications.

There was an observable trade-off between training dataset size and prediction accuracy, most notably between training stages of 1 min and 3 min. Increasing training data acquisition time improved mean and upper bound errors, particularly for the stomach and its subregions, whereas shorter training windows introduced greater uncertainty and error spread both within and across patients. Nevertheless, with approximately 1.5 min of training data, prediction accuracy remained within clinically relevant bounds for most datasets, suggesting that KRR-BMP can support suitable prediction performance with limited training data when paired with appropriate quality assurance (QA) or outlier rejection strategies.

A slight but consistent increase in EPE was observed for test set 2 compared to test set 1 across all evaluated organs (figure 4 and table 2). One potential contributing factor is variability in contrast washout dynamics, which may not be fully captured by the current feature representation. Test set 2 comprised samples 501–1000 and used T1000 training data from samples 1001–2000, whereas test set 1 contained samples 1501–2000 with training data from samples 501–1500. Although the EstKRR intensity features were L2-normalized to reduce sensitivity to intensity scaling, subtle temporal intensity variations may persist after normalization and introduce mild dataset-dependent variability. It is possible that residual differences in contrast enhancement between the training and testing data contributed to the observed performance differences. Nevertheless, both test sets exhibited consistent trends and low overall prediction errors, supporting the robustness of the prediction framework.

The observed reduction in prediction error at the 2040 ms horizon (figure 6) likely reflects the periodic nature of respiratory motion rather than improved prediction model performance. A horizon of approximately 2 s corresponds to roughly one half of one observed respiratory cycle in our datasets, such that phase-aligned or cyclic repetition effects may contribute to reduced EPE at this prediction horizon. Therefore, this behavior should not be interpreted as improved predictive accuracy at longer horizons, but rather more likely a consequence of periodicity and temporal structure in the underlying respiratory motion. This effect is sensitive to respiratory cycle variability across subjects and should be interpreted cautiously when evaluating multi-step forecasting.

Extending the original KRR-BMP framework to predict multiple PCA coefficients revealed that a PCA rank of two yielded slight improvements in accuracy compared to rank-1 PCA predictions. Higher PCA ranks reduced median prediction errors, at the cost of large increases in inter-subject variability and upper bound errors (figure 7), likely due to mismatches between PCA bases derived from limited training data and ground truth bases derived from the full dataset. This highlights an inherent limitation of PCA-based forecasting when training data incompletely capture breathing variability. More detailed analysis of the two reference reconstruction models, TrainPCA and FullPCA, further illustrates the limitations imposed by finite training data (figure 8(b) and table 3). The ground-truth oracle, FullPCA, decomposes the full breathing motion DVFs using PCA and performs a rank-1 reconstruction. The population-average FullPCA reconstruction error of 0.43 mm demonstrates the inherently low-rank structure of breathing motion and supports the use of PCA for dimensionality reduction. In contrast, the training oracle, TrainPCA, derives its motion model from a 1000-sample training block and represents the optimal approximation of test-sample DVFs achievable using a PCA model learned from limited data. TrainPCA also corresponds to the hypothetical case of zero-error coefficient prediction by PredKRR. Consequently, differences between FullPCA and TrainPCA errors reflect modeling errors arising from limited training data, whereas differences between TrainPCA and PredKRR errors reflect inaccuracies introduced during prediction. For breathing motion of the stomach MA, the average performance gap between TrainPCA and FullPCA was approximately 0.7 mm, while the gap between PredKRR and TrainPCA was approximately 0.27 mm. These results indicate that errors introduced during coefficient prediction are substantially smaller than the inherent modeling errors caused by limited training data. Accurate compensation of respiratory motion is essential for studying gastric contractile motion, which occurs independent of breathing and at slower temporal frequencies. Respiratory-induced stomach displacements have been reported to exceed 10 mm in previous studies (Mostafaei *et al* 2018, Umezawa *et al* 2021, Zhang *et al* 2021). To address this, prior MRI studies of gastric motility included breathing motion compensation. However, KRR-BMP of the stomach, particularly its centerline and OW, have not been thoroughly investigated. Our results demonstrate that reliable respiratory motion prediction of the stomach and its centerline is feasible across a population of patients, supporting its role as a necessary precursor to real-time gastric motility characterization.

A key aspect of this work is the separation of respiratory-induced motion from intrinsic gastric motion, including filling dynamics and (peristaltic) contractile motion. The respiratory surrogate, defined as superior-inferior liver translation, is used solely for respiratory phase estimation and does not impose a rigid or anatomical coupling between liver and stomach motion. The subsequent breathing motion model is derived from whole-abdomen deformable registration of respiratory-resolved image volumes, such that stomach deformations are represented within a global respiratory motion subspace rather than being directly inferred from the surrogate signal.

A potential concern is that gastric motility occurring concurrently with respiration could contaminate the estimated breathing DVFs during DIR. In this work, several aspects of the reconstruction framework mitigate this effect. First, *k*-space data are retrospectively sorted according to the respiratory surrogate prior to reconstruction, which removes temporal coherence between respiratory phase and gastric contractile activity, thereby decorrelating peristaltic motion from the breathing trajectory. Second, the respiratory states used to construct canonical DVFs span several minutes (approximately 3–9 gastric contraction cycles), and the wide VS reconstruction window effectively averages out non-respiratory cyclic components.

Under the assumption that respiratory and gastric motions are largely unsynchronized, these procedures yield respiratory state images in which peristaltic motion is substantially temporally incoherent and therefore attenuated in the reconstructed breathing model. This is further supported by manual QA of reconstructed respiratory state images from full datasets (2000–7000 samples), which confirmed physiologically plausible breathing motion without visible structured non-respiratory artifacts. Additionally, preliminary comparisons (data not included here) between PCA models derived from full-abdomen and stomach-specific DVFs show low-rank approximation differences below 0.1 mm, suggesting that the dominant respiratory subspace is robust to residual gastric motion and appropriate for modeling breathing-induced stomach deformation. Future work could investigate whether localized stomach-derived respiratory surrogate signals further improve characterization of breathing-induced stomach deformation, particularly in the presence of substantial gastric filling or contractile activity.

Several practical considerations and limitations were identified. Analysis of the sample-level prediction EPEs (data not shown) indicated that prediction errors were consistently higher near the inhale phase, consistent with prior respiratory motion imaging studies (Stemkens *et al* 2016, Liu *et al* 2021, Huttinga *et al* 2022). These findings suggest a role for phase-aware prediction models and, more generally, for the integration of QA and outlier detection steps into the prediction framework. Incorporating

estimates of forecasting uncertainty could aid in on-the-fly QA, by enabling the identification of potentially erroneous model predictions. Huttinga *et al* (2023) introduced a Gaussian Process framework for real-time 3D motion estimation and uncertainty quantification in MRgRT, using an uncertainty-based rejection criterion to discard implausible predicted dynamics. They evaluated accuracy using voxel-wise EPE (the residual magnitude of each prediction) and demonstrated in simulations that applying the rejection criterion reduced 75th percentile errors below 1 mm, whereas the same experiments produced errors exceeding 6 mm without it. These results highlight the potential for integrated uncertainty estimation to provide effective QA mechanisms within respiratory prediction workflows. Future work will aim to develop and integrate QA and outlier detection steps into the prediction framework.

Contrast enhancement in the DCE-MRI data introduced time-varying intensity changes that could confound the intensity-based features in the EstKRR stage. However, signal intensities in most datasets reached equilibrium within the first 85 s, thus these early samples were excluded from our analysis. Prior to feature vector construction, voxel-intensity vectors were L2-normalized to reduce sensitivity of the kernel representation to global intensity scaling and emphasize relative spatial patterns rather than absolute signal magnitude or contrast dynamics. As a result, the learned features are expected to remain robust to slow intensity variations, including potential signal drift in non-contrast MRgRT settings.

Although this study was performed using contrast-enhanced MRI acquisitions, the proposed framework is expected to generalize well to standard non-contrast MRgRT imaging. In the absence of contrast dynamics, liver voxel intensities are expected to exhibit greater temporal stability, potentially providing more consistent intensity-based features for motion estimation. This expectation is further supported by the absence of significant differences in the cumulative mean and 95th percentile EPEs for either the breathing DVF estimation (EstKRR) or prediction (PredKRR) stages between the Eovist and Multihance cohorts (supplementary figure S1), suggesting that the intensity-based features are robust to differences in contrast agent pharmacokinetics once signal equilibrium has been reached. Consequently, performance under routine non-contrast imaging conditions would be expected to be at least comparable, although direct validation using non-contrast MR-Linac data remains an important area for future work. Alternative feature representations that are less sensitive to intensity variations may nevertheless further improve robustness and facilitate broader clinical deployment. Finally, it should be noted that the proposed framework was evaluated using a T1-weighted spoiled gradient-echo (GRE) GAR acquisition, which avoids the off-resonance banding artifacts commonly associated with balanced steady-state free precession cine imaging and would likewise not be expected when using comparable T1-weighted GRE acquisitions on an MR-Linac.

The primary error metric employed in this study was the EPE between predicted breathing DVFs and reference DVFs. The reference DVFs were obtained by applying the motion modeling framework (Zhang *et al* 2021) to the full dataset, using DIR of reconstructed breathing state image volumes. Consequently, these reference DVFs represent estimates of respiratory motion derived from the dynamic MRI data rather than absolute physical ground truth. Any inaccuracies introduced during image reconstruction, motion modeling, or DIR are therefore propagated into the reference data. Accordingly, the reported EPE should be interpreted as a measure of predictive agreement with the reference motion model rather than absolute anatomical accuracy. Despite this limitation, an absolute ground truth for internal abdominal deformations is currently not attainable *in vivo*. The adopted motion modeling framework has been extensively investigated in previous studies (Johansson *et al* 2016, 2018a, 2018b, 2021, Zhang *et al* 2021) and has been used in prior motion prediction and tracking studies (Zhang *et al* 2023b, Liu *et al* 2021, 2022), where it was shown to produce physiologically plausible respiratory motion that generalized across subjects.

To reduce computational burden, we introduced a simple liver voxel cropping strategy that significantly reduced EstKRR feature dimensionality while preserving strong respiratory surrogates near the top of the liver and right hemidiaphragm (Veldman *et al* 2023), which undergo large respiratory motion amplitudes (Keall *et al* 2006). Future work will investigate more elegant feature extraction approaches, such as PCA-compressed intensity features and randomized linear algebra techniques like matrix sketching (Oscanoa *et al* 2024, Plummer *et al* 2026), which may further improve efficiency and robustness. Another promising alternative is the use of k-space self-navigators extracted from GAR trajectories, which have been previously utilized for direct estimation of surrogate respiratory motion signals (Feng *et al* 2016) but have not yet been evaluated in the proposed prediction framework. Encoding additional information in the PredKRR feature vectors, such as the surrogate respiratory signal and/or numerical gradients of the PCA coefficient trajectories, could enhance expressiveness and robustness to noise while incurring little additional computation burden. Alternatively, a generalized radial basis function (RBF) kernel, $k^{\gamma}\left(\mathbf{x}_i, \mathbf{x}_j\right) = \exp\left(-\left\|\mathbf{x}_i - \mathbf{x}_j\right\|_2^{\gamma} / h\right)$, containing one additional parameter γ that controls the exponential distance scaling (Geifman *et al* 2020), could be included as an additional hyperparameter during

training cross-validation. Laplace kernels are specialized cases where $\gamma = 1$, thus the generalized RBF kernel may improve flexibility for patient-specific modeling. Future work will also compare the KRR-BMP framework to other prediction models, such as LSTM networks and their variants, which have been widely used for MRI- and CT-based motion prediction (Romaguera *et al* 2020b, Lombardo *et al* 2022, 2023a, 2023b, Numakura *et al* 2024, Weng *et al* 2024).

The measured online inference times (table 4) demonstrate the computational efficiency of the PCA- and KRR-based framework and its potential suitability for real-time MRgRT workflows. Across the evaluated datasets, the mean total online inference time was 13.02 ms per sample, substantially shorter than the 170 ms temporal resolution of the dynamic MRI sampling and within the range of reported MRgRT system latencies (200–500 ms) (Borman *et al* 2018). EstKRR inference time increased with the number of cropped liver voxels, reflecting the dependence of kernel evaluation on the feature vector dimension, while DVF reconstruction time increased with the number of body voxels, since the computational cost of the rank-r PCA reconstruction scales with the basis vector dimension. In contrast, PredKRR inference time remained below 0.6 ms for all datasets due to the low-dimensionality of the feature vectors and responses.

It should be noted that the reported computation times represent only the online inference stage and assume that the training and test feature vectors are already available in memory. The current framework requires an offline motion modeling step to generate the training data (figure 1), including estimation of the surrogate breathing signal, respiratory-resolved image reconstruction and deformable registration. The current implementation of the motion modeling stage can require up to 30 min and represents a limitation for fully real-time deployment of the motion prediction framework within the MRgRT workflow. Therefore, the reported inference times primarily demonstrate the computational efficiency and low-latency prediction capability of the PCA-KRR algorithm rather than the end-to-end latency of the complete motion modeling and prediction pipeline. Future work will investigate accelerated motion model generation strategies and optimized implementations capable of true real-time online deployment. The reported times were obtained using a CPU-based MATLAB implementation; additional reductions in inference time may be achievable through GPU acceleration or implementation using a high-performance programming language (e.g. C/C++ or Julia).

While the proposed framework demonstrated favorable geometric accuracy for KRR-BMP, geometric agreement alone does not necessarily translate to improved dosimetric accuracy for target volumes or surrounding organs at risk (Xian *et al* 2021). Future work should therefore investigate the dosimetric impact of the proposed motion prediction framework within MRgRT workflows to determine whether the observed geometric forecasting accuracy results in meaningful gains in target coverage and normal tissue sparing.

Finally, the retrospective nature of this study enabled comprehensive evaluation but relied on assumptions not fully representative of online clinical deployment, including precomputed motion models and unchanged patient anatomy between training and testing. Future work will focus on accelerating motion model generation, incorporating uncertainty estimation for online QA, enhancing feature representation, and comparing KRR-BMP to DL prediction models. Together, these developments will support translation of respiratory prediction frameworks toward real-time abdominal motion management and gastric motility modeling in MRgRT.

5. Conclusion

This study characterized the performance of KRR-BMP for abdominal organs, with a particular emphasis on the stomach and its clinically relevant substructures. Across 27 patient datasets, KRR-BMP achieved sub-voxel prediction accuracy at a 340 ms prediction horizon, with population-average mean endpoint errors below 1.3 mm for all evaluated organs and 95th percentile errors consistently below 3 mm. These results indicate that KRR-BMP can reliably capture short-term respiratory motion dynamics across heterogeneous abdominal anatomies.

Prediction accuracy was organ-dependent, with the lowest errors observed in the body and slightly higher errors in the stomach, including its OW and MA, reflecting the increased complexity of gastric motion dynamics. Nevertheless, error distributions were relatively small across subjects and test sets, underscoring the generalizability and stability of the proposed framework. Reducing the training set size led to gradual, predictable increases in error, particularly for gastric structures, highlighting the importance of sufficient training data for accurate prediction of more deformable organs. Importantly, even with reduced training data, population-level upper bound errors remained within clinically relevant

ranges. Using one and a half minutes of training data achieved volumetric prediction accuracies comparable to previous studies.

The analysis of prediction horizons and PCA rank revealed key trade-offs in model configuration. While longer prediction horizons generally increased error, population-level error spreads remained stable. The occasional improvement at longer horizons suggests subject-specific temporal regularities that merit further investigation. An optimal PCA rank of two achieved the best balance between accuracy and robustness, with higher ranks introducing increased inter-subject variability and upper bound errors.

Overall, these findings demonstrate KRR-BMP as a reliable approach for short-term respiratory motion prediction in abdominal MRgRT. The demonstrated accuracy in the stomach MA provides an important foundation for future integration of KRR-BMP with real-time gastric motility modeling, with potential applications in motion-aware image reconstruction, adaptive targeting, and treatment delivery. Future work will extend this framework to jointly model respiratory and gastric motion and to assess its performance in real-time MR-guided RT workflows.

Data availability statement

The data cannot be made publicly available upon publication because the cost of preparing, depositing and hosting the data would be prohibitive within the terms of this research project. The data that support the findings of this study are available upon reasonable request from the authors.

Supplementary data 1 available at: https://doi.org/10.1088/1361-6560/ae9117/data1.

Author contributions

Robert Jones 0000-0003-3543-2877
Conceptualization (equal), Formal analysis (lead), Investigation (lead), Methodology (equal), Software (equal), Visualization (lead), Writing – original draft (lead), Writing – review & editing (lead)

Lianli Liu 0000-0002-1095-0953
Methodology (equal), Software (equal), Validation (equal)

Daekeun You
Data curation (lead), Software (equal)

Jeffrey A Fessler 0000-0001-9998-3315
Conceptualization (equal), Methodology (equal), Supervision (equal), Writing – original draft (supporting), Writing – review & editing (supporting)

James M Balter 0000-0003-1513-4936
Conceptualization (equal), Investigation (equal), Methodology (equal), Resources (lead), Writing – original draft (supporting), Writing – review & editing (supporting)

References

Abbas H, Chang B and Chen Z Jay 2014 Motion management in gastrointestinal cancers *J. Gastrointest. Oncol.* **5** 223–35

Alam S, Veeraraghavan H, Tringale K, Amoateng 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

Ali Mirzapour S, Mazur T, Sharp G and Salari E 2019 Intra-fraction motion prediction in MRI-guided radiation therapy using Markov processes *Phys. Med. Biol.* **64** 195006

Antiga L, Piccinelli M, Botti L, Ene-iordache B, Remuzzi A and Steinman D A 2008 An image-based modeling framework for patient-specific computational hemodynamics *Med. Biol. Eng. Comput.* **46** 1097–112

Baker S, Scharstein D, Lewis J P, Roth S, Black M J and Szeliski R 2011 A database and evaluation methodology for optical flow *Int. J. Comput. Vis.* **92** 1–31

Balter J M, Ten Haken R K, Lawrence T S, Lam K L and Robertson J M 1996 Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing *Int. J. Radiat. Oncol. Biol. Phys.* **36** 167–74

Barger A V, Block W F, Toropov Y, Grist T M and Mistretta C A 2002 Time-resolved contrast-enhanced imaging with isotropic resolution and broad coverage using an undersampled 3D projection trajectory *Magn. Reson. Med.* **48** 297–305

Beatty P J, Nishimura D G and Pauly J M 2005 Rapid gridding reconstruction with a minimal oversampling ratio *IEEE Trans. Med. Imaging* **24** 799–808

Bernstein M A, King K F and Zhou X J 2004 *Handbook of MRI Pulse Sequences* (Elsevier Science & Technology) (available at: http://ebookcentral.proquest.com/lib/umichigan/detail.action?docID=312727)

Block K T *et al* 2014 Towards routine clinical use of radial stack-of-stars 3D gradient-echo sequences for reducing motion sensitivity *J. Korean Soc. Magn. Reson. Med.* **18** 87
Block K T and Uecker M 2011 Simple method for adaptive gradient-delay compensation in radial MRI *Proc. 19th Annual Meeting of ISMRM (Montreal, Canada)* vol 2816
Blum H 1967 A transformation for extracting new descriptors of shape *Models for the Perception of Speech and Visual Form* (Boston, Massachusetts, USA)11-14 Nov 1964 Wathen-Dunn W 362–80 (MIT Press)
Borman P T S, Raaymakers B W and Glitzner M 2019 ReconSocket: a low-latency raw data streaming interface for real-time MRI-guided radiotherapy *Phys. Med. Biol.* **64** 185008
Borman P T S, Tijssen R H N, Bos C, Moonen C T W, Raaymakers B W and Glitzner M 2018 Characterization of imaging latency for real-time MRI-guided radiotherapy *Phys. Med. Biol.* **63** 155023
Bourque A E, Bedwani S, Carrier J-F, Ménard C, Borman P, Bos C, Raaymakers B W, Mickevicius N, Paulson E and Tijssen R H N 2018 Particle filter–based target tracking algorithm for magnetic resonance–guided respiratory compensation: robustness and accuracy assessment *Int. J. Radiat. Oncol. Biol. Phys.* **100** 325–34
Brey W W and Narayana P A 1988 Correction for intensity falloff in surface coil magnetic resonance imaging *Med. Phys.* **15** 241–5
Burton A, Beveridge S, Hardcastle N, Lye J, Sanagou M and Franich R 2022 Adoption of respiratory motion management in radiation therapy *Phys. Imaging Radiat. Oncol.* **24** 21–29
Cardoso M J, Li W, Brown R, Ma N, Kerfoot E, Wang Y, Murrey B, Myronenko A, Zhao C and Yang D 2022 Monai: an open-source framework for deep learning in healthcare (arXiv:2211.02701)
Chandarana H, Block K T, Winfeld M J, Lala S V, Mazori D, Giuffrida E, Babb J S and Milla S S 2014 Free-breathing contrast-enhanced T1-weighted gradient-echo imaging with radial *k*-space sampling for paediatric abdominopelvic MRI *Eur. Radiol.* **24** 320–6
Cuppen J and van Est A 1987 Reducing MR imaging time by one-sided reconstruction *Magn. Reson. Imaging* **5** 526–7
Driever T, Hulshof M C C M, Bel A, Sonke J-J and van der Horst A 2023 Quantifying intrafractional gastric motion using auto-segmentation on MRI: deformation and respiratory-induced displacement compared *J. Appl. Clin. Med. Phys.* **24** e13864
Fast M F, Cao M, Parikh P and Sonke J-J 2024 Intrafraction motion management with MR-guided radiation therapy *Semin. Radiat. Oncol.* **34** 92–106
Fast M F, Nill S, Bedford J L and Oelfke U 2014 Dynamic tumor tracking using the Elekta Agility MLC *Med. Phys.* **41** 111719
Fayad H, Pan T, Pradier O and Visvikis D 2012 Patient specific respiratory motion modeling using a 3D patient's external surface *Med. Phys.* **39** 3386–95
Fayad H, Pan T, Roux C and Visvikis D 2010 A generic respiratory motion model for motion correction in PET/CT *IEEE Nuclear Science Symposium and Medical Imaging Conf. IEEE Nuclear Science Symposium & Medical Imaging Conf.* pp 2455–8 (available at: https://ieeexplore.ieee.org/abstract/document/5874229)
Feng L 2023 Live-view 4D GRASP MRI: a framework for robust real-time respiratory motion tracking with a sub-second imaging latency *Magn. Reson. Med.* **90** 1053–68
Feng L, Axel L, Chandarana H, Block K T, Sodickson D K and Otazo R 2016 XD-GRASP: golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing *Magn. Reson. Med.* **75** 775–88
Geifman A, Yadav A, Kasten Y, Galun M, Jacobs D and Basri R 2020 On the similarity between the laplace and neural tangent kernels *Proc. 34th Int. Conf. on Neural Information Processing Systems NIPS'20* (*Curran Associates Inc*)
Grimbergen G, Pötgens G G, Eijkelenkamp H, Raaymakers B W, Intven M P W and Meijer G J 2023 Feasibility of delivered dose reconstruction for MR-guided SBRT of pancreatic tumors with fast, real-time 3D cine MRI *Radiother. Oncol.* **182** 109506
Hähn D 2009 *Integration of the Vascular Modeling Toolkit in 3d Slicer* (University of Heidelberg Germany)
Hofmann T, Schölkopf B and Smola A J 2008 Kernel methods in machine learning *Ann. Stat.* **36** 1171–220
Hotelling H 1933 Analysis of a complex of statistical variables into principal components *J. Educ. Psychol.* **24** 417
Hu P, Li X, Liu W, Yan B, Xue X, Yang F, Ford J C, Portelance L and Yang Y 2022 Dosimetry impact of gating latency in cine magnetic resonance image guided breath-hold pancreatic cancer radiotherapy *Phys. Med. Biol.* **67** 055008
Huttinga N R F, Bruijnen T, van den Berg C A T and Sbrizzi A 2022 Real-time non-rigid 3D respiratory motion estimation for MR-guided radiotherapy using MR-MOTUS *IEEE Trans. Med. Imaging* **41** 332–46
Huttinga N R F, Bruijnen T, van den Berg C A T and Sbrizzi A 2023 Gaussian processes for real-time 3D motion and uncertainty estimation during MR-guided radiotherapy *Med. Image Anal.* **88** 102843
Jassar H *et al* 2022 Real-time motion monitoring using orthogonal cine MRI during MR-guided radiation therapy for liver cancer *Int. J. Radiat. Oncol. Biol. Phys.* **114** S80–1
Jassar H *et al* 2023 Real-time motion monitoring using orthogonal cine MRI during MR-guided adaptive radiation therapy for abdominal tumors on 1.5T MR-Linac *Med. Phys.* **50** 3103–16
Jayachandran P, Minn A Y, Van Dam J, Norton J A, Koong A C and Chang D T 2010 Interfractional uncertainty in the treatment of pancreatic cancer with radiation *Int. J. Radiat. Oncol. Biol. Phys.* **76** 603–7
Johansson A, Balter J M and Cao Y 2018b Abdominal DCE-MRI reconstruction with deformable motion correction for liver perfusion quantification *Med. Phys.* **45** 4529–40
Johansson A, Balter J M and Cao Y 2021 Gastrointestinal 4D MRI with respiratory motion correction *Med. Phys.* **48** 2521–7
Johansson A, Balter J and Cao Y 2018a Rigid-body motion correction of the liver in image reconstruction for golden-angle stack-of-stars DCE MRI *Magn. Reson. Med.* **79** 1345–53
Johansson A, Balter J, Feng M and Cao Y 2016 An overdetermined system of transform equations in support of robust DCE-MRI registration with outlier rejection *Tomography* **2** 188
Keall P J *et al* 2006 The management of respiratory motion in radiation oncology report of AAPM task group 76a) *Med. Phys.* **33** 3874–900
Keiper T D, Tai A, Chen X, Paulson E, Lathuilière F, Bériault S, Hébert F, Cooper D T, Lachaine M and Li X A 2020 Feasibility of real-time motion tracking using cine MRI during MR-guided radiation therapy for abdominal targets *Med. Phys.* **47** 3554–66
Kielbasa J E, Meeks S L, Kelly P, Willoughby T R, Zeidan O and Shah A P 2020 Evaluation of cine imaging during multileaf collimator and gantry motion for real-time magnetic resonance guided radiation therapy *J. Appl. Clin. Med. Phys.* **21** 178–87
Kim T, Lewis B C, Price A, Mazur T, Gach H M, Park J C, Cai B, Wittland E, Henke L and Kim H 2020 Direct tumor visual feedback during free breathing in 0.35 T MRgRT *J. Appl. Clin. Med. Phys.* **21** 241–7
Kim T, Lewis B, Lotey R, Barberi E and Green O 2021 Clinical experience of MRI4D QUASAR motion phantom for latency measurements in 0.35 T MR-Linac *J. Appl. Clin. Med. Phys.* **22** 128–36

Klinder T and Lorenz C 2012 Respiratory motion compensation for image-guided bronchoscopy using a general motion model *2012 9th IEEE Int. Symp. on Biomedical Imaging (ISBI) 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* pp 960–3 (available at: https://ieeexplore.ieee.org/abstract/document/6235716)
Li R, Lewis J H, Jia X, Gu X, Folkerts M, Men C, Song W Y and Jiang S B 2011a 3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy *Med. Phys.* **38** 2783–94
Li R, Lewis J H, Jia X, Zhao T, Liu W, Wuenschel S, Lamb J, Yang D, Low D A and Jiang S B 2011b On a PCA-based lung motion model *Phys. Med. Biol.* **56** 6009
Li Y, Li Z, Zhu J, Li B, Shu H and Ge D 2023 Online prediction for respiratory movement compensation: a patient-specific gating control for MRI-guided radiotherapy *Radiat. Oncol.* **18** 149
Lin H, Zou W, Li T, Feigenberg S J, Teo B-K-K and Dong L 2019 A super-learner model for tumor motion prediction and management in radiation therapy: development and feasibility evaluation *Sci. Rep.* **9** 14868
Liu L, Johansson A, Cao Y, Lawrence T S and Balter J M 2021 Volumetric prediction of breathing and slow drifting motion in the abdomen using radial MRI and multi-temporal resolution modeling *Phys. Med. Biol.* **66** 175028
Liu L, Shen L, Johansson A, Balter J M, Cao Y, Chang D and Xing L 2022 Real time volumetric MRI for 3D motion tracking via geometry-informed deep learning *Med. Phys.* **49** 6110–9
Lombardo E *et al* 2023a Experimental comparison of linear regression and LSTM motion prediction models for MLC-tracking on an MRI-Linac *Med. Phys.* **50** 7083–92
Lombardo E *et al* 2023b Evaluation of real-time tumor contour prediction using LSTM networks for MR-guided radiotherapy *Radiother. Oncol.* **182** 109555
Lombardo E *et al* 2024 Real-time motion management in MRI-guided radiotherapy: current status and AI-enabled prospects *Radiother. Oncol.* **190** 109970
Lombardo E, Rabe M, Xiong Y, Nierer L, Cusumano D, Placidi L, Boldrini L, Corradini S, Niyazi M and Belka C 2022 Offline and online LSTM networks for respiratory motion prediction in MR-guided radiotherapy *Phys. Med. Biol.* **67** 095006
Lu K, Liu Z, Jaffey D, Wo J M, Mosier K M, Cao J, Wang X and Powley T L 2022 Automatic assessment of human gastric motility and emptying from dynamic 3D magnetic resonance imaging *Neurogastroenterol. Motil.* **34** e14239
Martini N, Santarelli M F, Giovannetti G, Milanesi M, De Marchi D, Positano V and Landini L 2010 Noise correlations and SNR in phased-array MRS *NMR Biomed.* **23** 66–73
McClelland J R, Hawkes D J, Schaeffter T and AP K 2013 Respiratory motion models: a review *Med. Image Anal.* **17** 19–42
Modat M, Ridgway G R, Taylor Z A, Lehmann M, Barnes J, Hawkes D J, Fox N C and Ourselin S 2010 Fast free-form deformation using graphics processing units *Comput. Methods Programs Biomed.* **98** 278–84
Molitoris J K, Diwanji T, Snider J W 3, Mossahebi S, Samanta S, Badiyan S N, Simone C B 2 and Mohindra P 2018 Advances in the use of motion management and image guidance in radiation therapy treatment for lung cancer *J. Thorac. Dis.* **10** S2437–50
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 ed Q Zhang *PLoS One* **13** e0205917
Nataraj G, Nielsen J-F, Scott C and Fessler J A 2018 Dictionary-free MRI PERK: parameter estimation via regression with kernels *IEEE Trans. Med. Imaging* **37** 2103–14
Ng J, Gregucci F, Pennell R T, Nagar H, Golden E B, Knisely J P, Sanfilippo N J and Formenti S C 2023 MRI-LINAC: A transformative technology in radiation oncology *Front. Oncol.* **13**
Numakura K *et al* 2024 Application of motion prediction based on a long short-term memory network for imaging dose reduction in real-time tumor-tracking radiation therapy *Phys. Med.* **125** 104507
Oscanoa J A, Ong F, Iyer S S, Li Z, Sandino C M, Ozturkler B, Ennis D B, Pilanci M and Vasanawala S S 2024 Coil sketching for computationally efficient MR iterative reconstruction *Magn. Reson. Med.* **91** 784–802
Pham J, Harris W, Sun W, Yang Z, Yin F-F and Ren L 2019 Predicting real-time 3D deformation field maps (DFM) based on volumetric cine MRI (VC-MRI) and artificial neural networks for on-board 4D target tracking: a feasibility study *Phys. Med. Biol.* **64** 165016
Piccinelli M, Veneziani A, Steinman D A, Remuzzi A and Antiga L 2009 A framework for geometric analysis of vascular structures: application to cerebral aneurysms *IEEE Trans. Med. Imaging* **28** 1141–55
Plummer J W, Daudé P, Tsakirellis A, Taylor J, Moss J, Ramasawmy R, Javed A and Campbell-Washburn A E 2026 Coil sketching for fast and efficient 4D lung MRI reconstruction *Magn. Reson. Med.* **95** 2241–53
Renner A, Gulyas I, Buschmann M, Heilemann G, Knäusl B, Heilmann M, Widder J, Georg D and Trnková P 2024 Explicitly encoding the cyclic nature of breathing signal allows for accurate breathing motion prediction in radiotherapy with minimal training data *Phys. Imaging Radiat. Oncol.* **30** 100594
Romaguera L V, Plantefève R and Kadoury S 2020a Spatiotemporal Motion Prediction in Free-breathing Liver Scans via a Recurrent Multi-scale Encoder Decoder Online (arXiv:2005.13071)
Romaguera L V, Plantefève R, Romero F P, Hébert F, Carrier J-F and Kadoury S 2020b Prediction of in-plane organ deformation during free-breathing radiotherapy via discriminative spatial transformer networks *Med. Image Anal.* **64** 101754
Salari E, Wang J, Wynne J F, Chang C, Wu Y and Yang X 2024 Artificial intelligence-based motion tracking in cancer radiotherapy: a review *J. Appl. Clin. Med. Phys.* **25** e14500
Schölkopf B and Smola A J 2002 *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond* (MIT Press)
Seregni M, Paganelli C, Lee D, Greer P B, Baroni G, Keall P J and Riboldi M 2016 Motion prediction in MRI-guided radiotherapy based on interleaved orthogonal cine-MRI *Phys. Med. Biol.* **61** 872–87
Shaffer N, Dresner A, Ying Q, Alberts E, Kruiskamp M, Caster J, Hyer D, Snyder J and St-aubin J 2025 Validating a T1-weighted cine MRI for a 1.5 T MR-Linac with temporal resolution appropriate for respiratory motion *Front. Oncol.* **15** 1575001
Siddiq S, Murray V, Tyagi N, Borman P, Gui C, Crane C, Wu C and Otazo R 2024 MR signature matching (MRSIGMA) implementation for true real-time free-breathing volumetric imaging with sub-200 ms latency on an MR-Linac *Magn. Reson. Med.* **92** 1162–76
Stemkens B, Tijssen R H, De Senneville B D, Lagendijk J J and Van Den Berg C A 2016 Image-driven, model-based 3D abdominal motion estimation for MR-guided radiotherapy *Phys. Med. Biol.* **61** 5335
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
Sung J, Choi Y, Kim J, Kim J W and Kim J 2025 Development of in-house software to process real-time cine magnetic resonance images acquired during 1.5 T MR-guided radiation therapy *Sci. Rep.* **15** 29515
Umezawa R, Wakita A, Katsuta Y, Ito Y, Nakamura S, Okamoto H, Takahashi K, Inaba K, Murakami N and Igaki H 2021 A pilot study of synchronization of respiration-induced motions in the duodenum and stomach for the primary tumor in radiation therapy for pancreatic cancer using 4-dimensional computed tomography *Adv. Radiat. Oncol.* **6** 100730

van Harten L D, De jonge C S, Struik F, Stoker J and Išgum I 2025 Quantitative analysis of small intestinal motility in 3D cine-MRI using centerline-aware motion estimation *J. Magn. Reson. Imaging* **61** 1956–66

Veldman J K, van Kesteren Z, Gunwhy E R, Parkes M J, Stevens M F, van den Aardweg J G, van Tienhoven G, Bel A and van Dijk I W E M 2023 Accuracy of abdominal organ motion estimation in radiotherapy using the right hemidiaphragm top as a surrogate during prolonged breath-holds quantified with MRI *Med. Phys.* **50** 3299–310

Weng J, Bhupathiraju S H V, Samant T, Dresner A, Wu J and Samant S S 2024 Convolutional LSTM model for cine image prediction of abdominal motion *Phys. Med. Biol.* **69** 085024

Wysocka B, Kassam Z, Lockwood G, Brierley J, Dawson L A, Buckley C A, Jaffray D, Cummings B, Kim J and Wong R 2010 Interfraction and respiratory organ motion during conformal radiotherapy in gastric cancer *Int. J. Radiat. Oncol. Biol. Phys.* **77** 53–59

Wysocka B, Moseley J, Brock K, Lockwood G, Wilson G, Simeonov A, Haider M A, Menard C, Bissonnette J-P and Dawson L A 2014 Assessment of nonrespiratory stomach motion in healthy volunteers in fasting and postprandial states *Pract. Radiat. Oncol.* **4** 288–93

Xian L, Li G, Xiao Q, Li Z, Zhang X, Chen L, Hu Z and Bai S 2021 Clinically oriented target contour evaluation using geometric and dosimetric indices based on simple geometric transformations *Technol. Cancer Res. Treat.* **20** 15330338211036325

Zhang Q, Pevsner A, Hertanto A, Hu Y-C, Rosenzweig K E, Ling C C and Mageras G S 2007 A patient-specific respiratory model of anatomical motion for radiation treatment planning *Med. Phys.* **34** 4772–81

Zhang W *et al* 2023a Real-time, volumetric imaging of radiation dose delivery deep into the liver during cancer treatment *Nat. Biotechnol.* **41** 1160–7

Zhang Y, Cao Y, Kashani R, Lawrence T S and Balter J M 2023b Real-time prediction of stomach motions based upon gastric contraction and breathing models *Phys. Med. Biol.* **68** 015001

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