Evaluation of Liver Respiratory Biomechanics using 4D-MRI

by

Xiao Liang

Graduate Program in Medical Physics
Duke University

Date:_____

Approved:

_____

Jing Cai, Supervisor

_____

Fang-Fang Yin, Chair

_____

James R MacFall

_____

Mustafa R Bashir

Thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the
Graduate Program in Medical Physics
in the Graduate School of
Duke University

2014

ABSTRACT

Evaluation of Liver Respiratory Biomechanics using 4D-MRI

by

Xiao Liang

Graduate Program in Medical Physics
Duke University

Date:_____

Approved:

_____
Jing Cai, Supervisor

_____
Fang-Fang Yin, Chair

_____
James R MacFall

_____
Mustafa R Bashir

An abstract of a thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the
Graduate Program in Medical Physics
in the Graduate School of
Duke University

2014

Abstract

Purpose: It is of clinical interest to study liver deformation during breathing by applying deformable image registration (DIR) on respiratory correlated four-dimensional (4D) images. This study aims to evaluate and compare the accuracy of DIR-derived liver deformation based on 4D computed tomography (CT) and 4D magnetic resonance imaging (MRI).

Methods: 4D CT, 4D MRI and cine magnetic resonance (MR) images of liver region were acquired from 5 patients with liver cancer under an IRB-approved protocol. ROIs containing tumors in each patient were tracked multiple times (3~5) in cine MR images. The trajectories from tracking, covering several breathing cycles, were converted to trajectories in one breathing cycle through phase sorting. The average phase sorted trajectories for each patient were used as reference trajectories after manual verification. Deformation vector fields (DVFs) from 4D CT and 4D MRI were generated via DIR implemented in Velocity AI. To enable comparison between DVFs and reference tumor trajectories, deformation vectors from each frame were linked together, forming DVF-based trajectories at each voxel. All DVF-based trajectories within each ROI were averaged to represent tumor motion. The single-phase difference, the trajectory difference and the correlation coefficient between each pair of DVF-based trajectory and reference trajectory were calculated. Wilcoxon signed-rank tests were conducted to

determine whether there was significant difference between the single-phase differences, the trajectory differences and the correlation coefficients for 4D CT and 4D MRI.

Results: In the superior-inferior (SI) direction, 4D CT trajectories exhibit smaller trajectory differences (traj. diff.) in millimeters on average (traj. diff. (mm)= 2.09±0.75mm) but larger trajectory differences in number of voxels (traj. diff. (voxels)= 0.87±0.29) and smaller correlation coefficients (c. c. = 0.89±0.09) than 4D MRI trajectories (traj. diff. (mm)= 2.23±1.46mm, traj. diff. (voxels)= 0.45±0.29, c. c. = 0.93±0.06) whereas 4D MRI (traj. diff. = 1.09±1.23mm, traj. diff. (voxels)= 0.60±0.65, c. c. = 0.59±0.30) surpasses 4D CT (traj. diff. = 1.30±1.36mm, traj. diff. (voxels)= 1.02±1.07, c. c. = 0.15±0.64) in every metric in the right-left (RL) direction. In the anterior-posterior (AP) direction, 4D MRI trajectories have smaller trajectory differences in millimeters (traj. diff. (mm) = 1.11±0.70mm) and smaller trajectory differences in voxels (traj. diff. (voxels) = 0.61±0.36) but slightly smaller correlation coefficients (c. c. = 0.72±0.26) than 4D CT trajectories (traj. diff. (mm) = 1.47±0.63mm, traj. diff. (voxels) = 1.15±0.50, c. c. = 0.77±0.26). A trend that the trajectory differences in voxels for 4D MRI are smaller than those for 4D CT in every direction has been observed, though the results of Wilcoxon signed-rank tests do not support there is any significant difference between the accuracy of DVFs from 4D CT and 4D MRI.

Conclusion: We have implemented a novel approach for evaluating accuracy of DVFs based on 4D imaging for studying liver deformation. Current results indicate that the accuracy of DVFs from 4D CT and 4D MRI are comparable. Trends suggesting that the DVF from 4D MRI can be potentially more accurate than the DVF from 4D CT have been observed. Further study on more patients is warranted to determine whether there is significant difference between 4D CT and 4D MRI and to what degree the accuracy of the DVF from 4D MRI can be improved.

Contents

List of Tables

List of Figures

List of Abbreviations

3D – three-dimensional

4D – four-dimensional

AP – Anterior-Posterior

c. c. – correlation coefficient

CT - Computed Tomography

DIR – Deformable Image Registration

DVF – Deformation Vector Field

EBRT – External Beam Radiation Therapy

EOI - End-of-Inhalation

GTV – Gross Tumor Volume

GUI - Graphical User Interface

MR - Magnetic Resonance

MRI - Magnetic Resonance Imaging

OAR – Organ-at-risk

ph. diff. – single-phase difference

RL – Right-Left

ROI – Region of Interest

RPM - Real-time Position Management™

SI – Superior-Inferior

SNR - Signal-to-Noise Ratio

TPS - Treatment Planning System

traj. diff. – trajectory difference

Acknowledgements

I would like to express my gratitude to Dr. Jing Cai for his insightful comments and patient guidance for this study. I would also like to thank the other members of my thesis committee, Dr. Fang-Fang Yin, Dr. James MacFall, and Dr. Mustafa Bashir, for their time and effort. I would like to give a special thanks to Juan Yang for her help throughout my work and specifically in DVF generation. Last but not least, I would like to thank all the students in our laboratory for their constructive suggestions and perspectives.

1. Introduction

1.1 Respiratory Motion – A Challenge in External Beam Radiation Therapy

External beam radiation therapy (EBRT) utilizes ionizing radiation beams, most commonly photon and electron beams, delivered from outside the patient body to kill cancer cells. Tumor control probability (TCP), a metric for the efficacy of radiation therapy, can be calculated using the following equation,

$$TCP = e^{-\sum_{i=1}^{N} e^{-f\left(\alpha d_i + \beta d_i^2\right)}}$$

where N=initial number of cancer cells, f=number of fractions, d_i=dose to cancer cell *i*, α and β are parameters in the linear quadratic model for radiation effects[1-4]. Ionizing radiation is toxic to not only cancer cells but also normal cells. Concerns about excessive cell killing to normal cells prevent physicians from prescribing higher dose to the tumor, especially when organs-at-risk (OARs) are near or overlapping the target. Normal tissue complication probability (NTCP) serves as a metric for toxicity of radiation to normal tissues, which is calculated using the following equation,

$$p(EUD) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{EUD} e^{-\frac{(x-TD_{50})^2}{2\sigma^2}} dx$$

where

$$EUD = \left(\frac{\sum_i V_i d_i^{1/n}}{\sum_i V_i}\right)^n$$

$$\sigma = mTD_{50}$$

TD_{50}=tolerance dose for 50% probability of injury, d_i=dose to cancer cell *i* or region *i*, V_i=volume at dose d_i, m and n are tissue specific parameters in Lyman-Kutcher-Burman model[1, 5, 6].

As implied in the above equations, dose distribution is crucial for evaluation of the outcomes of radiation therapy. However, it is much more complicated to calculate the accumulated dose for targets and normal tissues that keep moving during beam delivery than for static ones. The main barrier is the lack of an accurate method to map dose distribution in a specific phase to the reference phase. Theoretically, if motion trajectories of each region that is sufficiently small are known, dose warping[7-13] can be accomplished. Although the breath-hold approach has been used to circumvent the problem by freezing the motion of moving targets and normal tissues, a large number of patients still have to receive free-breathing and phase-gating radiation therapy, during which targets and normal tissues keep moving. Therefore, it is of clinical interest to obtain quantitative motion information at different phases for accurate calculation of accumulated dose distribution.

1.2 Overview of Four-dimensional (4D) Computed Tomography (CT), 4D Magnetic Resonance Imaging (MRI), and Cine MRI

Various techniques have been developed to image the respiratory motion of tumors and normal tissues by adding a time dimension to or improving the temporal resolution of existing imaging modalities. Images obtained via these techniques contain motion information that, if properly extracted, can be used for dose warping.

4D CT has been widely used for evaluation of respiratory motion in clinical radiation therapy[14-19]. 4D CT works in cine mode, meaning that, instead of one image at each slice, multiple images will be acquired at each slice location for a minimal duration of a respiratory cycle. Each image will be subsequently assigned to a frame according to the phase in a respiratory cycle, at which it is acquired. Images from all slices in the same frame form a volume and anatomic changes can be seen between different frames. A unique advantage of 4D CT is that it provides information of electron density, which is necessary for dose calculation.

In contrast to the well-accepted and widely-used 4D CT technique, 4D MRI[20-23] is relatively new and not yet a routine procedure for motion evaluation. However, the role of 4D MRI in 4D radiation therapy is expected to be promising because it provides motion information and meanwhile retains the advantages of conventional MRI including high soft-tissue contrast and absence of radiation hazard. High soft-tissue contrast may be beneficial especially in the abdominal region where 4D CT images tend to exhibit a poor contrast. Absence of radiation hazard allows prolonged acquisition, which may lead to better image quality.

Currently, 4D MRI can be implemented in either cine mode or sequential mode. Cine mode acquisition is similar to the acquisition method used in 4D CT. Images at a slice location are acquired repeatedly for no shorter than a respiratory cycle and are subsequently sorted into different frames based on their phases. Sequential mode is also

commonly used, especially for those sequences that do not support cine mode acquisition. In sequential acquisition, images are acquired from the first slice location to the last and the acquisition will be repeated from the first slice location again until sufficient images have been collected for every slice location. The images will be sorted into corresponding frames by matching the time stamps of the images with the associated respiratory curve. Sequential mode typically takes longer than cine mode, implying that redundant information is acquired. This redundancy makes sequential mode more robust to irregular breathing. Removal of images acquired during irregular breathing will render the 4D MRI reconstruction impossible if the raw images are acquired in cine mode but is not likely to cause the same consequence if the raw images are acquired in sequential mode.

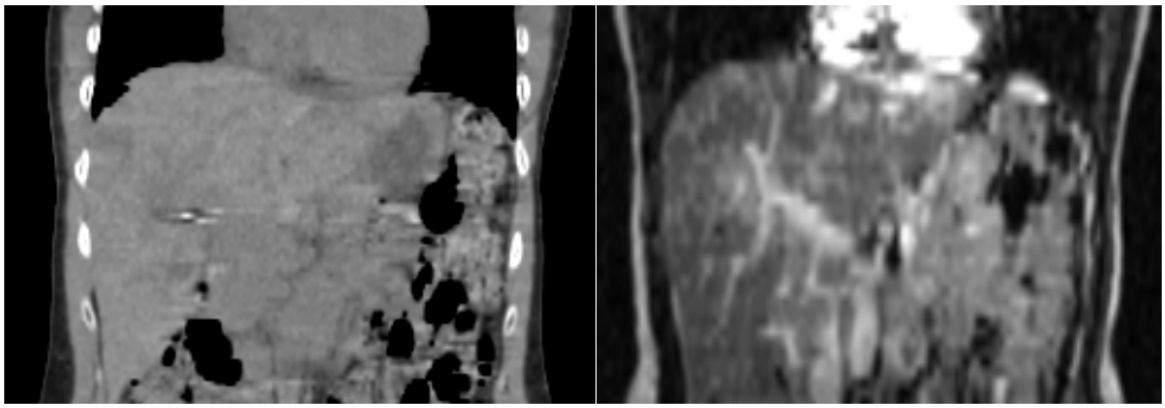

Figure 1: Coronal images of 4D CT (left) and 4D MRI (right) at the same location.

Cine magnetic resonance (MR) images are acquired by repeatedly imaging at the same location. The acquisition of cine MR images is actually the basic procedure in cine

mode 4D MRI acquisition but can extend to multiple respiratory cycles. The frame rate of cine MRI typically ranges from 3 to 7 frames per second. Because phase sorting is not involved in cine MRI, cine MR images are free from artifacts caused by irregular breathing. Cine MR images that cutting through the center of target are commonly taken in the sagittal, coronal, and axial views for motion evaluation.

1.3 The Displacement Vector Field (DVF)

The DVF, representing a voxel-based correlation between two volumes or frames, is introduced to describe deformation between the primary and the secondary volumes. As of now, the DVF can be generated via a variety of commercially available software and open-source deformable image registration (DIR) programs. Typically, displacement vectors are defined at each voxel in the primary volume and end at the most probable positions that voxels in the primary volume correspond to in the secondary volume. Since deformation in anatomy observed in 4D images essentially results from the respiratory motion, in theory, the DVF represents how each voxel in the primary volume move from the phase of the primary volume to the phase of the secondary volume. Therefore, the DVF can be used as quantitative motion information for accumulated dose calculation.

Because images are based on voxels with fixed physical size whereas tissues may change in shape, the anatomy represented in a single voxel in a phase can expand to several voxels in another phase and tissue occupying multiple voxels in a phase can

shrink to only one voxel in another phase. Due to this characteristic, it is noteworthy that the correlation described by the DVF is not voxel-to-voxel or readily invertible.

The DVF from 4D CT images has been extensively investigated[24-28]. With the availability of 4D MRI techniques, it is warranted to study the DVF from 4D MRI, especially in the abdominal region, considering the improved soft-tissue contrast in 4D MRI images.

1.4 Current Approaches to DVF Evaluation

Accuracy of the DVF should be scrutinized before conclusions are made from calculation based on the DVF. A variety of techniques[29-37], which can be categorized into two approaches, phantom and human subject, have been proposed to evaluate the accuracy of DVFs.

It is common to use digital and physical phantoms for DVF evaluation. Easy access to the ground truth, or the actual displacements, and the controllability of motion are beneficial. The phantoms can also tolerate virtually any amount of radiation dose without raising concerns about overdose, enabling high quality images to be acquired for DVF generation. Limitations of phantoms are as obvious. Most phantoms, especially physical phantoms, are not likely to contain as detailed anatomy as in human body. Tissues are often assumed to be homogeneous, lacking internal variation in density, composition, elasticity, etc. This homogeneity naturally leads to motion homogeneity,

meaning that an organ tends to move as a whole as a rigid body, which is not the case in the human body.

Patient or volunteer study is preferred for comprehensive evaluation. Implanted markers, anatomical landmarks, and contrast agents such as hyperpolarized gas have been used to provide the ground truth. Displacements derived from DIR algorithms are compared with displacements obtained from tracking or measurement in images. Limitations still present in these approaches. Implanting markers is invasive and may not be justifiable for health volunteers and many patients. Land markers may not be suitable for images with low quality such as low contrast, high noise. For instance, the liver in 4D CT images tends to be homogeneous without many landmarks. The major limitation of the hyperpolarized method is its limited application to the lungs.

1.5 Aims of This Study

This study aims at evaluating the accuracy of liver DVFs from 4D CT and 4D MRI. 4D CT has been long developed for assessing internal body motion in the lung region. It was not applied in liver region until recently[38, 39]. 4D MRI is expected to provide better contrast and motion information in the liver region. In 4D radiation therapy, the DVF has been proposed to be used for dose warping between different frames for calculation of accumulated dose. This potential application, however, is not yet widely accepted because of the difficulty in verifying DVFs from clinical 4D CT and

4D MRI images due to lack of ground truth. In this study, cine MRI images are used as reference for evaluation of DVFs from 4D CT and 4D MRI in the liver region.

The underlying principle that enables cine MR images to serve as reference is that they contain respiratory motion information, although only available for voxels in a few individual slices, typically three orthogonal slices cutting through the center of tumor. In addition, because ideally the deformation information obtained from DIR should reflect the actual respiratory motion information, the motion information in cine MR images can be utilized to evaluate the accuracy of DVFs. In practice, both DVFs and cine MR images need to be converted to a common form before evaluation, which is 10-phase single-cycle tumor trajectory in this study.

2. Materials and Methods

2.1 Patient Selection

Since this study aims at evaluating the accuracy of liver DVFs from 4D CT and 4D MRI using cine MR images as reference, an eligible patient should have 4D CT, 4D MRI, and cine MR images of the liver. As of now, 4D MRI is not yet a standard imaging procedure for patients with liver cancer. The 4D MR images used in this study are from patients accrued in an IRB-approved protocol for a 4D MRI study.

Because tumor motion trajectories obtained directly from tracking in cine MR images would be used as reference trajectories, tumors in cine MR images had to be of good contrast for the tracking algorithm to work. Among all patients with complete image data, those with visually identifiable tumors in cine MR images were tested for tracking regardless the type and the size of tumor. Patients with reasonable trajectories from tracking in cine images were finally enrolled in this study. Reasonability of trajectories from tracking in cine images was examined case by case. Specific reasons for exclusion of certain patients will be explained in the Results section.

Ideally, tumor visibility in 4D CT and 4D MR images should also be counted as one of the criteria for eligibility check. However, it is not uncommon that, even for the same patient, images of different modalities do not exhibit consistently high image quality. Considering the already small size of the pool from which patients were chosen from and the existence of alternative approach to localize tumors of poor contrast in 4D

CT and 4D MR images, tumor visibility in 4D MR and 4D cine was not checked for patient inclusion and exclusion.

2.2 Image Acquisition

4D CT images used in this study were acquired in cine mode, using a phase gating protocol at 120kVp, on a GE Medical Systems LightSpeed RT® CT scanner and a Siemens Biograph™ 40 PET/CT scanner. The 4D CT images were generated using phase sorting by matching the time stamp of each image with the synchronized respiratory curve recorded by a Varian® Real-time Position Management™ (RPM) system.

4D MR images used in this study were acquired in cine mode, using the FIESTA sequence on a GE Medical Systems Signa HDxt 1.5T MR scanner. FIESTA, which stands for Fast Imaging Employing Steady-state Acquisition is a GE developed sequence that uses the T2 steady state contrast mechanism to provide high signal-to-noise ratio (SNR) images with strong signal from fluid tissues while suppressing background tissue for contrast and anatomic detail of small structures. In addition, the ultra-short TR and TE enable extremely short acquisition times, which satisfy the requirement for temporal resolution of 4D MRI. The body area, defined as the number of voxels inside the skin, was used as the internal surrogate of respiration and the phase of each image was determined by the body area in the image.

Cine MR images were acquired using the FIESTA sequence on a GE Medical Systems Signa HDxt 1.5T MR scanner. Detailed parameters are listed in Table.2.

Table 1: Voxel size of 4D CT and 4D MR images used in this study.

Patient	Modality	In-plane Resolution (mm)	Slice Thickness (mm)
1	4D CT	1.270	2.500
	4D MRI	1.875	5.000
2	4D CT	1.270	2.500
	4D MRI	1.406	5.000
3	4D CT	1.367	2.000
	4D MRI	1.875	5.000
4	4D CT	1.270	2.500
	4D MRI	1.875	5.000
5	4D CT	1.270	2.500
	4D MRI	1.875	5.000

Table 2: Parameters of cine MR images used in this study.

Patient	Direction	In-plane Resolution (mm)	Slice Thickness (mm)	Frame Rate (frames/second)	No. of Frames
1	Coronal	1.875	5.000	2.53	180
	Sagittal	1.875	5.000	4.07	180
2	Coronal	1.406	5.000	2.30	180
	Sagittal	1.875	5.000	4.07	180
3	Coronal	1.875	5.000	2.84	100
	Sagittal	1.875	5.000	2.84	100
4	Coronal	1.875	5.000	2.84	100
	Sagittal	1.875	5.000	3.47	100
5	Coronal	1.875	5.000	2.44	180
	Sagittal	1.875	5.000	3.03	180

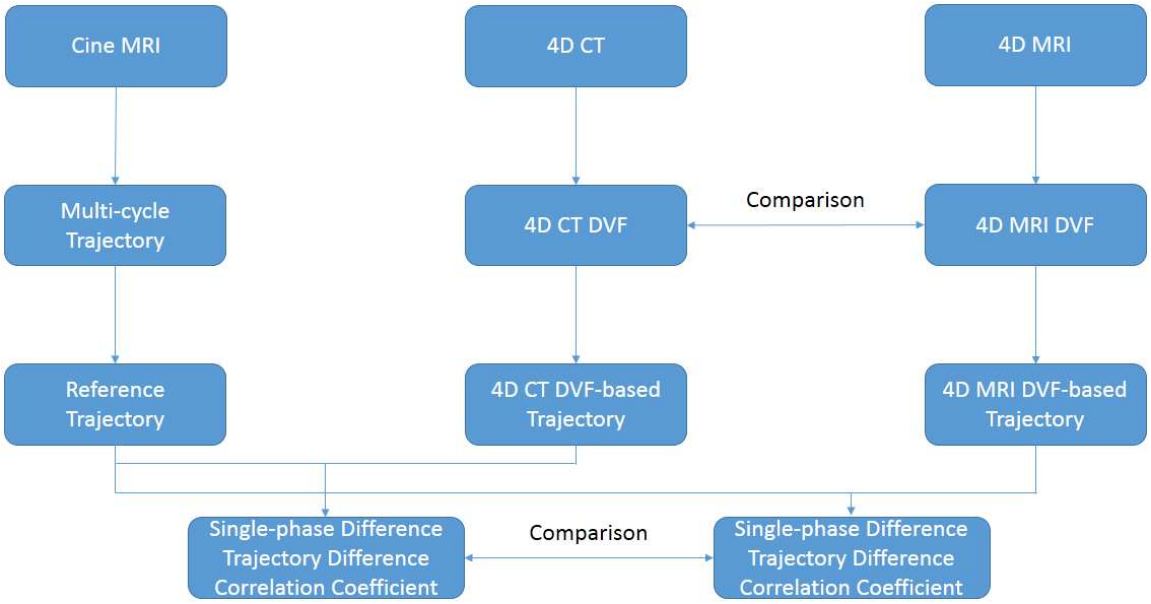

Figure 2: The study design. The reference trajectories were acquired by tracking tumors in the cine MR images. DVF-based trajectories were generated from 4D CT and 4D MRI DVFs.

2.3 Generation of Reference Tumor Trajectories from Tracking in Cine MR Images

2.3.1 Software

Tracking for tumor trajectories is implemented via an in-house developed tracking graphical user interface (GUI) based on a cross-correlation image registration algorithm. After manual selection of a region of interest (ROI), the program will determine the displacement of the ROI in each frame, which shifts the ROI from the position in the reference (the first frame) to the position where cross-correlation is maximized in the current frame. Programs for post-processing including peak detection and phase sorting were developed.

2.3.2 ROI selection

Tumors in cine MR images were localized visually. Since the tracking algorithm relies on the characteristic difference in pixel intensity, the tumor region must have a decent contrast such that it can be distinguished from the surrounding tissues by the tracking algorithm. If tumor contrast is too low, the patient has to be excluded because, even if tumor position is determined with the aid of extra information, e.g. gross tumor volume (GTV) contours in the treatment planning system (TPS), such knowledge of tumor position still cannot be utilized by the tracking GUI. However, it is worth noting that quality of tracking results may not agree with image quality estimated visually. Therefore, it is warranted to perform tracking on cine MR images of an extensive number of patients, even those patients whose cine MR images do not appear to be promising visually, as long as the tumor can be seen in the cine MR images, and a patient should not be excluded until tracking results justify the exclusion.

2.3.3 Tracking and Correction

Cine MR images were first loaded into the tracking GUI and the first frame would be displayed in the interface. Next, a rectangular ROI would be drawn manually to include the tumor in the image. The ROI would be used as the template later during the tracking process. To obtain the optimal result, the ROI was drawn as small as possible provided the entire tumor was encompassed. The larger the ROI is, the more pixels that represent the surrounding tissues instead of the tumor will be included,

rendering the tracking result less likely to reflect the motion of tumor accurately. The GUI allows a polygon to be drawn after ROI selection. The polygon serves as a real-time indicator of where the template is shifted to.

Tracking results may not maintain reliable throughout the entire course of tracking due to unfavorable structural characteristics and suboptimal image quality. The most undesirable and obvious effect of this imperfect tracking is discontinuities in displacement curves. For instance, the displacement of a ROI at a certain frame could be as high as more than 20mm to the superior whereas the displacements at adjacent frames were only approximately 8mm. If a discontinuous point was observed in the real-time tracking results and was accompanied by a sudden large change of the polygon's position, the point would be classified as abnormal.

If a displacement curve did not have any abnormal point, it would be accepted as a tumor trajectory from tracking in cine MR images. If a displacement curve did have abnormal points but did not have successive abnormal points of three or more, the curve will be corrected for acceptance. Otherwise, the curve would not be accepted and tracking would be repeated if necessary. The values of abnormal points were set to the interpolated values rounded to the nearest multiples of pixel size, e.g., 1.875mm, 5.625mm, for images with an in-plane resolution of 1.875mm.

To minimize the uncertainty that may be caused by the tracking algorithm and tumor localization, each tumor were tracked 3 to 5 times and average tumor trajectories from tracking in cine MR images were used for following processing.

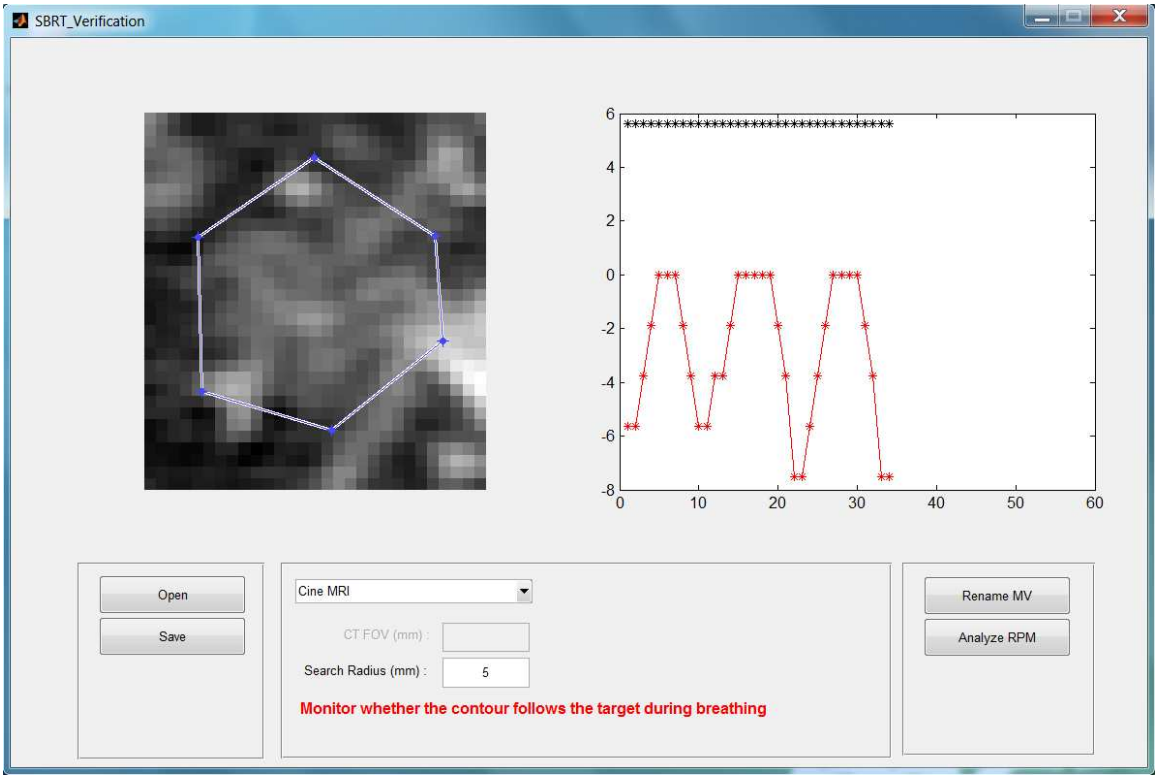

Figure 3: Tracking GUI used for tumor trajectory tracking in cine MR images. The image in the left is the ROI in a coronal frame that is currently being processed with the polygon used for real-time shift indication. The plot in the right shows the SI (red) and the RL (black) trajectories from tracking.

2.3.4 Generation of Reference Tumor Trajectories from Tracking in Cine MR Images (Single-cycle Tumor Trajectories) via Phase Sorting

2.3.4.1 Peak Detection and Correction

A peak detection program capable of ignoring temporally small fluctuation was used to identify the peaks and valleys of tumor trajectories from tracking. The program

rarely incorrectly identified peaks and valleys of average tumor trajectories from tracking in the SI direction but occasionally picked wrong points in the RL and the AP directions.

Since tumor motion in all the three directions were correlated and the trajectories in the SI direction were the most reliable, peaks and valleys of the RL and the AP trajectories were checked against the peaks and valleys in the SI trajectories. Only those in the RL and the AP trajectories that had corresponding peaks or valleys in the SI direction were retained whereas the others were treated as intra-cycle fluctuation and deleted.

2.3.4.2 Phase Sorting

Cine MR images are typically acquired for multiple respiratory cycles and, thus, tumor trajectories from tracking last for multiple cycles as well. To enable comparison, multi-cycle tumor trajectories from tracking have to be converted into trajectories that last for only one respiratory cycle. This conversion was implemented by phase sorting, the same method used for 4D MRI reconstruction. The core step in phase sorting is binning. Each interval between two adjacent valleys were divided into 10 sub-intervals, or bins. Each frame was subsequently assigned to one of the ten bins according to its frame number. The magnitude of displacement for a bin was calculated by averaging the magnitudes of displacement in all the frames in the bin. The magnitude of displacement at frame 10, representing the EOI phase, was set to zero. By doing so, 10-phase tumor

motion trajectories lasting for a single respiratory cycle were generated and used as reference tumor trajectories.

2.4 Generation of DVF-based Tumor Trajectories

2.4.1 Software

An in-house developed GUI was used for 4D MRI image reconstruction. Velocity AI was used to generate DVFs from 4D CT and 4D MRI. Programs for post-processing including DVF display and generation of DVF-based tumor trajectories were developed.

2.4.2 Image Preparation

4D CT images were directly imported into Velocity AI for DVF generation because 4D CT data had already been sorted into 10 phases when they were obtained from the CT scanner.

4D MRI used in this study was acquired in cine mode. Reconstruction was needed to convert raw data into 4D MRI images. The respiratory peaks were first detected by using body area as surrogate. The body area was determined according to the following steps. First, draw an ROI that covers a segment of body surface in the first frame of raw 4D MR data and leave enough room in the upper and lower part of the ROI such that the surface will not go beyond the upper or lower boundaries. Second, set the values of pixels with a value greater than a certain threshold to 1 and the values of rest of the pixels to 0. The number of pixels with a value of 1 in the ROI is defined as the body area. Ideally, if the threshold is properly chosen, all the pixels inside the body

surface will be 1 and those outside will be 0. This method, however, cannot give error-free results. The peaks of body area curves may not be consistent with the peaks of respiratory curves. Therefore, a manual check was followed to ensure the peaks were accurate.

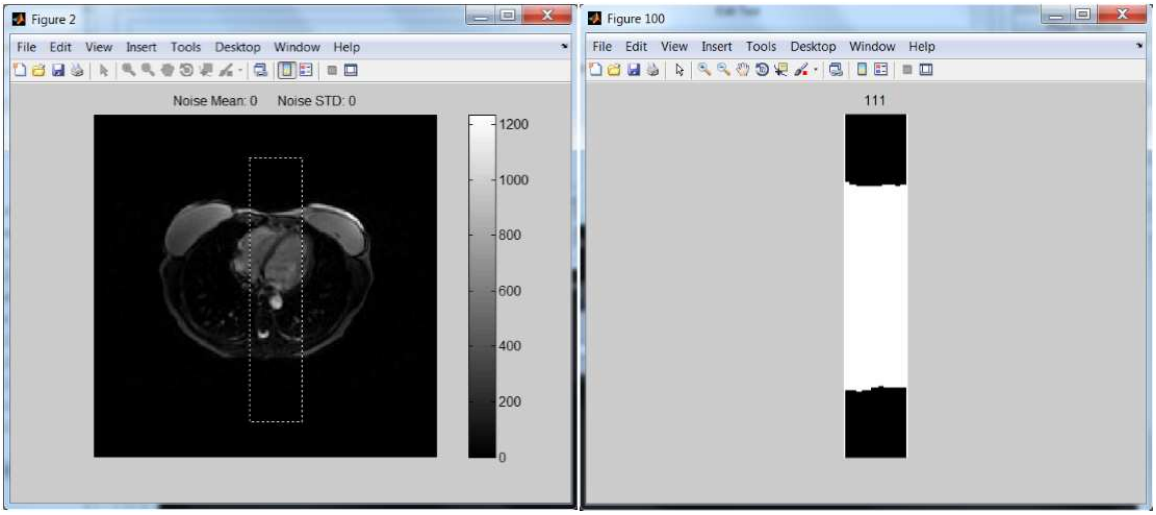

Figure 4: The rectangular region in the left figure is selected for body area calculation. The rectangle in the right figure is the region selected in the left figure. Voxels inside the body are set to white and voxels outside the body to black. Number of white voxels in the right figure is defined as the body area.

The raw data of 4D MRI images and the corrected peaks were input into the GUI and 4D MRI images were reconstructed using phase sorting. The first two frames were discarded due to improper contrast.

2.4.3 ROI selection

All the tumors in 4D CT and 4D MR images were localized in the EOI phase using GTV contours in the TPS as guidance. The first step was finding the "zero" of the images at the EOI phase. The sagittal, coronal, and axial planes that cut through the

image “zero” of the CT and MR images in the TPS were snipped and used to help determine the “zero” of 4D images by comparing 4D images slice-by-slice with the orthogonal planes. The distances from the “zero” to the most right, left, anterior, posterior, inferior, and superior slices that the GTV contour extended to were also acquired from the TPS. With prior knowledge of voxel sizes, the location of the most right, left, anterior, posterior, inferior, and superior slices that the GTV contour extended to in 4D images could be determined. In this study, the tumor region in 4D CT and 4D MR images was represented by the rectangular region encompassed by the six slices.

A potential error in tumor localization lies in that the planning images in the TPS were breath-hold images whereas the 4D CT and 4D MR images in which the tumor was localized are EOI images. Since patients tend to be asked to take a deep breath and hold it during the breath-hold imaging, the diaphragm can be lower than in EOI phase. To correct this discrepancy, the distances from the center of the rectangular volume encompassed by the six slices to the diaphragm, the right side of the body, the posterior side of the body in the TPS were recorded and the rectangular tumor region in the EOI images was shifted to match the distances with those measured in the TPS, the rationale being that, since the tumor moves during respiration, it is more reasonable to define the tumor location with respect to the moving tissues such as diaphragm.

2.4.4 Generation of DVFs

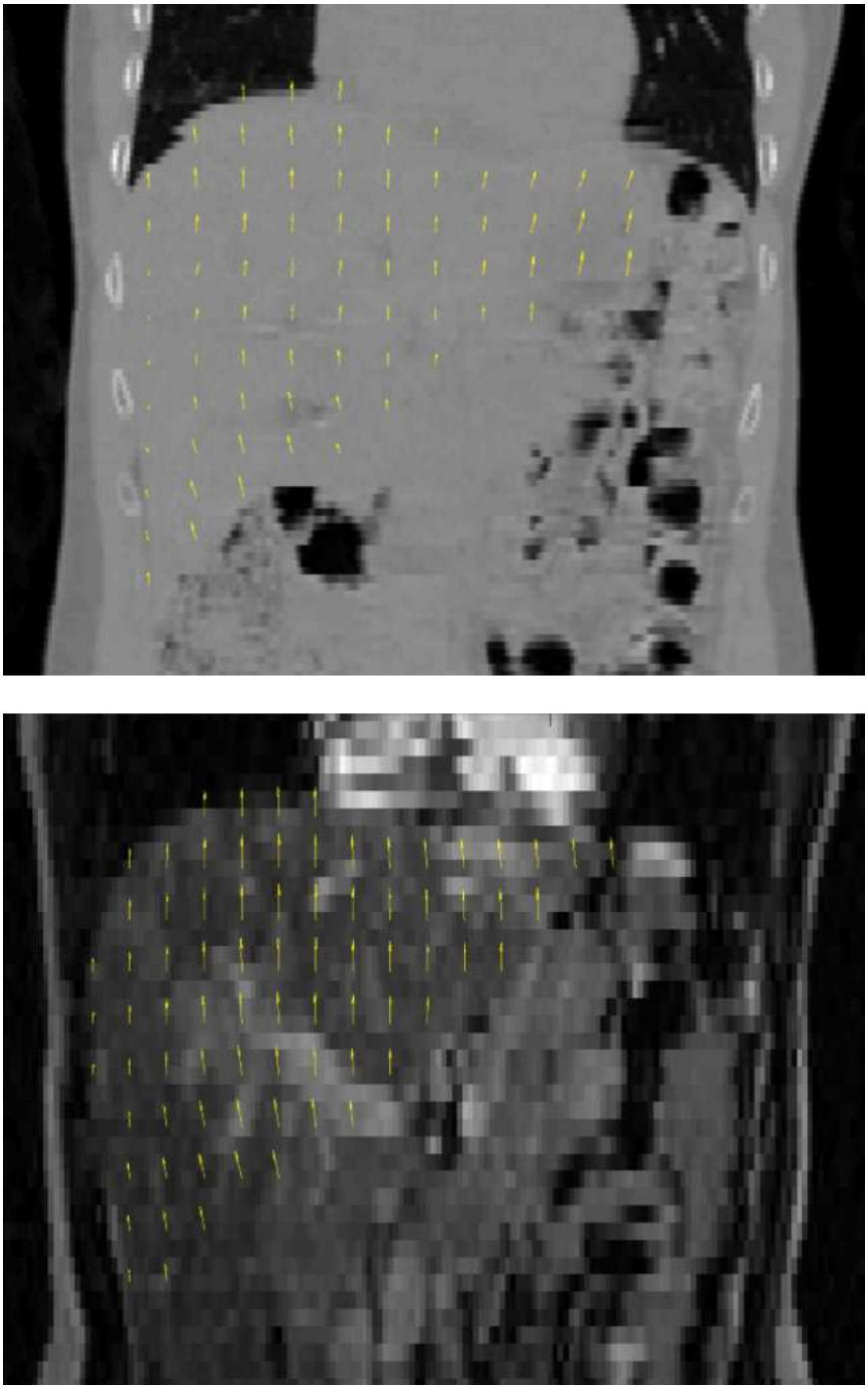

Figure 5: DVFs from 4D CT (top) and 4D MRI (bottom) between the EOI and the EOE phase in the coronal view in the liver.

DVFs used in this study were generated via Velocity AI. A set of 4D images consisting of 10 frames were imported into the Velocity database frame by frame. Each frame was renamed according to its frame number. For both 4D CT and 4D MR images, the end-of-inhalation (EOI) phase was defined as the 95%-5% phase (frame 1). In reality, however, the frame in which the diaphragm reaches its lowest position may not be exactly the frame 1 but could possibly range from frame 10 to frame 2. This discrepancy may come from possible phase shift between the tumor motion and external surrogate used in 4D CT and 4D MRI reconstruction. To facilitate comparison, the frame with the most inferior diaphragm was defined as the EOI phase, or reference frame, and used as primary volume for registration. All the other frames were defined as secondary volumes and each of them was registered with the reference frame.

To eliminate ambiguity, displacement vectors generated from Velocity AI point from each voxel in the primary volume to the secondary volume. In other word, if the voxels in the primary volume move according to the DVF, the deformed volume will have maximum similarity with the secondary volume.

For a set of 4D CT or 4D MRI images with 10 frames, 9 DVFs were generated between the reference frame and each of the remaining 9 frames. Deformable Multipass algorithm was used for 4D CT deformation and MR Corrected Deformable for 4D MRI deformation.

2.4.5 Generation of DVF-based Tumor Trajectories

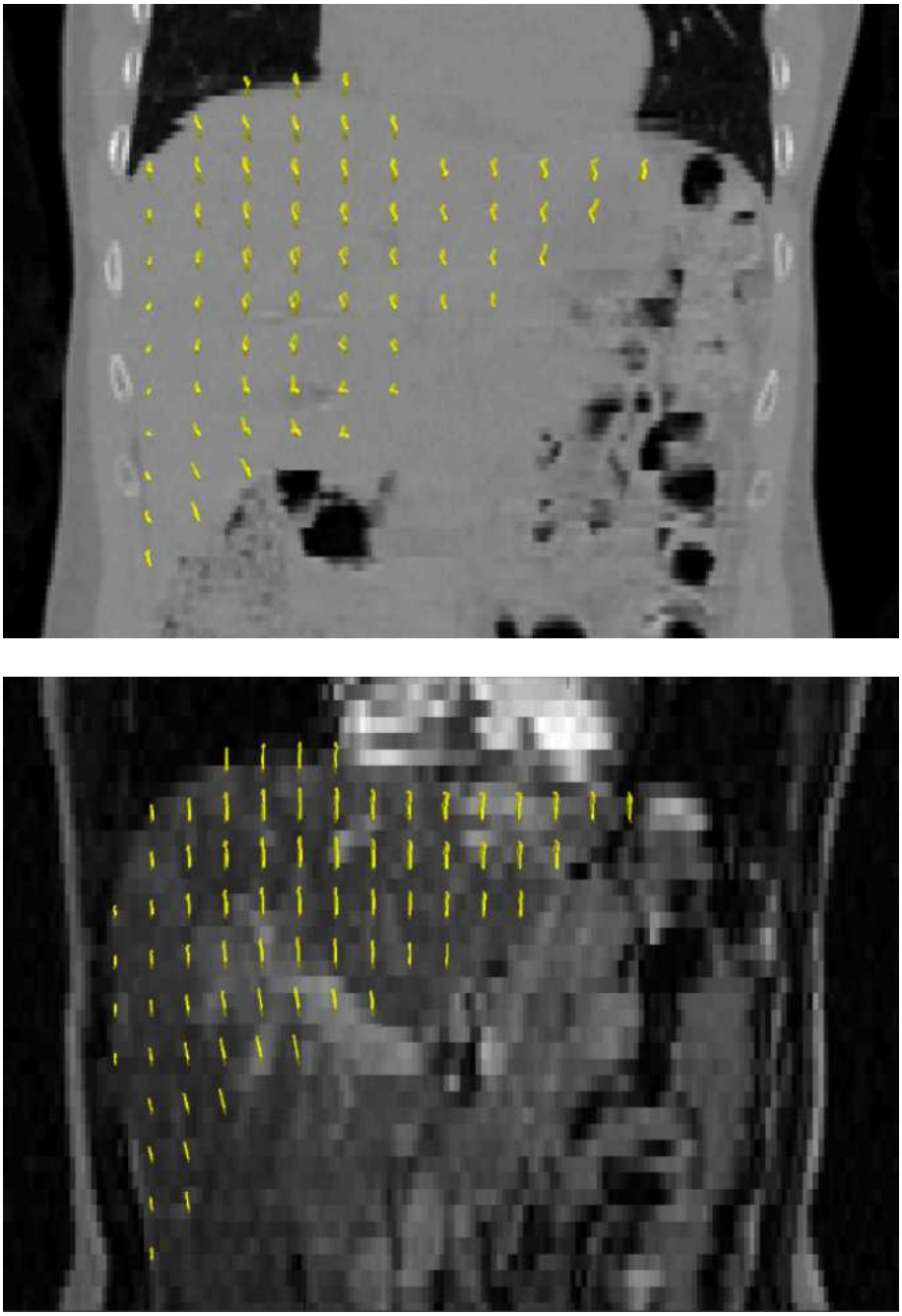

Figure 6: An illustration of 4D CT DVF- (top) and 4D MRI DVF- (bottom) based tumor trajectories. The color fades as the phase proceeds from the EOI to the EOE phase. The ascending part and the descending part overlap because of the small magnitude of displacement in the RL direction.

By connecting the endpoints of all the displacement vectors in temporal order, a motion trajectory were created for each voxel. All the motion trajectories in the ROI were subsequently averaged. To make the average trajectories comparable with tumor trajectories obtained from tracking in cine MR images, an addition data point with a value of zero was added to the end of each average trajectory to represent the position of the tumor in the reference frame. The 10-phase average trajectories were defined as DVF-based tumor trajectories.

2.5 Evaluation of DVFs from 4D CT and 4D MRI

This study aims at evaluating the accuracy of DVFs from 4D CT and 4D MRI using cine MR images as reference. Theoretically, under the assumption of regular breathing, the respiratory motion reconstructed from 4D images via deformation should agree with that from tracking in cine MR images. Therefore, the discrepancy between 4D DVF-based trajectories and cine MR trajectories can indicate DVFs from which modality are more accurate. Metrics used in this study include single-phase difference (ph. diff.), trajectory difference (traj. diff.), and correlation coefficient (c. c.), which are defined as follows.

Suppose r_1, r_2 ... r_{10} are the displacements of the reference trajectory in a certain direction for a patient, and a_1, a_2 ... a_{10} are the displacements of a 4D CT DVF-based or 4D MRI DVF-based trajectory for the same patient.

The ph. diff. at phase i is defined as

$$ph.diff. = a_i - r_i$$

The traj. diff. is defined for each DVF-based trajectories as

$$traj.diff. = \frac{\sum_{i=1}^{n} |a_i - r_i|}{n}$$

where n is the number of bins, which is 10 in this study.

The c.c. is defined as

$$c.c. = \frac{\sum_{i=1}^{n}(a_i - \bar{a})(r_i - \bar{r})}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2 \sum_{i=1}^{n}(r_i - \bar{r})^2}}$$

where n is the number of bins, which is 10 in this study, and $\bar{a}$ is

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$$

and $\bar{r}$ is

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i$$

Three-dimensional (3D) and 4D analysis were accomplished using the metrics mentioned above and the accuracy of the DVFs from 4D CT and 4D MRI was assessed from different perspectives.

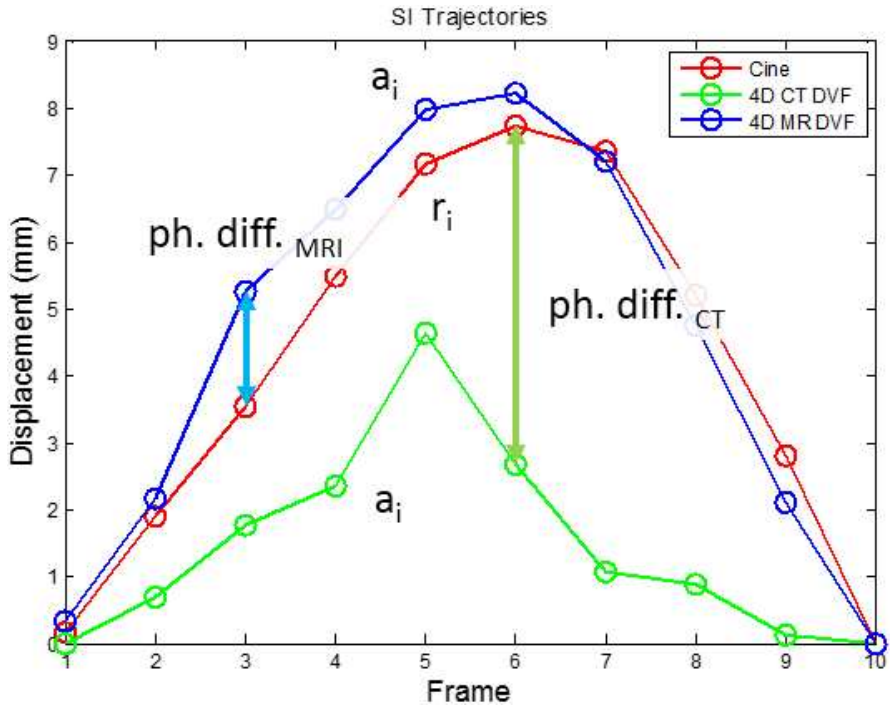

Figure 7: Relevant quantities illustrated in a figure showing a 4D CT DVF-based, a 4D MRI DVF-based and a reference trajectory in the SI direction.

2.5.1 3D Analysis

In 3D analysis, DVFs at each phase were treated as DVFs obtained from independent 3D images. Since the magnitude of displacement between the reference phase and each non-reference phase varies, it is possible to examine how the discrepancy between ROI motion estimated from 4D imaging and that estimated from cine MR is related to the magnitude of displacement. This comparison aimed at evaluation of potential susceptibility of the accuracy of the DVF to actual displacement magnitude. Considering that displacements in the SI direction were the greatest among the three directions, the single-phase difference was plotted against the displacement in the reference trajectories in the SI direction, for both 4D CT DVF-based and 4D MRI

DVF-based trajectories. A two tailed Wilcoxon signed-rank test was conducted to determine if the expected medians of single-phase difference of 4D CT and 4D MRI were different.

2.5.2 4D Analysis

In 4D analysis, accuracy of DVFs was evaluated based on the agreement of the DVF-based trajectory and the corresponding reference trajectory, measured by the trajectory difference and correlation coefficient. Two tailed Wilcoxon signed-rank tests were conducted to determine if the expected medians of trajectory difference and correlation coefficient of 4D CT and 4D MRI were different.

3. Results

3.1 Patient Selection

A total of 12 patients were screened for eligibility, seven of which had complete image data of 4D CT, 4D MRI and cine MR images, and visible tumors in cine MR images. According to the quality of their tumor trajectories from tracking in cine MR images, five were rated as eligible. Reasons for the exclusion of the other two are as follows.

The first excluded patient had a tumor too close to the heart. Theoretically, the motion of a liver tumor is a superposition of respiratory motion and heart motion. For most tumors, the effect from respiratory motion is dominant because the magnitude of heart motion is smaller and the already small motion is further mitigated by liver tissues between the heart and the tumor. The above statement, however, does not hold for the tumor in this patient due to its proximity to the heart. As can be seen from Figure 5, the magnitude of high frequency tumor motion caused by heart motion is comparable with the magnitude of low frequency tumor motion caused by respiratory motion, making identification of respiratory peaks and valleys of the displacement curve extremely difficult, if not impossible. Even if respiratory peaks and valleys are located, because the heart motion adds considerable fluctuation to the displacement curve, it will not be qualified to serve as reference. A satisfactory tumor displacement curve is shown in Figure 4 as reference.

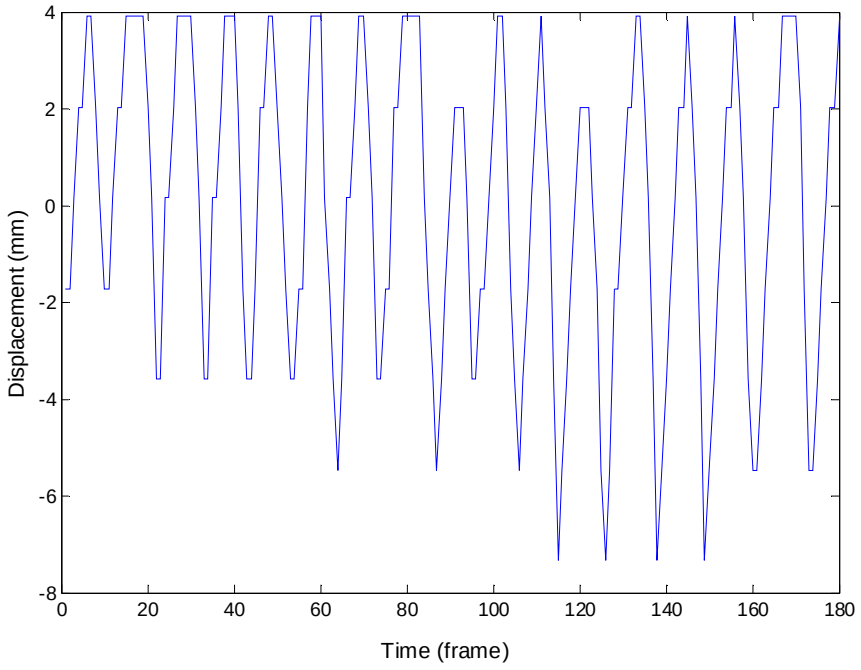

Figure 8: A tumor displacement curve in the SI direction tracking in the cine MR images of a selected patient.

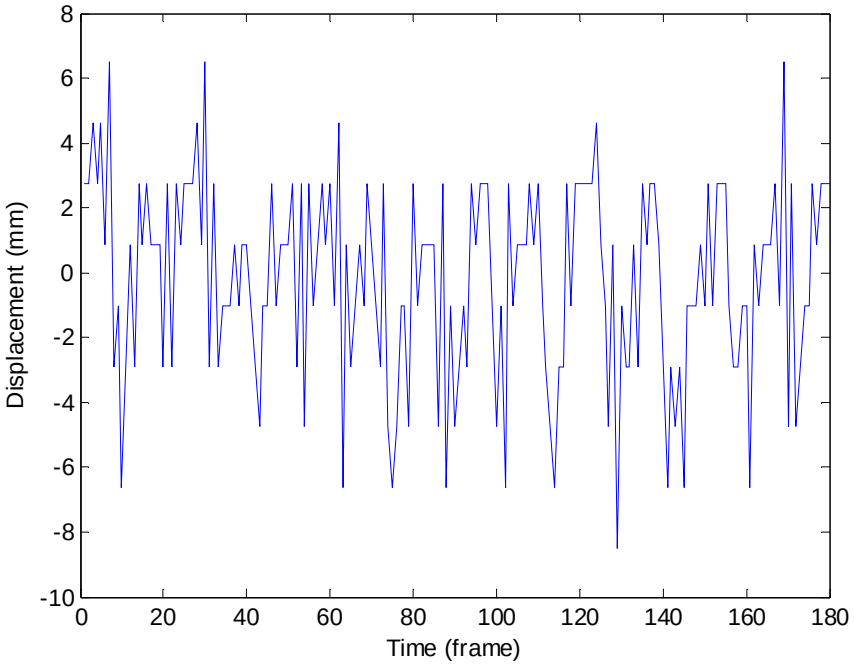

Figure 9: A tumor displacement curve in the SI direction from tracking in the cine MR images of the excluded patient with a tumor affected by heart motion.

The second patient was excluded because, in the cine MR images, his tumor was not moving with respiration, as shown in Figure 3. Since this study aims at evaluating accuracy of DVFs from 4D images by comparing the motion information extracted from 4D images with that information from cine MR images, a stationary tumor that cannot provide such information is not eligible for use.

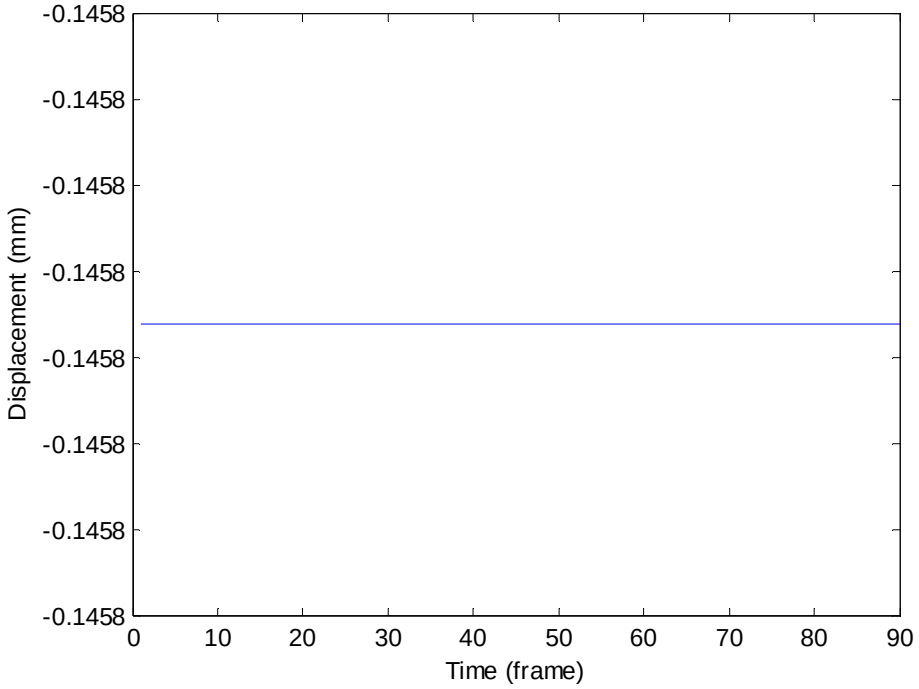

Figure 10: A tumor displacement curve in the SI direction from tracking in the cine MR images of the excluded patient with a stationary tumor.

Table 3: A summary of patient characteristics. The primary site is shown in the parentheses if the lesion in the liver is secondary.

Patient	Age (Years)	Gender	Weight (kg)	Primary or Secondary
1	45	F	61	Secondary (Breast)
2	67	F	74	Primary
3	53	M	109	Secondary (Liver)
4	74	M	102	Secondary (Liver)
5	67	M	113	Primary

3.2 Verification of Agreement between ROI in 4D CT and 4D MR Images

The validity of comparing 4D CT DVF-based and 4D MRI DVF-based tumor trajectories with reference tumor trajectories, respectively, relies on the assumption that the ROI in 4D CT images and 4D MR images contain the same tissues. The volume of ROI in 4D CT and 4D MRI has been guaranteed to be the same by the localization method used in this study. It was the position of the center of ROI that required further verification. D_{right} is the distance between the tumor and the right side of body, D_{back} is the distance between the tumor and the back side of body, and D_{dia} is the distance between the tumor and the diaphragm. These three distances were measured in 4D CT and 4D MRI and corrected to the distances measured in planning images in the TPS. The total percent difference (tot. % diff.) was calculated using the following equation,

$$tot.\,\%\,diff. = \frac{|D_{right,4D} - D_{right,TPS}|}{l_{RL}} + \frac{|D_{back,4D} - D_{back,TPS}|}{l_{AP}} + \frac{|D_{dia,4D} - D_{dia,TPS}|}{l_{SI}}$$

When the D_{4D} is approximately equal to D_{TPS}, it can be easily demonstrated that

$$\frac{\Delta V}{V} \approx tot.\,\%\,diff.$$

where ΔV is the volume ROI in 4D images that is not overlapping with the ROI in planning images.

Table 4: Edge lengths of rectangular ROIs determined in the planning images in the TPS. **Unit: mm**

Patient	l_{RL}	l_{AP}	l_{SI}
1	47.0	45.7	35.0
2	33.0	29.2	30.0
3	26.4	39.5	27.0
4	27.9	24.1	22.5
5	76.1	94.7	92.6

Table 5: Distance between the center of tumor to the right side of body, the posterior side of body, and the diaphragm, measured in the planning images in the TPS and in the 4D images at the EOI phase.

			Distance to Center of Tumor Region (mm)			
Patient	Modality		D_{right}	D_{back}	D_{dia}	Tot. % Diff.
1	4D CT	4D	65.6	108.1	67.5	3.00
		TPS	65.7	108.2	66.6	
	4D MR	4D	64.0	100.8	60.0	3.73
		TPS	64.9	101.5	59.9	
2	4D CT	4D	81.5	72.7	40.0	2.38
		TPS	81.6	73.2	39.9	
	4D MR	4D	97.0	63.0	35.0	5.93
		TPS	97.4	62.4	35.8	
3	4D CT	4D	68.5	161.3	16.5	1.69
		TPS	68.2	161.3	16.4	
	4D MR	4D	62.6	160.8	12.5	6.99
		TPS	63.0	160.1	13.5	
4	4D CT	4D	105.3	179.4	59.8	4.01
		TPS	105.7	178.9	59.9	
	4D MR	4D	105.7	171.5	57.3	7.27
		TPS	105.9	171.1	58.4	
5	4D CT	4D	90.9	140.7	158.0	1.87
		TPS	91.5	140.7	159.0	
	4D MR	4D	80.5	145.5	168.0	2.62
		TPS	80.7	144.8	166.5	

It can be observed that the ROI in 4D MR images has uniformly worse degree of agreement than ROI in 4D CT. This can be explained by the larger voxel sizes of 4D MR images because the largest possible value of $|D_{4D}-D_{TPS}|$ equals half of the voxel length in the same direction. Another observation is that ROI with larger edge lengths is less susceptible to imperfect alignment. Tot. % Diff. of ROIs in 4D CT and 4D MRI images for patient 4, whose tumor is the smallest in edge lengths, are 4.01% and 7.27%, both the largest among the five patients whereas Tot. % Diff. of ROIs in 4D CT and 4D MRI images for patient 5, who has the largest tumor, are 1.87% and 2.62%, considerably lower than average.

3.3 Comparison of DVF Distribution in Tumor Region for 4D CT and 4D MRI

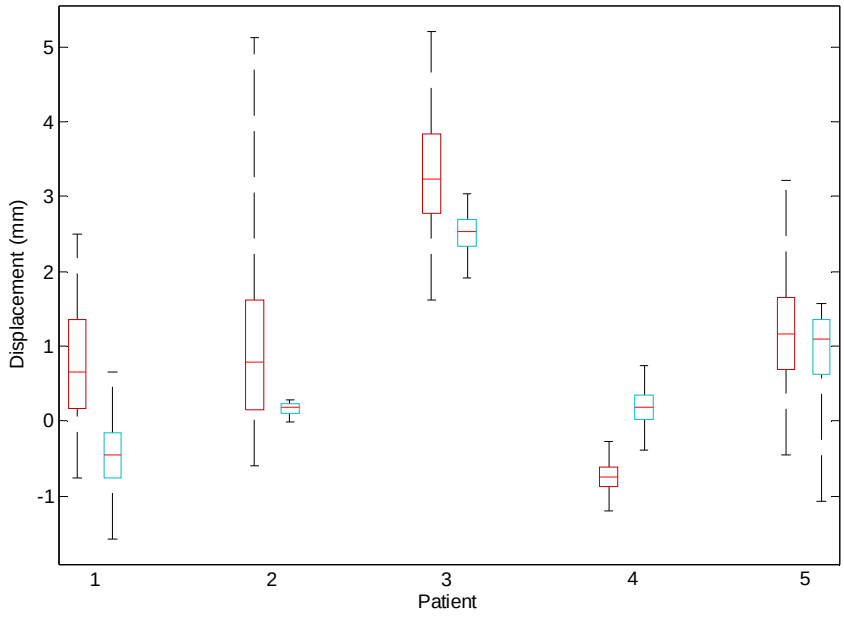

Figure 11: RL displacement vectors between the EOI and the EOE phases in the tumor region for all patients. Red boxes are for CT and blue for MRI. The red bars stand for the medians. The central 50% of the data points are included in the box.

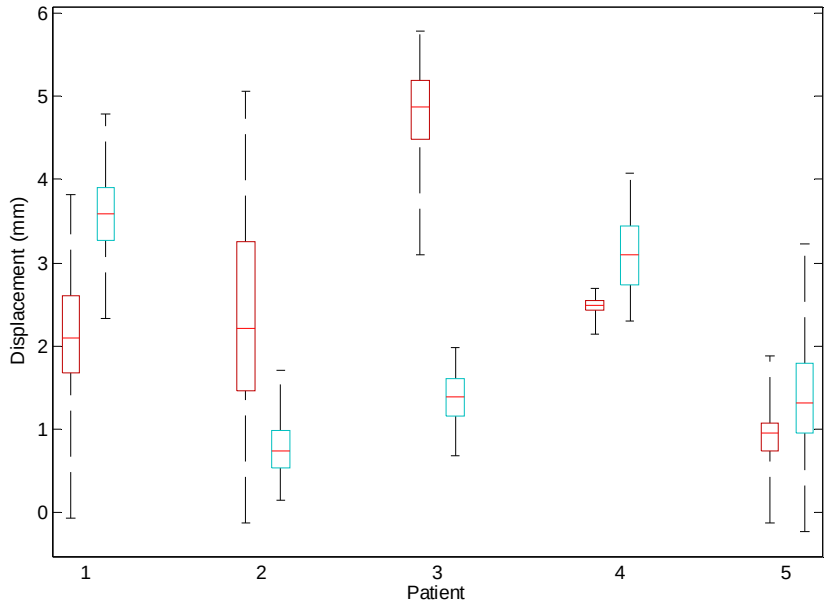

Figure 12: AP displacement vectors between the EOI and the EOE phases in the tumor region for all patients. Red boxes are for CT and blue for MRI. The red bars stand for the medians. The central 50% of the data points are included in the box.

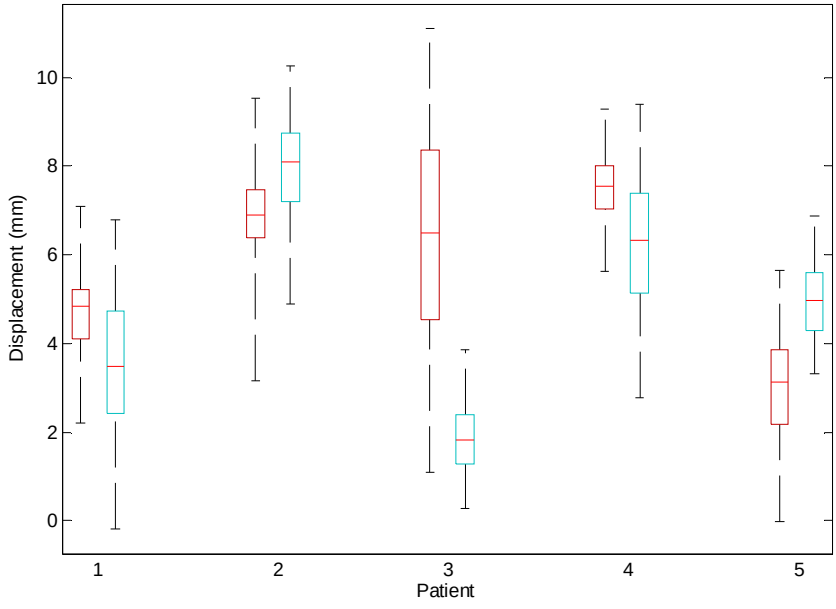

Figure 13: SI displacement vectors between the EOI and the EOE phases in the tumor region for all patients. Red boxes are for CT and blue for MRI. The red bars stand for the medians. The central 50% of the data points are included in the box.

The range of the magnitude of displacement vector varies considerably from patient to patient, noticeably in the RL direction, where the ranges from two patients can have no overlapping. The agreement of the ranges of the magnitude of displacement vector for 4D CT and 4D MRI is better in the SI direction than in the other two.

3.4 Reference Tumor Trajectories and DVF-based Tumor Trajectories

3.4.1 Reference Tumor Trajectories

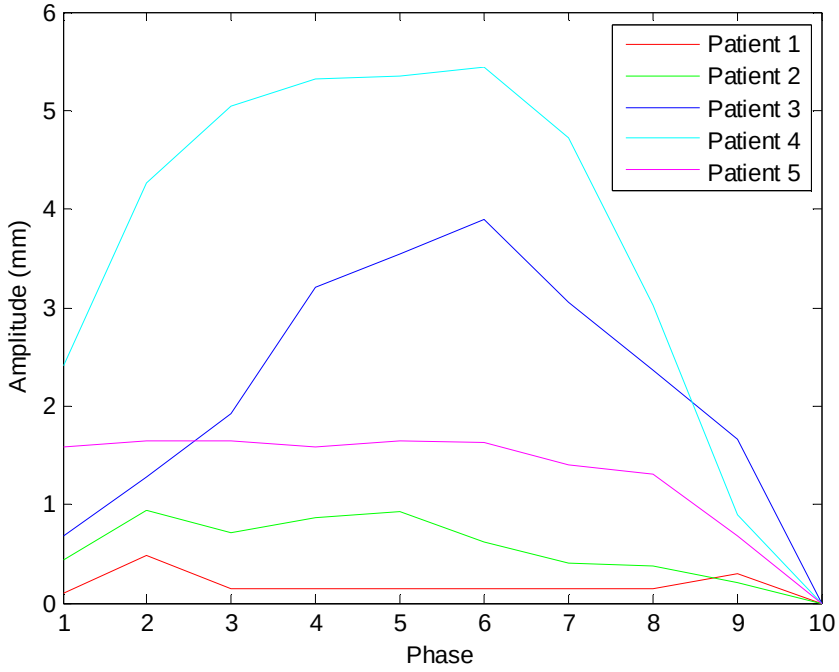

Figure 14: Reference tumor trajectories in the RL direction for all patients.

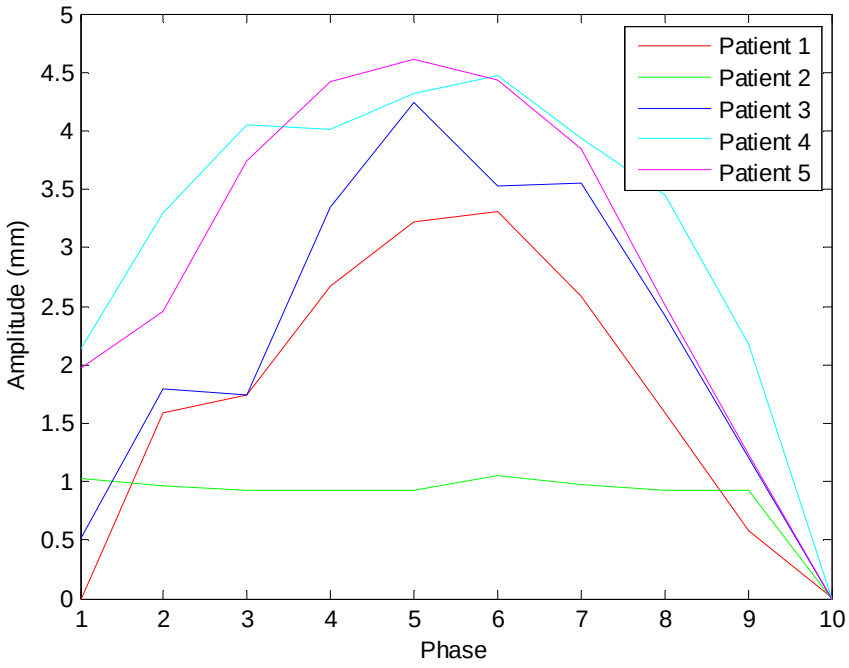

Figure 15: Reference tumor trajectories in the AP direction for all patients.

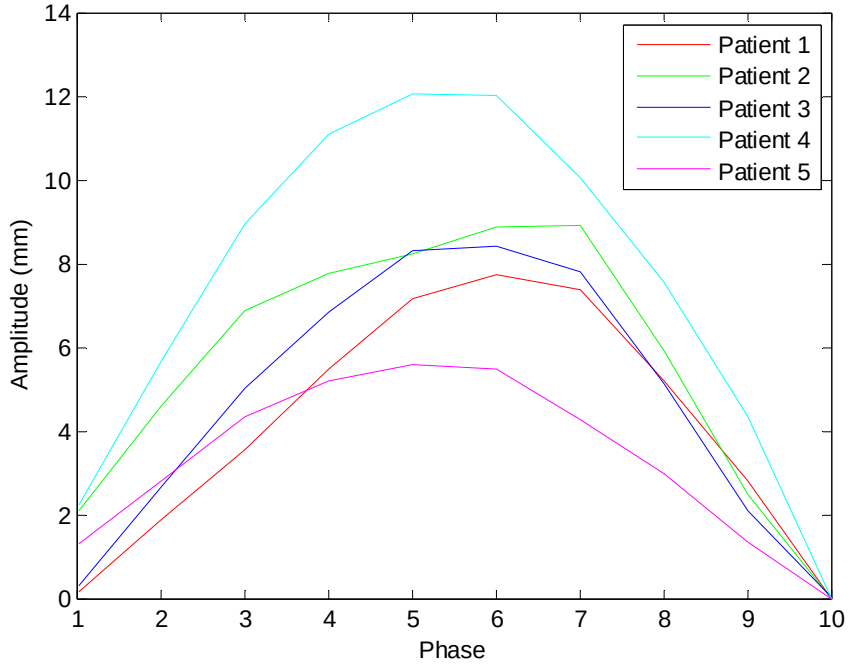

Figure 16: Reference tumor trajectories in the SI direction for all patients.

3.4.2 DVF-based Tumor Trajectories

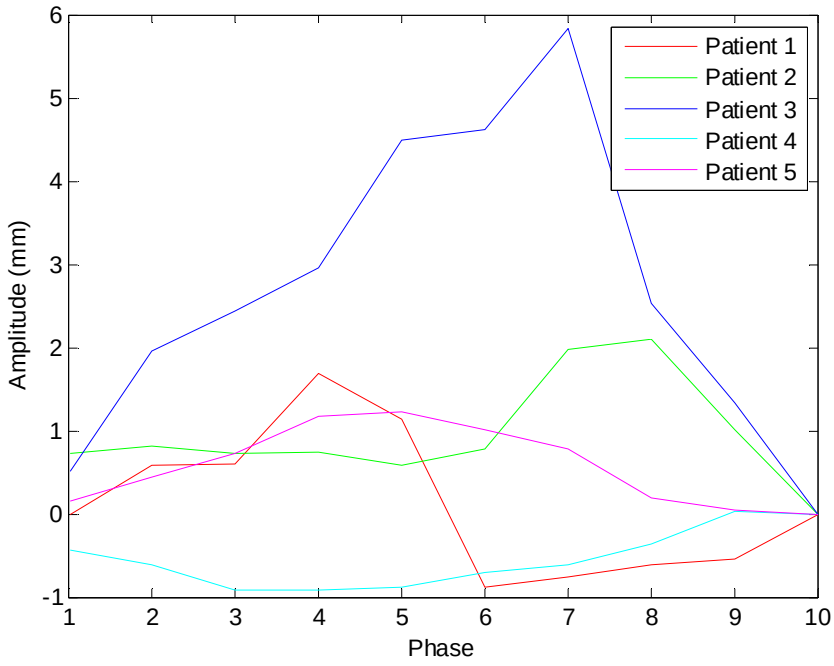

Figure 17: 4D CT DVF-based tumor trajectories in the RL direction for all patients.

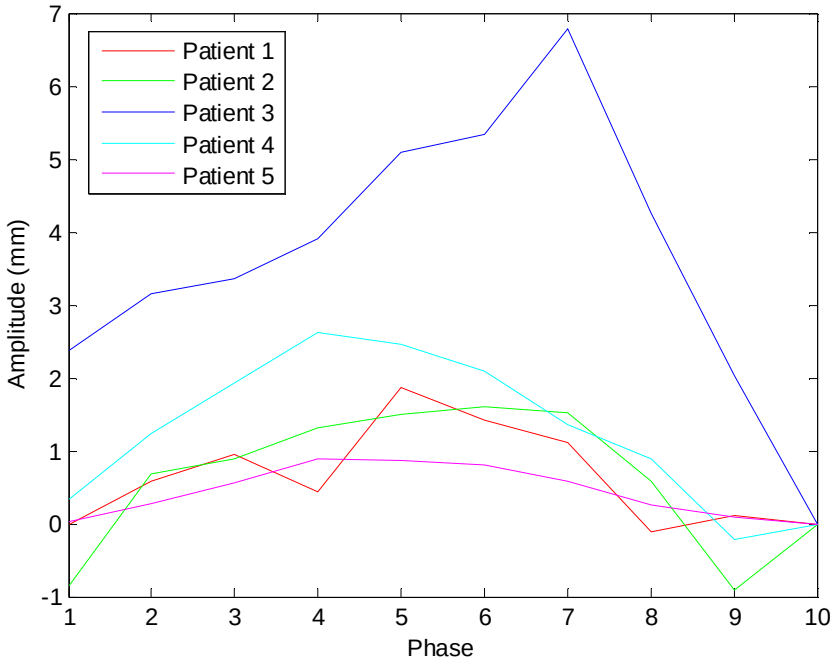

Figure 18: 4D CT DVF-based tumor trajectories in the AP direction for all patients.

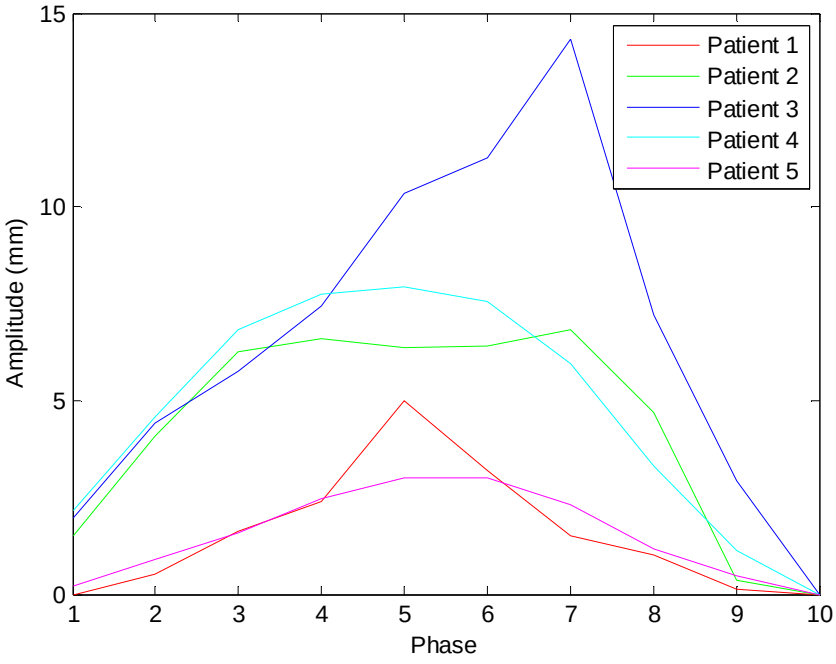

Figure 19: 4D CT DVF-based tumor trajectories in the SI direction for all patients.

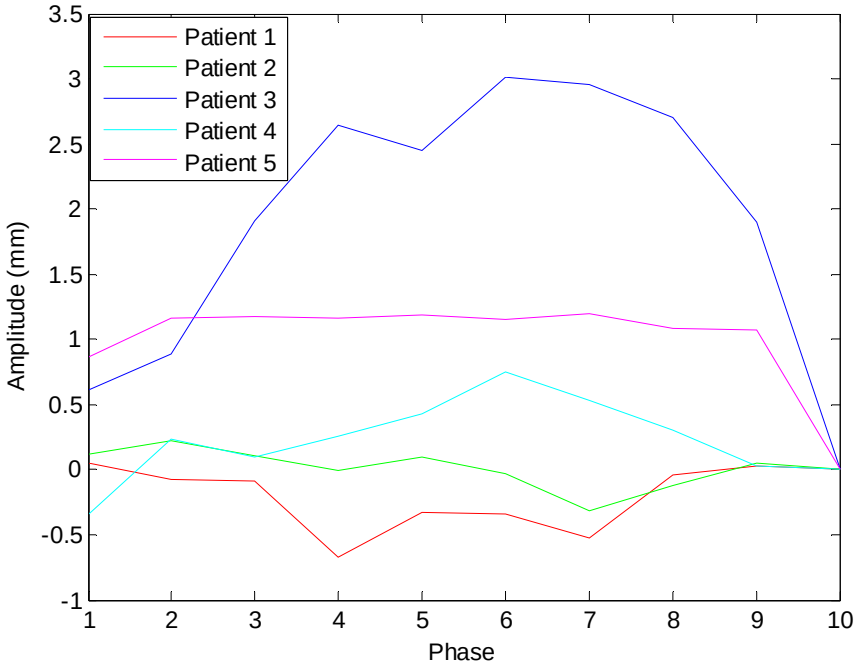

Figure 20: 4D MRI DVF-based tumor trajectories in the RL direction for all patients.

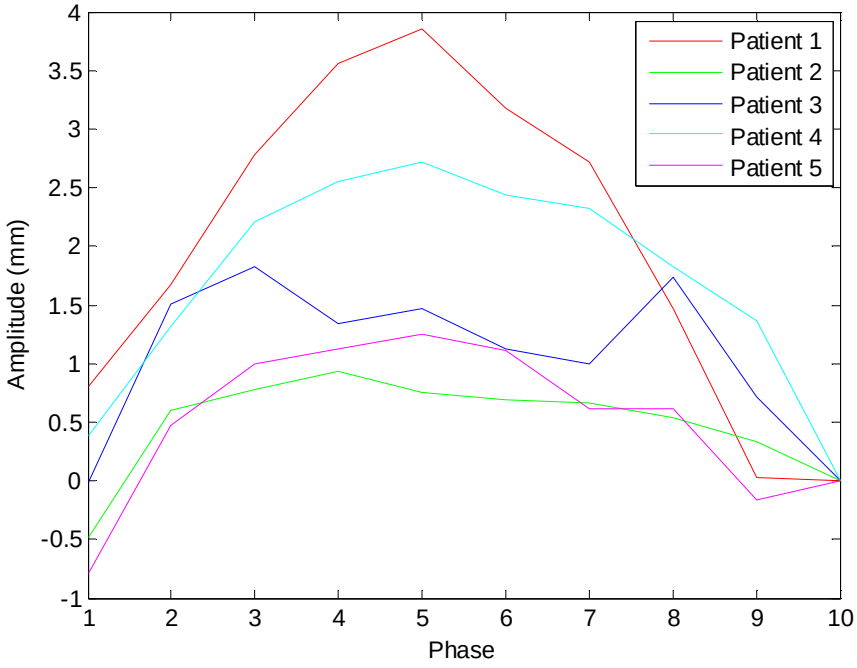

Figure 21: 4D MRI DVF-based tumor trajectories in the AP direction for all patients.

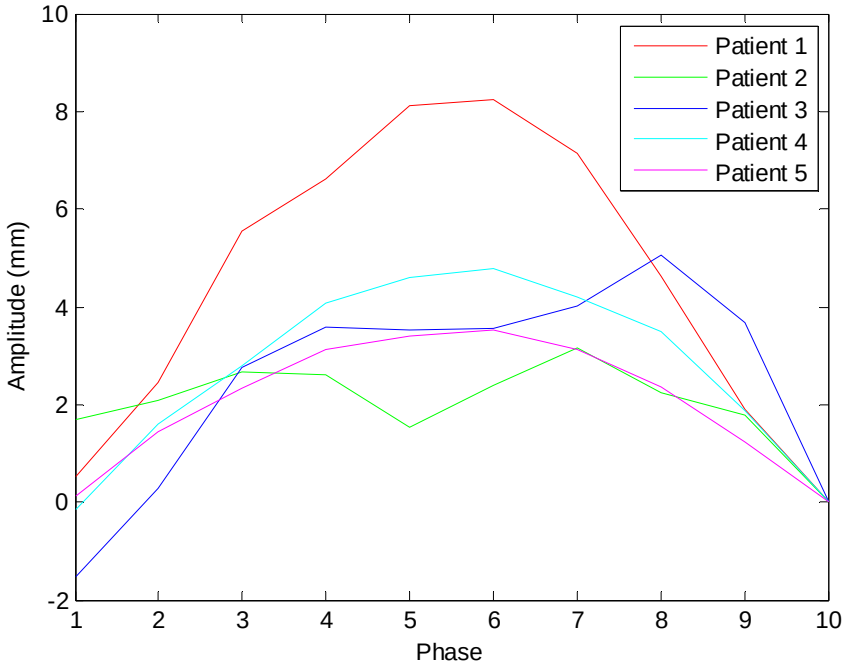

Figure 22: 4D MRI DVF-based tumor trajectories in the SI direction for all patients.

3.4 Evaluation of DVFs from 4D CT and 4D MRI

3.4.1 3D Analysis

3.4.1.1 Single-phase Difference vs. Magnitude of Displacement

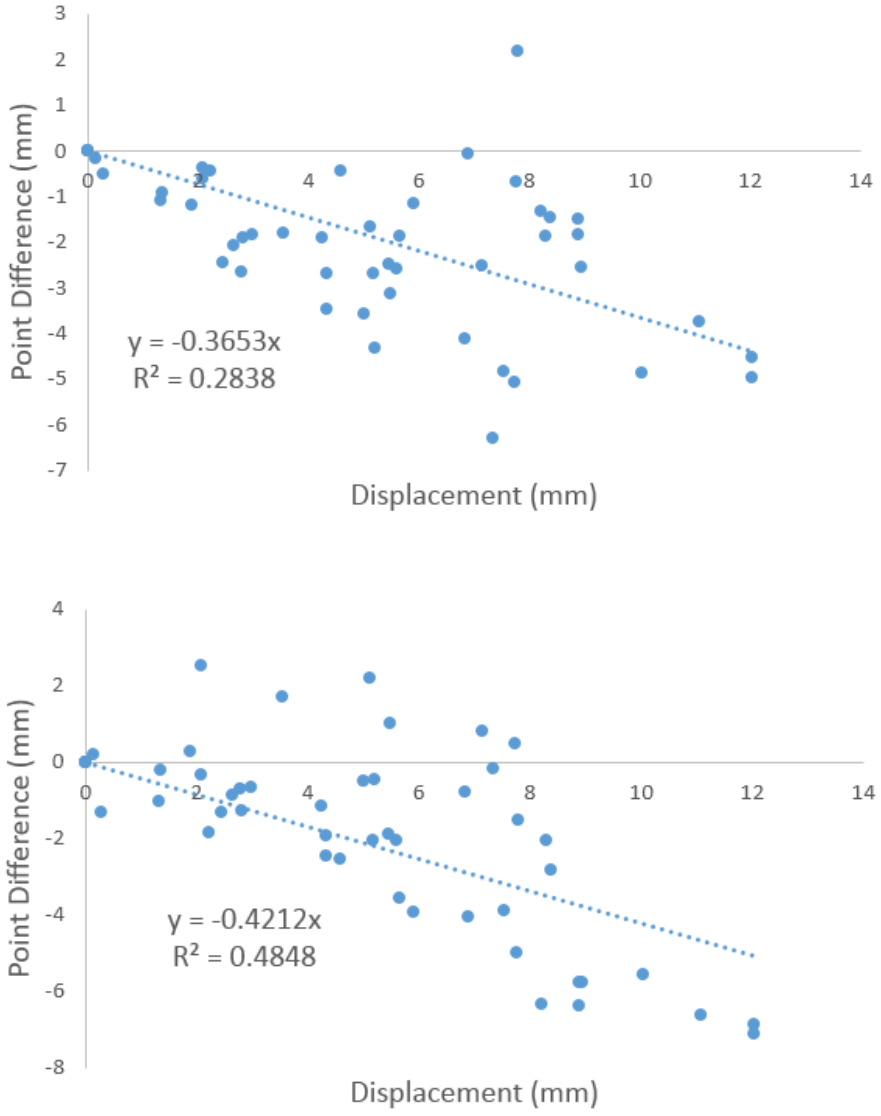

Figure 23: The single-phase difference for 4D CT (top) and 4D MRI (bottom) is plotted against the tumor displacement in the SI direction. Data from all the five patients are reflected in this figure.

According to the results of linear regression, the single-phase difference for 4D MRI shows a greater correlation with the magnitude of displacement than the single-phase difference for 4D CT because the slope of the regressed line for 4D MRI is -0.42, larger than the slope for 4D CT (-0.37) in terms of the absolute value. That being said, the coefficients of determination for both regressed line are not large, meaning that a single-phase difference can fall at somewhere quite distant from the value predicted by the regressed line. The linear dependence of the single-phase difference on the displacement is not conclusive for neither 4D CT nor 4D MRI. It is interesting to see that, although the regression result indicates a weaker dependence on displacement and a more scattered distribution with respect to the regressed line, the single-phase difference for 4D CT are uniformly below zero except for one point. Considering the regulation term in a DIR algorithm, which is intended to eliminate drastic change in the DVF, has a smoothing effect on the DVF, it is expected that the deformation vectors with large magnitudes are more susceptible to this effect and therefore tend to be smaller than actual displacements.

3.4.2.2 Statistical Study on the Absolute Single-phase Difference

A Wilcoxon signed-rank test was accomplished on the absolute single-phase difference in millimeter and in number of voxels for all the three directions to test if the median of X-Y equals zero, where X stands for the absolute single-phase difference for 4D CT combining the ten phases in all the five patients together, and Y stands for the

absolute single-phase difference for 4D MRI combining the ten phases in all the five patients together. In this test, the difference between the DVF-based tumor trajectories and the corresponding reference trajectory at each phase is treated equally, rather than averaged for each patient. The validity relies on the assumption that the single-phase difference does not depend on any patient-specific characteristics.

Table 6: P-values of a Wilcoxon signed-rank test on the absolute ph. diff. in mm, and in number of voxels for the RL, the AP, and the SI directions.

	RL	AP	SI
absolute ph. diff. (mm)	0.0093	0.0011	0.8170
absolute ph. diff. (voxels)	<5E-5	<2E-6	<2E-4

All p-values in Table. 6 reach the significant level of 0.05 except for the p-value for the test on the absolute single-phase difference in millimeters in the SI direction, suggesting that the median of the absolute single-phase difference in millimeters in the RL and the AP directions, and the single-phase difference in voxels in the RL, the AP, and the SI directions are likely to be different for 4D CT and 4D MRI. By comparing the trajectory difference in millimeters, and in number of voxels, as shown in Table. 7, which are also the mean values of the absolute single-phase differences in millimeters and in voxels, the results of the signed-rank tests are in favor of the hypotheses that accuracy of the 4D MRI DVF is potentially superior to the accuracy of the 4D CT DVF except for the SI direction when evaluated in length.

3.4.2 4D Analysis

3.4.2.1 Quantitative Results for All Patients

Table 7: Results of agreement measurement between DVF-based trajectories and trajectories from tracking in cine MR images in the RL, AP and SI directions.

Patient	Metric	Modality	RL	AP	SI
1	traj. diff. (mm)	4D CT	0.54	0.90	2.72
		4D MR	0.48	0.36	0.57
	traj. diff. (voxels)	4D CT	0.43	0.71	1.09
		4D MR	0.26	0.19	0.11
	c. c.	4D CT	0.06	0.90	0.74
		4D MR	0.08	0.94	0.97
2	traj. diff. (mm)	4D CT	0.67	1.00	1.01
		4D MR	0.42	0.43	3.56
	traj. diff. (voxels)	4D CT	0.52	0.79	0.40
		4D MR	0.30	0.31	0.71
	c. c.	4D CT	0.04	0.26	0.97
		4D MR	0.68	0.30	0.86
3	traj. diff. (mm)	4D CT	0.52	1.02	1.78
		4D MR	0.45	1.24	1.45
	traj. diff. (voxels)	4D CT	0.38	0.74	0.89
		4D MR	0.24	0.66	0.29
	c. c.	4D CT	0.92	0.92	0.85
		4D MR	0.92	0.54	0.85
4	traj. diff. (mm)	4D CT	4.01	1.98	3.13
		4D MR	3.55	1.22	4.36
	traj. diff. (voxels)	4D CT	3.16	1.56	1.25
		4D MR	1.89	0.65	0.87
	c. c.	4D CT	-0.92	0.81	0.94
		4D MR	0.48	0.96	0.98
5	traj. diff. (mm)	4D CT	0.76	2.45	1.81
		4D MR	0.54	2.30	1.22
	traj. diff. (voxels)	4D CT	0.60	1.93	0.72
		4D MR	0.29	1.22	0.24
	c. c.	4D CT	0.66	0.96	0.97
		4D MR	0.80	0.84	0.97

Table 8: Averaged traj. diff. in mm, traj. diff. in number of voxles and correlation coefficients for the five patients.

Metric	Modality	RL	AP	SI
traj. diff. (mm)	4D CT	1.30±1.36	1.47±0.63	2.09±0.75
	4D MR	1.09±1.23	1.11±0.70	2.23±1.46
traj. diff. (voxels)	4D CT	1.02±1.07	1.15±0.50	0.87±0.29
	4D MR	0.60±0.65	0.61±0.36	0.45±0.29
c. c.	4D CT	0.15±0.64	0.77±0.26	0.89±0.09
	4D MR	0.59±0.30	0.72±0.26	0.93±0.06

This study aims at evaluating the accuracy of DVFs from 4D CT and 4D MRI by comparing similarity between the DVF-based tumor trajectories and the tumor trajectories from tracking in cine MR images. The similarity is quantified using trajectory difference and correlation coefficient.

Table 8 shows the trajectory differences, and correlation coefficients averaged over the five patients. The trajectory differences in millimeters for both 4D CT and 4D MRI increase as the direction goes from RL, AP to SI. Since typically the respiratory motion has the largest amplitude in the SI direction, less in the AP direction and the smallest in the RL direction, the increase in trajectory difference agrees with the observation that the larger difference between the primary and the secondary images, the larger the deformation error.

As for the correlation coefficient, an increasing trend is observed for both 4D CT and 4D MRI as the direction for which correlation coefficient is calculated varies from RL, AP to SI. Meaning that the trends of the reference trajectories and DVF-based

trajectories match best in the SI direction, less good in the AP direction and worst in the RL direction. It is worth noting that curves with good correlation may still have large difference.

Although the metrics show consistent results as the direction varies, they fail to indicate which modality is better. All metrics did not give consistent indications of which modality can provide more accurate DVFs. In the SI direction, 4D CT trajectories exhibited smaller trajectory difference in millimeters (traj. diff. (mm)= 2.09±0.75mm) but larger trajectory difference in voxels (traj. diff. (voxels)= 0.87±0.29) and smaller correlation coefficient (c. c. = 0.89±0.09) than 4D MR trajectories (traj. diff. (mm)= 2.23±1.46mm, traj. diff. (voxels)= 0.45±0.29, c. c. = 0.93±0.06) whereas 4D MR (traj. diff. (mm)= 1.09±1.23mm, traj. diff. (voxels)= 0.60±0.65, c. c. = 0.59±0.30) surpassed 4D CT (traj. diff. (mm)= 1.30±1.36mm, traj. diff. (voxels)= 1.02±1.07, c. c. = 0.15±0.64) in every metric in the RL direction. In the AP direction, 4D MR trajectories have a smaller trajectory difference in millimeters (traj. diff. (mm) = 1.11±0.70mm) and a smaller trajectory difference in voxels (traj. diff. (voxels) = 0.61±0.36) but a slightly smaller correlation coefficient (c. c. = 0.72±0.26) than 4D CT trajectories (traj. diff. (mm) = 1.47±0.63mm, traj. diff. (voxels) = 1.15±0.50, c. c. = 0.77±0.26). This inconsistency indicates that, by using 4D analysis, for the five patients studied, no evidence shows that DVFs from 4D CT are more accurate than DVFs from 4D MRI or vice versa.

Although no definite conclusion can be made in terms of the accuracy of 4D CT and 4D MRI from the patient-averaged results shown in Table 8, patient-specific results in Table 7 show an interesting results that, for patient 1, accuracy of DVFs from 4D MRI is dominantly better than that of DVFs from 4D CT.

3.4.2.2 Statistical Study on the Trajectory Difference and the Correlation Coefficient

A Wilcoxon signed-rank test was conducted to test if the median of X-Y equals zero, where X stands for the trajectory difference in millimeters for 4D CT for each patient, and Y stands for the trajectory difference in millimeters for 4D MRI for each patient. The same test was also conducted for the trajectory differences in number of voxels for 4D CT and 4D MRI, and for the correlation coefficients for 4D CT and 4D MRI. The p-values are summarized in Table. 9.

Table 9: P-values of a Wilcoxon signed-rank test on traj. diff. in mm, traj. diff. in voxels, and c. c for the RL, the AP, and the SI directions.

	RL	AP	SI
traj. diff. (mm)	0.0625	0.1875	1.0000
traj. diff. (voxels)	0.0625	0.0625	0.1250
c. c.	0.1250	0.1250	1.0000

No p-value is below the significant level of 0.05 in Table. 9, meaning that the hypotheses that the median of X-Y equals zero cannot be rejected. Therefore, the results of the trajectory difference in millimeters, the trajectory difference in number of voxels, and the correlation coefficient do not support that the accuracy of 4D CT DVF is significantly different from the accuracy of 4D MRI DVF.

4. Discussion

4.1 Interpretation

4.1.1 DVF Distribution

It was expected that the variation of the magnitude of displacement vectors inside the ROI was relatively small considering that the DVF was smoothed during registration. The magnitude distribution of the DVF for some patients, however, covers a wide range. The overlapping between the ranges of the magnitude of displacement vectors from 4D CT and 4D MRI varies considerably with the patient. No clear trend in terms of the pattern and the degree of overlapping has been observed. It should also be noted that a good agreement in magnitude distribution cannot be used as evidence of good agreement between two DVFs. Spatial distributions must be compared before two DVFs can be considered in agreement with each other. In sum, the lack of similarity between the DVFs from 4D CT and 4D MRI may necessitate patient-by-patient examination of the accuracy of 4D CT and 4D MRI DVF if the DVF is to be applied for research and clinical use.

4.1.2 Trajectory Differences in Length and in Number of Voxels

As shown in Table. 8, the trajectory difference in millimeters for 4D CT is greater than that for 4D MRI in the RL and the AP directions whereas the trajectory difference in millimeters for 4D CT is smaller in the SI direction. The statistical results do not support there is significant difference between the median of the trajectory difference in

millimeters for 4D CT and 4D MRI. Therefore, the accuracy of the DVFs from 4D CT and 4D MRI is concluded as comparable with current results. However, if the discrepancy between the DVF-based trajectories and the reference trajectories is measured in number of voxels, the DVF from 4D MRI performs better in all the three directions.

4D MR images have better soft-tissue contrast and more features, which should lead to improved registration accuracy. The comparable results shown in this study suggest that the benefit brought by high soft-tissue contrast is nullified by other factors. It is hypothesized that registration errors are also related to resolution and it is the inferior resolution of 4D MRI compared to that of 4D CT that results in the comparable accuracy of DVFs from 4D CT and 4D MRI. This statement is supported by the data showing that 4D MRI has smaller single-phase differences and trajectory differences in number of voxels. This hypothesis is subject to further investigation and possible approaches are proposed in 4.3 Future Work.

From the voxel-based perspective, the DVFs from 4D MRI show more accuracy in terms of trajectory difference. From the perspective of application, however, similar results in trajectory difference of 4D CT and 4D MR indicates the magnitude of possible error that may be introduced to calculation based on the DVF are likely to be the same.

4.2 Limitations

4.2.1 Limited Number of Patients

The major limitation of this study is the small number of patients, which is partially because the small number of patients that have taken 4D MR images and the suboptimal image quality. The low contrast of tumor and thus the low visibility in cine MR images result in the exclusion of a patient.

This inadequacy may leave potential statistically significant conclusions uncovered. The paired t-test, which is capable of yielding stronger conclusion than the two-tailed Wilcoxon signed-rank test, requires the data set to be normal distributed, which need to be tested first. This normality test is difficult for a data set of only five samples. Although not assuming normal distribution of data sets, the two-tailed Wilcoxon signed-rank test requires at least six pairs of data points to arrive at any statistically significant conclusion.

Under the assumption that the accuracy of the DVF does not depend on patient-specific characteristics, single-phase differences from all the five patients are analyzed as a whole, arriving at some statistically significant results. It should be noted that the validity of the assumption may be affected by different breathing patterns, different voxel sizes, etc.

4.2.2 Temporal Dependence of Patient's Breathing Patterns

An underlying assumption of this study is that patients have constantly regular breathing pattern independent of time, meaning information of ROI motion recorded in 4D CT, 4D MRI and cine MR images is identical. This assumption is, however, not exactly true. For some of the patients, they could maintain a relatively regular breathing period and amplitude, though small variation in period and amplitude can still be observed. For the other patients, apparent change in breathing period and amplitude were observed either in images of a single modality or between images of different modalities. This drift of baseline may contribute to the difference between DVF-based trajectories and the reference trajectories, complicating the evaluation of DIR accuracy. Breathing training is a possibly approach to reducing the variation of a patient's breathing pattern and should be considered as part of the protocol in future prospective studies.

4.3 Future work

A major limitation of this study is the small number of patients, making the discovery of potential statistically significant results extremely difficult. Despite the limited size of data, some trends are still observed and needed to be confirmed by a sample large enough for statistical analysis. Considering that an eligible patient for this study must have 4D CT, 4D MRI, and cine MR images at the same time, accruing sufficient number of eligible patients may take a fairly long time. I would suggest taking

orthogonal cine MR images at anatomical landmarks in addition to tumors in the liver. If the landmarks can be contoured on CT and MR images for the purpose of localization, these landmarks can be used as, to some extent, independent targets for evaluation of the accuracy of DVF. By doing so, the number of samples can be increased at a faster rate than the number of accrued patients.

There is also room for improvement of the strength of conclusion if the resolution of 4D MR images can be boosted to the same level as 4D CT images. As presented above, when measured in number of voxels, the results suggest that the accuracy of 4D MRI DVF may be higher than that of 4D CT DVF. Currently, the resolution of 4D MR images is mainly limited in the slice thickness. The 4D MR images used in this study were all acquired in the axial view. As a consequence, the large slice thickness in the SI direction, in which the respiratory motion is dominant, may undermine the accuracy of DIR. In the future, the 4D MR images can be acquired in the sagittal or coronal view such that the possible effect of coarse resolution on the accuracy of DIR can be excluded and more decisive results and conclusion are expected. In addition, the 4D MR images can be interpolated and re-sliced such that the 4D MR images have the same voxel size with the 4D CT images. A study comparing the DIR results from interpolated 4D MR images with those from original 4D MR images can be conducted to investigate whether DIR algorithms interpolate the input images.

It is also desirable to repeat the study using different DIR algorithms, considering that the accuracy of DVFs obtained from different algorithms varies significantly. Although for most commercially available DIR software, the registration process is more like a black box with no customizable parameters, DIR algorithms in open-source toolkits tend to provide flexibility of parameter optimization, which may considerably affect the accuracy of the DVF.[40] Therefore, the study should be repeated using more algorithms, including those allowing for parameter optimization.

5. Conclusion

We have implemented a novel approach for evaluating accuracy of DVFs based on 4D imaging for studying liver deformation using cine MR images as reference. Considerable variation in the distribution of the magnitude of deformation vectors in the ROIs encompassing tumors were observed between 4D CT DVF and 4D MR DVF, necessitating careful evaluation of the accuracy of the DVF before application of the DVF for research and clinical use. The discrepancy between the DVF and reference respiratory motion has been discovered to be moderately influenced by the magnitude of respiratory motion. Current results from patient-wise study suggest that DVFs from 4D CT and 4D MRI are comparable. Although not supported by statistically significant results, a trend that, if measured in number of voxels, the discrepancy in the motion of ROI between 4D MRI and cine MR imaging is smaller than that between 4D CT and cine MR imaging has been observed. When the registration for each two phases is considered as independent of patient-specific characteristics, 4D MRI shows possibly better DVF accuracy than 4D CT, when measured in number of voxels.

Further studies involving a sufficient number of patients for statistical analysis and using 4D MR images with resolution comparable with resolution of 4D CT images are warranted and expected to determine whether there is significant difference between the accuracy of 4D CT and 4D MRI DVFs and to confirm whether improved resolution of 4D MRI images can increase the accuracy of 4D MRI DVF.

References

1. Hall E. J., Giaccia A. J. *Radiobiology for the radiologist*. Lippincott Williams & Wilkins, 2011.

2. Fowler J. F. *The linear-quadratic formula and progress in fractionated radiotherapy*. Br. J. Radiol. 62:679-694, 1989.

3. Thames H. D., Hendry J. H. *Fractionation in Radiotherapy*. Taylor & Francis, 1987.

4. Dale R. G. *The application of the linear-quadratic dose-effect equation to fractionated and protracted radiotherapy*. Br. J. Radiol. 58:515-528, 1985.

5. Lyman J. T. *Complication probability as assessed from dose-volume histograms*. Radiat. Res. Suppl. 8:S13-S19, 1985.

6. Kutcher G. J., Burman C. *Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method*. Int. J. Radiat. Oncol. Biol. Phys. 16:1623-1630, 1989.

7. Schaly B., Kempe J. A., Bauman G. S., Battista J. J., Van Dyk J. *Tracking the dose distribution in radiation therapy by accounting for variable anatomy.* Phys. Med. Biol. 49:791–805, 2004.

8. Rosu M., Chetty I. J., Balter J. M., Kessler M. L., McShan D. L., Ten Haken R. K. *Dose reconstruction in deforming lung anatomy: dose grid size effects and clinical implications.* Med. Phys. 32:2487–2495, 2005.

9. Siebers J. V., Zhong H. *An energy transfer method for 4D Monte Carlo dose calculation.* Med. Phys. 35:4096–4105, 2008.

10. Zhong H., Siebers J. V. *Monte Carlo dose mapping on deforming anatomy.* Phys. Med. Biol. 54:5815–5830, 2009.

11. Peterhans M., Frei D., Manser P., Aguirre M. R., Fix M. K. *Monte Carlo dose calculation on deforming anatomy.* Z. Med. Phys. 21:113–123, 2011.

12. Heath E., Tessier F. and Kawrakow I. *Investigation of voxel warping and energy mapping approaches for fast 4D Monte Carlo dose calculations in deformed geometries using VMC++.* Phys. Med. Biol. 56:5187–5202, 2011.

13. Li H. S., Zhong H., Kim J., Glide-Hurst C., Gulam M., Nurushev T. S., Chetty I. J. *Direct dose mapping versus energy/mass transfer mapping for 4D dose accumulation: fundamental differences and dosimetric consequences.* Phys. Med. Biol. 59: 173–188, 2014.

14. Keall P. *4-dimensional computed tomography imaging and treatment planning*. Semin. Radiat. Oncol. 14:81-90, 2004.

15. Low D. A., Nystrom M., Kalinin E., Parikh P., Dempsey J. F., Bradley J. D., Mutic S., Wahab S. H., Islam T., Christensen G., Politte D. G., Whiting B. R. *A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing*. Med. Phys. 30:1254-1263, 2003.

16. Mageras G. S., Pevsner A., Yorke E. D., Rosenzweig K. E., Ford E. C., Hertanto A., Larson S. M., Lovelock D. M., Erdi Y. E., Nehmeh S. A., Humm J. L., Ling C. C. *Measurement of lung tumor motion using respiration-correlated CT*. Int. J. Radiat. Oncol. Biol. Phys. 60:933-941, 2004.

17. Pan T. *Comparison of helical and cine acquisitions for 4D-CT imaging with multislice CT*. Med. Phys. 32:627-634, 2005.

18. Rietzel E., Pan T., Chen G. T.. *Four-dimensional computed tomography: Image formation and clinical protocol*. Med. Phys. 32: 874-889, 2005.

19. Ford E. C., Mageras G. S., Yorke E., Ling C. C. *Respiration-correlated spiral CT: A method of measuring respiratory induced anatomic motion for radiation treatment planning*. Med. Phys. 30:88-97, 2003.

20. Cai J., Chang Z., Wang Z., Segars W. P., Yin F. F. *Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: A feasibility study.* Med. Phys. 38:6384, 2011.

21. Hu Y., Caruthers S. D., Low D. A., Parikh P. J., Mutic S. *Respiratory Amplitude Guided 4-Dimensional Magnetic Resonance Imaging.* Int. J. Radiat. Oncol. Biol. Phys. 86(1):198-204, 2013.

22. Tryggestad E., Flammang A., Han-Oh S., Hales R., Herman J., McNutt T., Roland T., Shea S.M., Wong J. *Respiration-based sorting of dynamic MRI to derive representative 4D-MRI for radiotherapy planning.* Med. Phys. 40, 2013.

23. Yang J., Cai J., Wang H., Chang Z., Czito B. G., Bashir M. R., Yin F. F. *Four-Dimensional Magnetic Resonance Imaging Using Axial Body Area as Respiratory Surrogate: Initial Patient Results.* Int. J. Radiat. Oncol. Biol. Phys. 88(4):907-912, 2014.

24. Hardcastle N., van Elmpt W., De Ruysscher D., Bzdusek K., Tomé W. A. *Accuracy of deformable image registration for contour propagation in adaptive lung radiotherapy*. Radiat. Oncol. 8:243, 2013.

25. Kadoya N., Fujita Y., Katsuta Y., Dobashi S., Takeda K., Kishi K., Kubozono M., Umezawa R., Sugawara T., Matsushita H., Jingu K. *Evaluation of various deformable image registration algorithms for thoracic images*. J. Radiat. Res. 55(1):175-182, 2014.

26. Brock K. K. *Results of a Multi-Institution Deformable Registration Accuracy Study (MIDRAS)*. Int. J. Radiat. Oncol. Biol. Phys. 76(2):583-596, 2010.

27. Latifi K., Zhang G., Stawicki M., van Elmpt W., Dekker A., Forster K. *Validation of three deformable image registration algorithms for the thorax*. J. Appl. Clin. Med. Phys. 14(1):3834, 2013.

28. Balik S., Weiss E., Jan N., Roman N., Sleeman W. C., Fatyga M., Christensen G. E., Zhang C., Murphy M. J., Lu J., Keall P., Williamson J. F., Hugo G. D. *Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy*. Int. J. Radiat. Oncol. Biol. Phys. 86(2):372-9, 2013.

29. Zhong H., Kim J., Chetty I. J. *Analysis of deformable image registration accuracy using computational modeling.* Med. Phys. 37(3):970–979, 2010.

30. Kashani R., Hub M., Balter J. M., Kessler M. L., Dong L., Zhang L., Xing L., Xie Y., Hawkes D., Schnabel J. A., McClelland J., Joshi S., Chen Q., Lu W. *Objective assessment of deformable image registration in radiotherapy: a multi-institution study.* Med. Phys. 35(12):5944–5953, 2008.

31. Schaly B., Baumen G. S., Battista J. J., Van Dyk J. *Validation of contour-driven thin-plate splines for tracking fraction-to-fraction changes in anatomy and radiation therapy dose mapping*. Phys. Med. Biol. 50(3):459–475, 2005.

32. Wang H., Dong L., O'Daniel J., Mohan R., Garden A. S., Ang K. K., Kuban D. A., Bonnen M., Chang J. Y., Cheung R. *Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy.* Phys Med Biol. 50(12):2887–2905, 2005.

33. Brock K. K. *Results of a multi-institution deformable registration accuracy study (MIDRAS).* Int J Radiat Oncol Biol Phys. 2010; 76(2):583–96.

34. Kaus M. R., Brock K. K., Pekar V., Dawson L. A., Nichol A. M., Jeffray D. A. *Assessment of a modelbased deformable image registration approach for radiation therapy planning.* Int. J. Radiat. Oncol. Biol Phys. 2007; 68(2):572–80.

35. Ostergaard Noe K., DeSenneville B. D., Elstram U. V., Tanderup K., Sarenson T. S. *Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy.* Acta. Oncol. 47(7):1286–93, 2008.

36. Zhang T., Chi Y., Meldolesi E., Yan D. *Automatic delineation of on-line head-and-neck computed tomography images: toward on-line adaptive radiotherapy.* Int. J. Radiat. Oncol. Biol. Phys. 68(2):522–530, 2007.

37. Castillo R., Castillo E., Guerra R., Johnson V. E., McPhail T., Garg A. K., Guerrero T. *A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets.* Phys. Med. Biol. 54(7):1849–70, 2009.

38. Jannin P., Fitzpatrick J. M., Hawkes D. J., Pennec X., Shahidi R., Vannier M. W. *Validation of medical image processing in image-guided therapy*. IEEE Trans. Med. Imaging. 21(12):1445-1449, 2002.

39. Liu F., Hu Y., Zhang Q., Kincaid R., Goodman K. A., Mageras G. S. *Evaluation of deformable image registration and a motion model in CT images with limited features*. Phys. Med. Biol. 57(9):2539-2554, 2012.

40. Yeo U. J., Supple J. R., Taylor M. L., Smith R., Kron T., Franich R. D. *Performance of 12 DIR algorithms in low-contrast regions for mass and density conserving deformation*. Med. Phys. 40(10):101701, 2013.