J Pathol Inform

Editor-in-Chief:
Anil V. Parwani , Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA

OPEN ACCESS
HTML format

For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

Original Article

3D prostate histology image reconstruction: Quantifying the impact of tissue deformation and histology section location

Eli Gibson[1,2], Mena Gaed[1,3,4], José A. Gómez[4], Madeleine Moussa[4], Stephen Pautler[3,5], Joseph L. Chin[5], Cathie Crukley[1,3], Glenn S. Bauman[6], Aaron Fenster[1,2,3,6,7], Aaron D. Ward[2,3,6,7]

[1]Robarts Research Institute, [2]Graduate Program in Biomedical Engineering, [4]Departments of Pathology, [5]Urology, [6]Oncology, and [7]Medical Biophysics, The University of Western Ontario, London, Ontario, [3]Lawson Health Research Institute, London, Canada

*E-mail: Eli Gibson - egibson@robarts.ca
*Corresponding author

Received: 26 June 13 Accepted: 03 August 13 Published: 31 October 13

This article may be cited as:
Gibson E, Gaed M, Gómez JA, Moussa M, Pautler S, Chin JL, *et al*. 3D prostate histology image reconstruction: Quantifying the impact of tissue deformation and histology section location. J Pathol Inform 2013;4:31.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2013/4/1/31/120874

Abstract

Background: Guidelines for localizing prostate cancer on imaging are ideally informed by registered post-prostatectomy histology. 3D histology reconstruction methods can support this by reintroducing 3D spatial information lost during histology processing. The need to register small, high-grade foci drives a need for high accuracy. Accurate 3D reconstruction method design is impacted by the answers to the following central questions of this work. (1) How does prostate tissue deform during histology processing? (2) What spatial misalignment of the tissue sections is induced by microtome cutting? (3) How does the choice of reconstruction model affect histology reconstruction accuracy? **Materials and Methods:** Histology, paraffin block face and magnetic resonance images were acquired for 18 whole mid-gland tissue slices from six prostates. 7-15 homologous landmarks were identified on each image. Tissue deformation due to histology processing was characterized using the target registration error (TRE) after landmark-based registration under four deformation models (rigid, similarity, affine and thin-plate-spline [TPS]). The misalignment of histology sections from the front faces of tissue slices was quantified using manually identified landmarks. The impact of reconstruction models on the TRE after landmark-based reconstruction was measured under eight reconstruction models comprising one of four deformation models with and without constraining histology images to the tissue slice front faces. **Results:** Isotropic scaling improved the mean TRE by 0.8-1.0 mm (all results reported as 95% confidence intervals), while skew or TPS deformation improved the mean TRE by <0.1 mm. The mean misalignment was 1.1-1.9° (angle) and 0.9-1.3 mm (depth). Using isotropic scaling, the front face constraint raised the mean TRE by 0.6-0.8 mm. **Conclusions:** For sub-millimeter accuracy, 3D reconstruction models should not constrain histology images to the tissue slice front faces and should be flexible enough to model isotropic scaling.

Key words: Correlative histopathology, image registration, prostate cancer imaging, validation, 3D histology reconstruction

Access this article online
Website: www.jpathinformatics.org
DOI: 10.4103/2153-3539.120874
Quick Response Code:

INTRODUCTION

The localization of aggressive prostate cancer tissue before therapy may support improved targeted biopsy and selection of treatment (whole gland vs. focal ablation) as well as the form of ablative energy.[1] Several *in vivo* imaging modalities are showing promise for staging and grading prostate cancer;[2-5] however, there is a knowledge gap in how to localize aggressive prostate cancer with these modalities.[6] The development of interpretation guidelines for localizing aggressive cancer on imaging ideally includes the comparison of *in vivo* images to spatially concordant post-prostatectomy histology images on which Gleason grade, an accepted surrogate for prostate cancer aggressiveness,[7] has been assessed. This comparison involves determining the 3D spatial relationship between corresponding histology and *in vivo* images, often performed in two steps: (1) a reconstruction of histology images to the 3D *ex vivo* spatial context and (2) an alignment of reconstructed histology to *in vivo* images.

The challenges in 3D histology reconstruction can be illustrated in the context of the process of collecting histology from radical prostatectomy specimens, which typically proceeds as follows [shown in the first row of Figure 1]. After surgery, the prostate is fixed in a formalin solution and then cut into 3-5 mm thick tissue slices at the pathology bench. These tissue slices proceed through a series of chemical baths to replace water in the tissue with paraffin and the slices are embedded in a block of translucent paraffin. This block is mounted to a microtome by hand, aligned by eye to square the tissue face (as seen through the translucent paraffin) with the microtome blade and tissue is cut until a full cross-section can be collected. Once a sufficient depth has been reached, the operator cuts a 4 μm histological section, allows it to expand on a water bath to flatten the section and mounts it on a glass slide.

The 3D reconstruction of histology consists of retroactively determining the positions of cutting and the deformation of the tissue to determine the original 3D spatial relationships of histological tissue, a process that remains an active area of research.[8-11] 2D to 3D deformable reconstructions for clinical prostate specimens have many degrees of freedom and sparse out-of-plane information content. A common approach to mitigate these challenges is to make simplifying assumptions about the spatial relationship of histological tissue to the corresponding tissue in the specimen. Some assumptions made in existing approaches for 3D histology reconstruction are enumerated below. The focus of this paper is on testing the strength of the first two assumptions in this list.

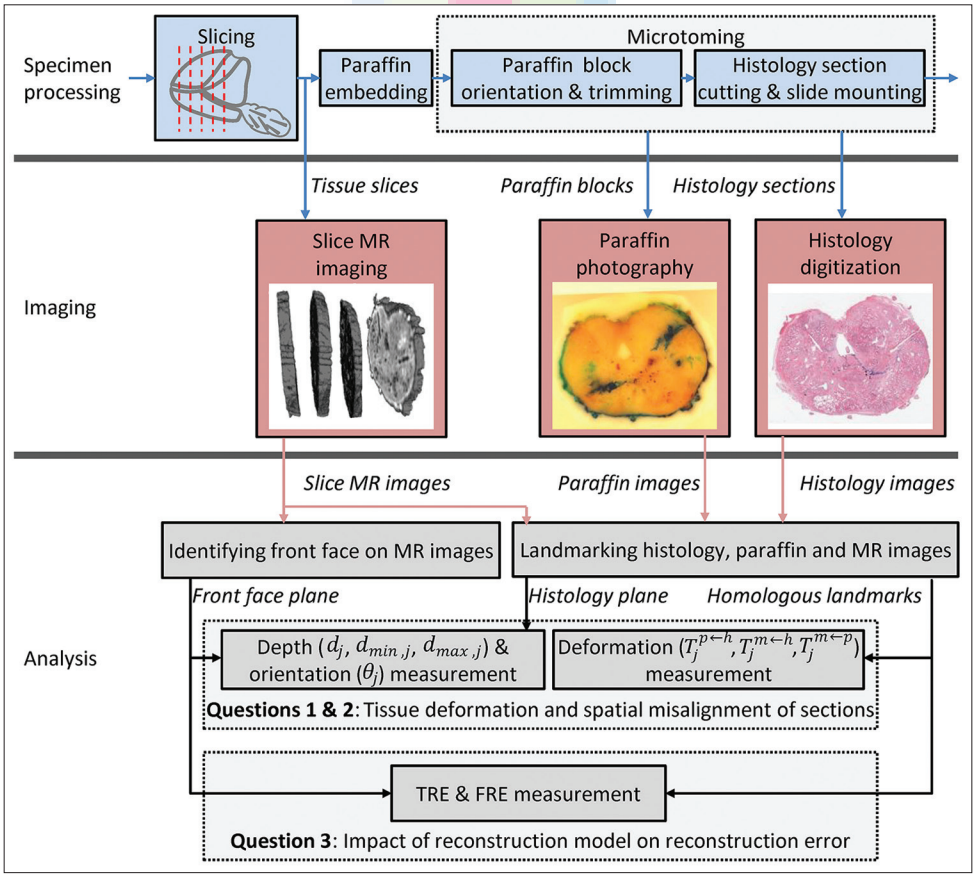

Figure 1: Overview of the specimen processing, imaging and analysis

- The deformation of the histological tissue after coarse slicing fits a specified constrained deformation model, such as the rigid,[8] rigid + isotropic scaling (referred to as similarity throughout this paper),[12,13] affine,[9,14] or thin-plate-spline (TPS)[15,16] deformation models. This is referred to as the *deformation model assumption* throughout this paper.
- Each histological section corresponds to the front face of the 3-5 mm thick tissue slice from which it was taken.[10,12-14,16-18] This is referred to as the *front face assumption* throughout this paper.
- Each histological section corresponds to a planar surface in the specimen.[8-10,14,17,18]
- Each histological section corresponds to a surface in the specimen defined prospectively by carefully controlling the position and orientation of the cuts made during specimen slicing.[10,17,19,20]
- The histological sections correspond to parallel, evenly spaced surfaces in the specimen (typically justified based on controlled cutting of tissue slices), which are determined retrospectively using additional imaging (typically photographs of the faces of tissue slices) during or after the slicing of specimens into thick tissue slices.[12,14,16,18]

By constraining the allowable spatial relationship between each histology section and the tissue from which it was cut, these assumptions may simplify the registration problem by decreasing the degrees of freedom but may also impact the accuracy of 3D reconstruction methods. In the context of imaging validation studies that evaluate imaging modalities by comparison to a 3D reconstructed histological reference, reconstruction accuracy affects the statistical power (i.e., the probability of a study finding an existing statistically significant effect) of studies that apply the reconstruction methods.[21] Because of this relationship, 3D histology image reconstructions with greater error create a requirement for more patients to be enrolled in the study, which can have a substantial impact on the cost of the study [see discussion for an illustrative case study] or run the risk of improperly evaluating the imaging modality if the study is underpowered for the error inherent in the technique. Thus, it is important to consider these simplifying assumptions, and their impact on reconstruction error, in the development and/or selection of reconstruction methods for such studies.

The strengths of these assumptions depend upon their fidelity to the processes the tissue undergoes throughout the preparation of histological sections. For example, the correspondence of the histological sections to the front faces of tissue slices depends in part on the skill and experience of the microtome operator determining the paraffin block face orientation and cutting depth. This task is complicated by the fact that the tissue face is hidden behind a translucent layer of paraffin, challenging the assessment of tissue face orientation until the tissue has already been exposed by the microtome blade and thus sectioned. The strengths of these assumptions and their impact on the reconstruction error have not, to the best of our knowledge, been quantified in the literature. This complicates the selection of appropriate assumptions for reconstruction method developers and also complicates the selection of reconstruction methods (that may incorporate such assumptions) for study designers.

In this work, our objective was to quantify the spatial relationships between histological sections, paraffin embedded blocks and the corresponding tissue slices from which the sections were taken to answer three questions, referred to according to the following enumeration throughout this paper.

- Question 1: How does prostate tissue deform during histology processing? Specifically, with what accuracy can rigid, similarity, affine or TPS deformation models align homologous landmarks on histology sections [Figure 2f, labeled $\boldsymbol{h}_{i,j}$], on the cut paraffin-embedded tissue blocks [Figure 2e, labeled $\boldsymbol{p}_{i,j}$] and on magnetic resonance (MR) images of formalin-fixed tissue slices [Figure 2c and d, labeled $\boldsymbol{m}_{i,j}$]. This question constitutes a test of the strength of the deformation model assumption.
- Question 2: What spatial misalignment of the tissue sections is induced by microtome cutting? Specifically, relative to the front faces of the tissue slices, from what depth [Figure 2c, labeled d_j] and at what orientation [Figure 2c, labeled θ_j] are histology sections taken? This question constitutes a test of the strength of the front face assumption.
- Question 3: How does the choice of reconstruction model affect the accuracy of histology reconstructions? A reconstruction model is defined in this paper by a choice of one of four deformation models (rigid, affine, similarity, TPS) and a choice of whether or not to make the front face assumption. This question resolves to the following two more specific questions: for all possible reconstruction models as defined above, (1) what are the target registration errors (TRE) and (2) the fiducial registration errors (FRE) of least squares best-fit landmark-based reconstructions? These questions constitute an evaluation of the impact on registration error of making the deformation model assumption and the front face assumption. The answers to these questions were quantified using homologous landmarks manually identified on histology images, paraffin block face images and MR images of the tissue slices, and the resulting errors were evaluated in the context of a type of imaging validation study that relies on histology image reconstruction. Preliminary experiments from this work were reported in a previous conference proceeding.[22]

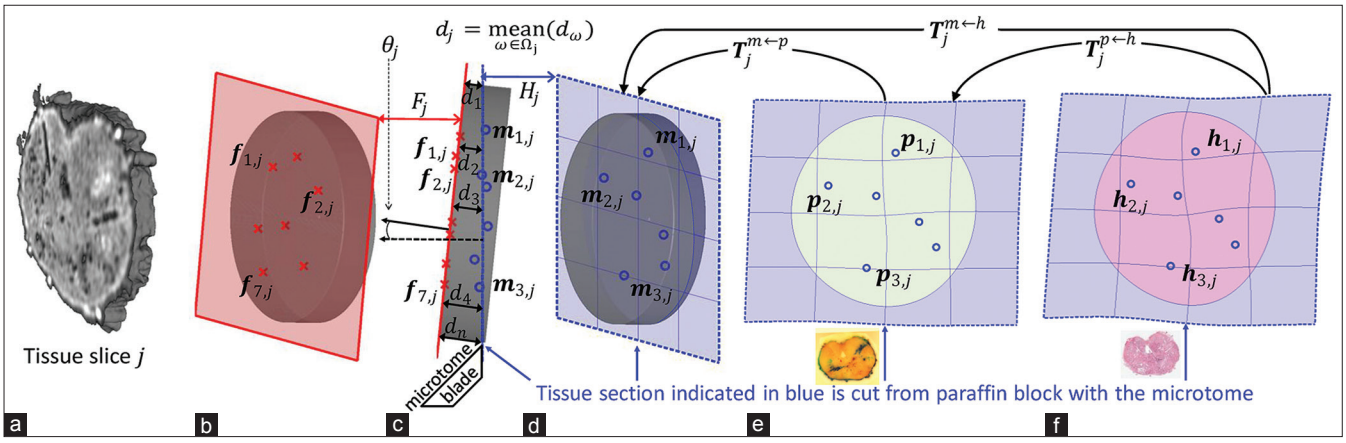

Figure 2: Schematic representations of tissue, landmarks and measurements, including (a) a surface rendering of a tissue slice magnetic resonance image, (b) a schematic rendering of the tissue slice with the front face fiducials $f_{i,j}$ and the best fit front face plane F_j, (c) a projected side view of the tissue slice as oriented over the microtome blade with the front face fiducials $f_{i,j}$, the best fit front face plane F_j, the histology-visible landmarks $m_{i,j}$, the best fit histology section plane H_j, the orientation θ_j and the depth measurement d_j, (d) a schematic rendering of the tissue slice with the histology-visible landmarks and the best fit histology section plane H_j, (e) a schematic rendering of the paraffin block face after histological sectioning showing the homologous landmarks $p_{i,j}$ and (f) a schematic rendering of the corresponding histology section with the homologous landmarks $h_{i,j}$

MATERIALS AND METHODS

Materials and Imaging

As part of an ongoing prospective imaging validation study, we obtained prostate specimens from 6 subjects after radical prostatectomy with the following inclusion criteria: (1) male, (2) age 18 years or older and (3) clinical prostate cancer stage T1 or T2 with histological confirmation from biopsy. The exclusion criteria were (1) prior therapy for prostate cancer, (2) use of 5-alpha reductase inhibitors within 6 months of the study start, (3) inability to comply with pre-operative imaging, (4) allergy to contrast agents, (5) sickle cell or other anemias, (6) hip prosthesis, (7) sources of artifact within the pelvis and (8) contraindications to MR imaging. This research was approved by our institutional human subjects research ethics board and informed consent was obtained from each subject.

An overview of the processing, imaging and measurement of these data is shown in Figure 1. After resection, fixation (10% buffered formalin for 48 h) and marking with fiducial strands,[10] the prostatic apex was removed and the mid-gland was gross-sectioned into 4.4-mm thick tissue slices (3-5/specimen, 21 total). MR images of these tissue slices were acquired using a Discovery MR750 (GE Healthcare, Waukesha, WI, USA) at 3T using an endorectal coil (Prostate eCoil, Medrad, Inc., Warrendale, PA, USA). Tissue slices were immobilized in tissue processing and embedding cassettes and immersed in Christo-Lube (Lubrication Technology Inc., Franklin Furnace, OH, USA) to provide a black background and minimize boundary artifacts on imaging. Imaging used a T1-weighted protocol (3D spoiled gradient recalled sequence, repetition time (TR) 6.5 ms, echo time (TE) 2.5 ms, bandwidth ± 31.25 kHz, eight averages, field of view (FOV) 14 cm × 14 cm × 6.2 cm, slice thickness 0.4 mm, 256 × 192 matrix, 312 slices, flip angle 15°, duration 25 min) and a T2-weighted protocol (3D fast spin echo sequence, TR 2000 ms, TE 151.5 ms, bandwidth ± 125 kHz, three averages, FOV 14 cm × 14 cm × 6.2 cm, slice thickness 0.4 mm, 320 × 192 matrix, 312 slices, duration 25 min). These images are referred to as *tissue slice MR images* throughout this paper.

Following MR imaging, formalin-fixed tissue slices were decalcified in a hydrochloric acid and chelating agent solution (Cal-Ex Decalcifier, Fisher Chemical, Ottawa, Canada) overnight and then dehydrated and embedded in paraffin using a series of chemical baths of formalin, ethanol, xylene and paraplast under our hospital's standard clinical pathology laboratory protocol for large blocks (wherein the duration of ethanol, xylene and paraplast are lengthened). The full processing schedule is given in Appendix A.

The embedded blocks were sectioned by one of fifteen clinical histotechnologists in our hospital's clinical pathology laboratory. Each block was mounted by hand on the chuck of a microtome (RM2245 or RM2255, Leica Biosystems, Nussloch, Germany) and the operator attempted to align the front face of the tissue, as seen through the semi-transparent paraffin covering, with the cutting axis by manually adjusting mechanical control knobs on the microtome. Sections were repeatedly cut until a full cross-section of the tissue block was reached. A final 4 μm section was cut from the block, floated on a hot water bath to remove distortion and mounted on a positively charged glass slide. All sections were stained with hematoxylin and eosin.

After the clinical pathology assessment was complete, stained histology sections were digitized on a ScanScope GL (Aperio Technologies, Vista, CA, USA) bright field slide scanning system with a 0.5 μm pixel size. These images are referred to as histology images throughout this paper.

Photographs of the exposed face of paraffin-embedded tissue blocks were acquired using a Pentax K200D with a 100 mm F2.8 macro lens (super-multi-coating Pentax-D FA, Pentax Imaging Company, Denver, CO, USA). The camera was attached to the camera-mount column of the photography table to ensure the optical axis was perpendicular to the tissue blocks. Labels containing a 4 mm long scale marker were affixed to the cut surfaces and used to calibrate the pixel size of the images. These images are referred to as paraffin images throughout.

Three sections were excluded from our analysis, because we identified an insufficient number of homologous landmarks on the three imaging modalities (histology, paraffin and tissue slice MR images). Five sets of homologous landmarks were necessary for our evaluation; see discussion for details.

Methods

Our method is illustrated in the "Analysis" portion of Figure 1 and is summarized at a high level as follows. To characterize prostate tissue deformation due to histology processing (Question 1), we determined the class of deformation that best mapped tissue on histology sections to the homologous tissue on the paraffin block faces and on the formalin-fixed tissue slices. Specifically, we assessed which of the four evaluated classes of transformation (rigid, similarity, affine or TPS) best mapped homologous landmarks on histology images [Figure 2f, labeled $\boldsymbol{h}_{i,j}$], on paraffin images [Figure 2e, labeled $\boldsymbol{p}_{i,j}$] and on tissue slice MR images [Figure 2c and d, labeled $\boldsymbol{m}_{i,j}$]. To measure the spatial misalignment of tissue sections induced by microtome cutting (Question 2), we characterized the locations from which histology sections were taken from within tissue slices, by estimating the depth [Figure 2c, labeled d_j] and the orientation [Figure 2c, labeled θ_j] relative to the front face of the tissue slice from which the histology sections were taken. To assess the impact of the choice reconstruction model (i.e., the choice of deformation model, plus the choice of whether or not to make the front face assumption) on 3D reconstruction error (Question 3), we estimated two reconstruction error measures, the TRE and the FRE, using different reconstruction models. The impact of these assumptions will depend, in part, on the reconstruction algorithm that is used. We used least-squares best-fit alignment of manually identified homologous intrinsic fiducials for these measurements. This approach is parameter-free, has an analytic solution for each considered reconstruction model and has an accuracy that depends only on the number and placement of fiducials and not on image properties. In the four subsections that follow, we describe the details of the selection of these landmarks, as well as the details of the methods used to address each of the three central questions of this work.

Identification of Landmarks and Tissue Slice Faces

All of these measurements rely on identifying homologous landmarks in histology images, paraffin images and tissue slice MR images. For each tissue slice, we identified 7-15 distinct landmarks (162 in total), comprising the centers of atrophic ducts, cysts and corpora amylaceae with diameters less than 1 mm. Illustrative examples of these images and identified homologous landmarks are shown in Figure 3. The positions of the landmarks (2D for histology and paraffin images, 3D for tissue slice MR images) on these modalities are denoted $\boldsymbol{h}_{i,j}$, $\boldsymbol{p}_{i,j}$ and $\boldsymbol{m}_{i,j}$, respectively, for the i^{th} landmark on the j^{th} tissue slice. Landmarks were interactively localized using 3D Slicer (Surgical Planning Lab, Harvard Medical School,

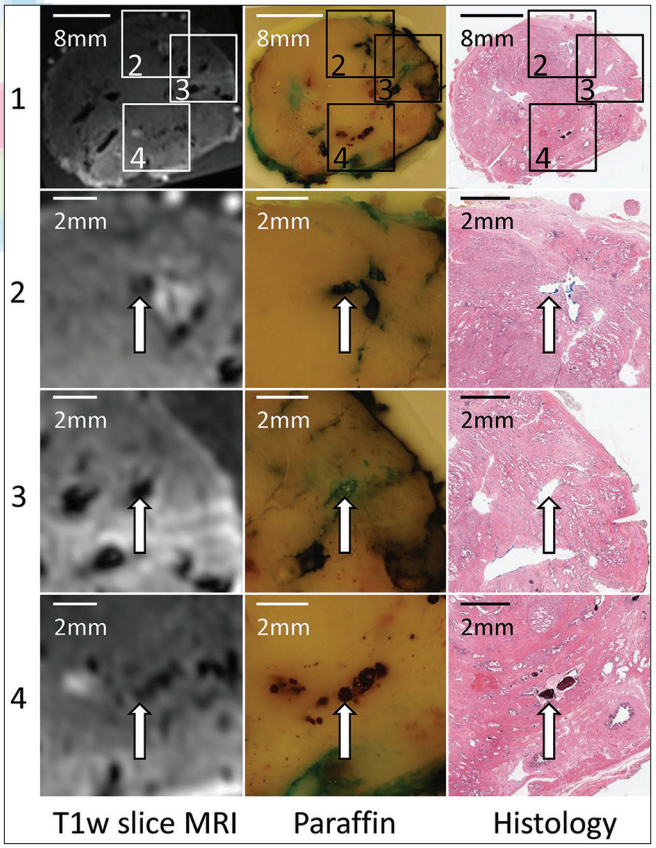

Figure 3: Illustrative examples of the T1-weighted tissue slice magnetic resonance (left), paraffin (middle) and histology (right) images transformed by a best-fit affine transformation aligning manually identified landmarks.The three highlighted regions in row 1 are shown magnified in rows 2, 3 and 4, with the corresponding landmarks denoted by arrows

Boston, USA), which required that the histology images be loaded into random-access memory; as the full resolution images typically occupy 15-20 GB, we downsampled them to a 30 μm × 30 μm voxel size for landmark identification, yielding images 10-40 MB in size.

We estimated the front face of the j^{th} tissue slice by first manually identifying seven 3D points $\{\boldsymbol{f}_{i,j} \mid i = 1\ldots7\}$ evenly distributed across the face of the j^{th} tissue slice on the tissue slice MR image [approximately in the configuration shown in Figure 2b] and then computing the least squares best-fit plane F_j to these points.

Because variability in landmark localization introduces uncertainty into spatial relationships measured in this work, the fiducial localization error (FLE) was estimated on histology images (denoted FLE_h), on paraffin images (denoted FLE_p) and on tissue slice MR images (denoted FLE_m). Based on previous measurements using these histology images and MR images of intact prostate specimens using the same protocols,[23] FLE_h and FLE_m were taken to be 0.05 mm and 0.16 mm respectively. These measurements quantified FLE as an unbiased estimator of the standard deviation (SD) of repeated localizations of landmarks. FLE_p was estimated to be 0.05 mm, the same as FLE_h, because the types of fiducials and the pixel sizes were similar (30 μm for histology images, 18 μm for paraffin images).

Tissue Deformation Due to Histology Processing (Question 1)

To assess the fidelity of the different deformation models, we quantified deformation between formalin-fixed tissue slices, paraffin blocks and histology sections under each of the models. Thus, we assessed the deformation due to three histology processes: (1) Paraffin processing and embedding (denoted with superscript $m \leftarrow p$, corresponding to the transformation $T_j^{m\leftarrow p}$ from the paraffin images [p] to the tissue slice MR images [m]), (2) histological sectioning and mounting (denoted with superscript $p \leftarrow h$, corresponding to the transformation $T_j^{p\leftarrow h}$ from the histology images [h] to the paraffin images [p]) and (3) the combination of both processes (denoted with superscript $m \leftarrow h$, corresponding to the transformation $T_j^{m\leftarrow h}$ from the histology images [h] to the tissue slice MR images [m]).

Deformation models were compared using the mean TRE of homologous landmark pairs identified on images before and after each process after landmark-based registrations constrained by four deformation models of increasing flexibility: rigid, similarity (rigid + isotropic scaling), affine (rigid + scaling + skewing) and non-linear TPS.[24] The mean TRE was estimated as the misalignment between homologous landmarks after transformation by a least squares best fit transformation constrained by the deformation model and was calculated using a leave-one-out cross-validation:

$$TRE^{t\leftarrow s,k} = \underset{j\in\{1\ldots J\},v\in\{1\ldots I_j\}}{\text{mean}} \left\| T^{t\leftarrow s,k}_{\hat{P}_{v,j},\hat{Q}_{v,j}}\left(\boldsymbol{s}_{v,j}\right) - \boldsymbol{t}_{v,j} \right\|, \quad (1)$$

where $T^{t\leftarrow s,k}_{\hat{P},\hat{Q}}$ is the transformation of type $k \in$ {rigid, similiarity, affine, TPS} that best maps the vector of image landmarks $\hat{Q}_{v,j} = \left\langle \boldsymbol{s}_{i,j} \mid i \in \{1\ldots I_j\}, i \neq v \right\rangle$ from the j^{th} slice on the source modality $s \in \{h, p\}$ to the vector of image landmarks $\hat{P}_{v,j} = \left\langle \boldsymbol{t}_{i,j} \mid i \in \{1\ldots I_j\}, i \neq v \right\rangle$ from the j^{th} slice on the target modality $t \in \{p, m\}$, $J = 21$ tissue slices and v denotes which landmark is left out. For example, $T^{m\leftarrow h,rigid}_{\hat{P}_{3,2},\hat{Q}_{3,2}}$ is the best-fit rigid transformation for the second tissue slice that maps from histology to tissue slice MR image coordinates that is fit to all but the 3rd identified fiducial.

The sensitivity of these measurements to FLE depends upon the spatial configuration and number of landmarks identified for each tissue slice. For example, the best fit transformation, and hence the leave-one-out TRE, of a section with few landmarks centrally clustered near the urethra and one landmark near the prostate boundary could be more sensitive to a misplaced landmark than that of a section with many widely spaced landmarks. Although relationships between TRE and FLE with respect to the spatial distribution of the landmarks have been characterized for rigid transformations,[25] to the best of our knowledge, there is no closed form solution for calculating this sensitivity for a leave-one-out TRE for all four deformation models. Thus, we assessed this sensitivity instead by Monte Carlo simulation. For each tissue slice, landmarks on histology, paraffin and tissue slice MR images were modeled as $\boldsymbol{h}_{i,j} + \mathbf{G}^{2D}$, $\boldsymbol{p}_{i,j} + \mathbf{G}^{2D}$ and $\boldsymbol{m}_{i,j} + \mathbf{G}^{3D}$, respectively, where $\mathbf{G}^{2D}$ is a 2D Gaussian random variable sampled for each landmark with x and y components distributed as $N\left(0, \frac{FLE_h^2}{2}\right)$ and $\mathbf{G}^{3D}$ is a 3D Gaussian random variable sampled for each landmark with x, y and z components distributed as $N\left(0, \frac{FLE_m^2}{3}\right)$. For each tissue slice, the TRE measurements were calculated for 5000 sets of perturbed landmarks and the SD of these measurements was calculated. The sensitivity of the TRE measurements to FLE was quantified as the average of these SDs across all tissue slices. The number of samples was chosen such that the standard error of the SD would be 1% of the SD itself.

Spatial Misalignment of Tissue Sections Induced by Microtome Cutting (Question 2)

To assess the strength of the front face assumption, we quantified the depth and orientation of histology sections relative to the front face of the tissue slices from which they were cut. The depth and orientation of each histology section were both estimated based on the

spatial relationship between two planes: the best fit plane F_j through the points identified on the front face of the tissue slice in the tissue slice MR image and the plane H_j, an estimate of the tissue from which the histology section was cut, computed as the best fit plane through the landmark points $\{\boldsymbol{m}_{i,j} \mid i = 1\ldots I_j\}$ in the tissue slice MR image corresponding to homologous landmarks visible on the histology image.

The orientation θ_j of the j^{th} histology section within the corresponding tissue slice was measured as the angle between the normal of plane F_j and plane H_j: specifically, $\theta_j = cos^{-1}\left(\left|\boldsymbol{n}\left(F_j\right)\cdot\boldsymbol{n}\left(H_j\right)\right|\right)$, where $\boldsymbol{n}(P)$ is the 3D unit normal of plane P.

The depth d_j of the j^{th} histology section from within the tissue slice was measured as the minimum, average and maximum distances from the tissue points Ωj (the intersection of plane H_j with tissue identified on the corresponding tissue slice MR image) to the front face plane F_j: specifically, $d_{\min,j} = \min_{\boldsymbol{\omega}\in\Omega_j}\left(D\left(F_j,\boldsymbol{\omega}\right)\right)$, $d_j = \text{mean}_{\boldsymbol{\omega}\in\Omega_j}\left(D\left(F_j,\boldsymbol{\omega}\right)\right)$ and $d_{\max,j} = \max_{\boldsymbol{\omega}\in\Omega_j}\left(D\left(F_j,\boldsymbol{\omega}\right)\right)$, where $D(P, \boldsymbol{\omega})$ is the distance from 3D point $\boldsymbol{\omega}$ to the plane P. Tissue points on the tissue slice MR image were identified by a threshold-based segmentation of the T1-weighted tissue slice MR image using a manually selected threshold, followed by manual editing.

The sensitivity of these measurements to FLE also depends on the spatial configuration and number of landmarks identified for each tissue slice. For example, the estimated orientation of a section with few landmarks centrally clustered near the urethra could be more sensitive to a misplaced landmark than that of a section with many widely spaced landmarks. Because, to the best of our knowledge, there is no closed form solution for calculating this sensitivity with respect to the spatial distribution of the landmarks, we assessed it instead by Monte Carlo simulation. For each tissue slice, histology-visible landmarks on tissue slice MR images were modeled as $\boldsymbol{m}_{i,j} + \mathbf{G}^{3D}$ and front face landmarks were modeled as $\boldsymbol{f}_{i,j} + \mathbf{G}^{3D}$. As was done for Question 1, the sensitivity of the depth and orientation measurements to FLE was measured as the mean, across all tissue slices, of the SD of each measure across 5000 sets of perturbed landmarks.

Impact of Reconstruction Model on 3D Reconstruction Error (Question 3)

For a reconstruction algorithm that uses a particular reconstruction model (i.e., a specified deformation model with or without the front face assumption), reconstruction accuracy may decrease if the true spatial relationships between histology sections and the tissue slices from which they were cut are different from the assumed constraints. Although the impact will depend on the 3D reconstruction methods used, it can be explored by examining reconstructions based on the least-squares best-fit transformation of identified homologous intrinsic landmarks under various reconstruction models. This reconstruction approach was chosen because the reconstructions are parameter-free, they can be solved analytically avoiding reconstruction errors due to local optima and their accuracy depends only on the number and placement of the fiducials and not on image properties.

The impact of the reconstruction model was quantified using the TRE (calculated in a leave-out-out manner). This is analogous to the TRE described in Eq. 1, but with an expanded set of transformation types $\ddot{K}$ that includes the deformation types {rigid, similiarity, affine, TPS} both with and without an additional constraint imposed by the front face assumption. Because the TPS transformation is an interpolating spline (i.e., source fiducials used to define the transformation are mapped exactly to target fiducials) and target fiducals may lie at a non-zero depth from the front face, the front face assumption cannot be directly applied. However, a transformation that does satisfy the front face assumption can be realized by first projecting the target fiducials used to define the transformation onto the front face and then defining a TPS transformation from source fiducials to the projected target fiducials. For rigid, similarity and affine transformations, the constrained least-squares fitting (constrained by the front face assumption) of transformed source fiducials to target fiducials is mathematically equivalent to the unconstrained least squares fitting of transformed source fiducials to the projected the target fiducials. Thus, for the reconstructions where the front face assumption was made, the target fiducials were projected onto the front face for all four deformation models. The sensitivity of TRE to FLE was quantified as for the TRE in Question 1.

In addition to quantifying the reconstruction error for these particular reconstructions, we can also calculate the lower bound on reconstruction error as measured by the identified landmarks for any possible reconstruction algorithm constrained by a particular reconstruction model. This lower bound is quantified as the FRE,

$$\text{FRE}^{t\leftarrow s,\ddot{K}} = \underset{j\in\{1\ldots J\},i\in\{1\ldots I_j\}}{\text{mean}}\left\|T^{t\leftarrow s,\ddot{K}}_{P_j,Q_j}\left(\boldsymbol{s}_{i,j}\right)-\boldsymbol{t}_{i,j}\right\|, \quad (2)$$

where $T^{t\leftarrow s,\ddot{K}}_{P_j,Q_j}$ is the transformation of type $\ddot{K}$ (the expanded set of transformation types described in the previous paragraph) that best maps the vector of image landmarks $Q_j = \left\langle \boldsymbol{s}_{i,j} \mid i\in\left\{1\ldots I_j\right\}\right\rangle$ from the j^{th} slice on the source modality $s \in \{h, p\}$ to the vector of image landmarks $P_j = \left\langle \boldsymbol{t}_{i,j} \mid i\in\left\{1\ldots I_j\right\}\right\rangle$ from the j^{th} slice on the

target modality $t \in \{p, m\}$ and $J = 21$ tissue slices. Note that unlike the TRE, the FRE includes all fiducials when fitting the transformation $T_{P_j,Q_j}^{t\leftarrow s,\ddot{K}}$ and represents a lower bound on the TRE as measured using the identified intrinsic landmarks. Because the TPS transformation is an interpolating spline (i.e., source fiducials used to define the transformation are mapped exactly to target fiducials) the FRE of an unconstrained TPS transformation is 0, by construction, for any configuration of fiducials. The sensitivity of FRE to FLE was quantified as for the TRE in Question 1.

Statistical Analysis

Statistical analyses were performed in SPSS 20 (IBM, Chicago, USA). The depth and orientation measurements were characterized with descriptive statistics (mean and SD) and 95% confidence intervals (CI) of the means were computed. Correlations of the depth measurements with the orientation were assessed using pairwise Spearman correlations.

The TRE measurements quantifying deformation during paraffin processing and embedding, histological sectioning and mounting and the combination of both processes were characterized with descriptive statistics. We assessed differences in mean TRE between the deformation models using separate 1-way repeated measures ANOVA tests with Greenhouse–Geisser correction for asphericity with the deformation model as the factor. Pairwise *post hoc* analysis of adjacent levels (i.e., rigid vs. similarity, similarity vs. affine and affine vs. TPS deformation models) was performed by constructing 95% CI on the differences in mean TRE.

To assess the impact of reconstruction assumptions, we assessed the differences in mean TRE using a 2-way repeated-measures ANOVA with a Greenhouse–Geisser correction for asphericity, with the two assumptions (the deformation model assumption and the front face assumption) as factors. Pairwise *post hoc* analysis was performed by constructing 95% CI on the difference in FRE and TRE due to the front face assumption under each deformation model and on the pairwise differences due to the deformation model between adjacent levels with and without the front face assumption. Note that FRE of a reconstruction under stricter assumptions is mathematically guaranteed to be equal to or higher than the FRE under relaxed assumption. For example, the rigid deformation assumption is stricter than the affine deformation assumption and the front face assumption is stricter than eliminating that assumption.

RESULTS

Tissue Deformation Due to Histology Processing (Question 1)

The key finding regarding tissue deformation was that modeling isotropic scaling as in the similarity deformation model improved the mean TRE by 0.8-1.0 mm, while modeling skew or TPS deformation improved the mean TRE by less than 0.1 mm [bolded intervals in the first row of Table 1]. The mean and SD of TRE for the three histology processes under the four deformation models are shown in Table 2. For the combined deformation from tissue slice to histology section [shown in the first row of Tables 1 and 2 and as a box plot in Figure 4], the similarity model had a significantly lower mean TRE than the rigid model and the affine model has a significantly lower mean TRE than the similarity model (by 0.9 mm and 0.06 mm differences respectively), but *post hoc* analyses failed to show a statistically significant difference between the affine and TPS deformation models (0.005 mm difference).

For the intermediate deformation due to paraffin embedding [second row of Tables 1 and 2], the rigid model had a significantly higher mean TRE than the similarity model (by a 1.2 mm difference), but *post hoc* analyses failed to show a statistically significant difference between the other models (differences < 0.03 mm). For the intermediate

Table 1: *Post hoc* analyses comparing mean TRE under varying deformation models: 95% CI of mean TRE for model A-model B

Deformation model A Deformation model B	Rigid Similarity	Similarity Affine	Affine TPS
Tissue slice MR to histology images	**(0.78, 0.98)**	**(0.03, 0.10)**	**(−0.02, 0.02)**
Tissue slice MR to paraffin images	(1.07, 1.31)	(−0.04, 0.01)	(−0.03, 0.01)
Paraffin to histology images	(0.12, 0.20)	(0.02, 0.07)	(−0.01, 0.02)

TRE: Target registration error, CI: Confidence interval, MR: Magnetic resonance, TPS: Thin-plate-spline. Key findings are shown in bold

Table 2: Mean (SD) TRE (mm) for four models of deformation during histological processing stages

Deformation model	Rigid	Similarity	Affine	Thin plate spline
Tissue slice MR to histology images	1.44 (0.73)	0.56 (0.31)	0.50 (0.28)	0.50 (0.28)
Tissue slice MR to paraffin images	1.71 (0.82)	0.54 (0.26)	0.54 (0.28)	0.54 (0.28)
Paraffin to histology images	0.42 (0.27)	0.26 (0.18)	0.21 (0.14)	0.21 (0.14)

SD: Standard deviation, TRE: Target registration error, MR: Magnetic resonance. Statistical comparisons (performed between adjacent columns) where the statistical tests failed to detect a significant difference are connected by lines

deformation due to histological sectioning [third row of Tables 1 and 2], the affine model had a significantly lower mean TRE (by 0.05 mm) than the similarity model and the similarity model had a significantly lower mean TRE (by 0.2 mm) than the rigid model. *Post hoc* analyses failed to show a statistically significant difference between the affine and TPS deformation models (0.009 mm difference). The sensitivities of the TRE to the observed FLE ranged from 0.05 to 0.13 mm.

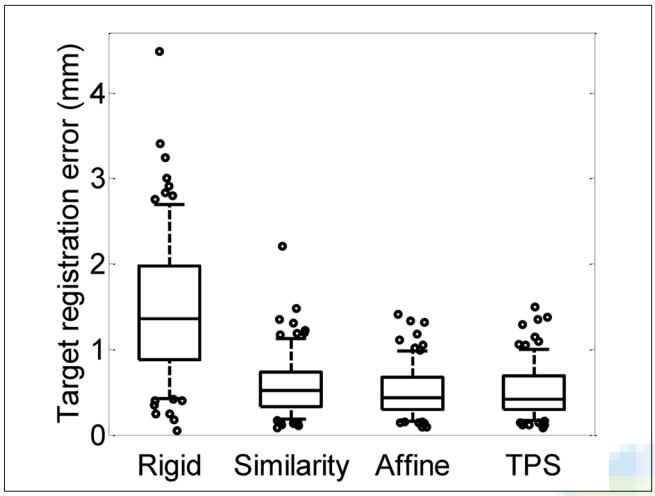

Figure 4: Boxplot showing the target registration errors of homologous landmarks under four deformation models for the tissue deformation due to histological processing and cutting. These results correspond to the descriptive statistics shown in the first row of Table 1

Spatial Misalignment of Tissue Sections Induced by Microtome Cutting (Question 2)

The key finding regarding the spatial misalignment of tissue sections was that the 95% CI on the mean of orientation was 1.1-1.9° and the 95% CI on the mean of mean depth was 0.9-1.3 mm [bolded intervals in the second column of Table 3]. The distributions of depth and orientation measurements are shown in Figure 5 and the correlation plots of orientation with the minimum (Spearman $r = -0.4$), mean (Spearman $r = 0.4$) and maximum (Spearman $r = 0.75$) section depth measures are shown in Figure 6. A subset of the tissue slices, chosen to illustrate the range of depths and orientations, are shown in Figure 7 with the front face F_i and the best fit plane to H_i superimposed. The SD, 95% CI on the mean and sensitivity to FLE for the orientation, minimum section depth, mean section depth and maximum section depth are shown in Table 3.

Impact of Reconstruction Model on 3D Reconstruction Error (Question 3)

The two key findings regarding the impact of the reconstruction model on 3D reconstruction error were as follows. (1) Modeling isotropic scaling (as in the similarity deformation model) improved the mean TRE by 0.5-0.7 mm if the front face assumption was made and by 0.8-1.0 mm if the front face assumption was not made, but modeling skew or TPS deformation improved mean TRE by less than 0.1 mm [bolded intervals in Table 4]. (2) Under a similarity deformation model, the front face assumption

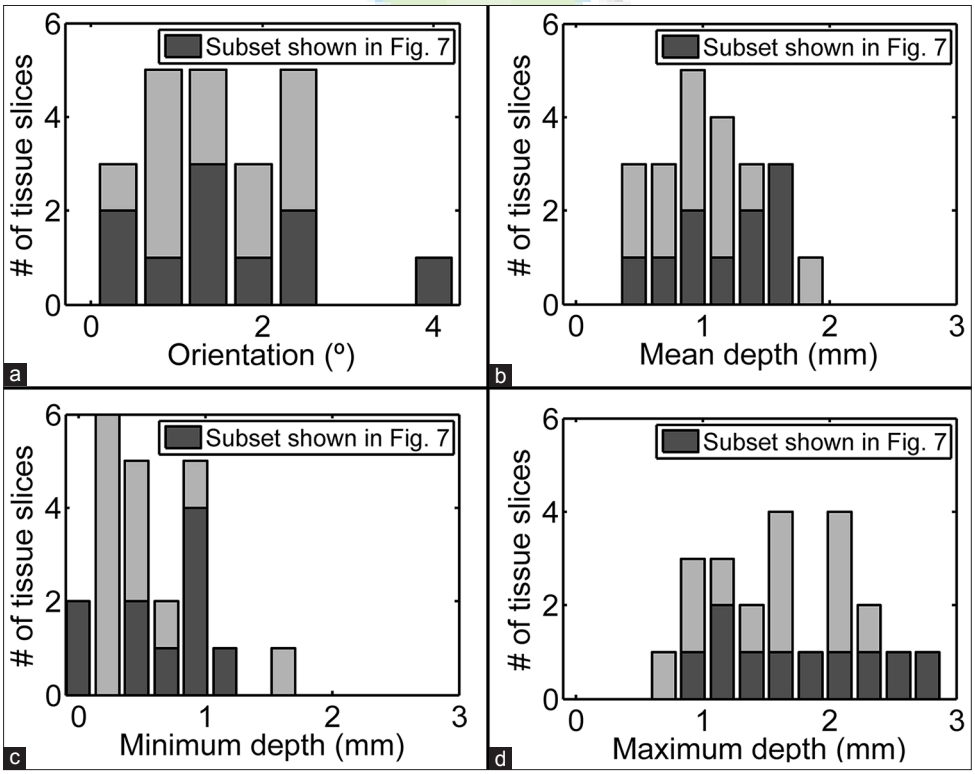

Figure 5: Histograms of histology section depths and orientations. The subset of tissue slices illustrated in Figure 7 is shown in dark gray

increased the mean TRE by 0.6 mm to 0.8 mm [bolded interval in Table 5]. The mean and SD of TRE and FRE for the intrinsic landmark-based reconstructions are shown in Tables 6 and 7, respectively. The 95% CI for the difference in TRE and FRE due to deformation model and due to the front face assumption are shown in Tables 4 and 5 respectively. The sensitivities of the TRE to the observed FLE ranged from 0.10 to 0.13 mm. The sensitivities of the FRE to the observed FLE were 0 mm (by construction) for the reconstruction model comprising the TPS deformation model without the front face assumption and 0.07-0.09 mm for the remaining reconstruction models.

DISCUSSION

In vivo prostate imaging is increasingly being validated against 3D reconstructed histology images.[26-29] Many algorithms for 3D reconstruction limit the degrees of freedom by making simplifying assumptions about the cutting of histology sections from the prostate gland which may affect the accuracy of reconstruction. This work explored two such assumptions: the deformation model assumption that histology sections have been deformed under a specified deformation model relative to the fixed tissue and the front face assumption that histology section corresponds to the front face of the tissue slice from which it was cut. Operator variability in sectioning could lead to histology sections that are not taken coincident with or parallel to the front face of the tissue slice and the cumulative deformation of the histology section due to dehydration, cutting, water bath expansion and slide-mounting processes may not be accurately modeled by the chosen transformation. In this work, we quantified the spatial relationship between histology images and the formalin-fixed tissue slices from which they were taken and evaluated the impact of the reconstruction model assumptions on 3D reconstruction error.

Tissue Deformation Due to Histology Processing (Question 1)

Modeling the deformation due to paraffin processing and histological sectioning as affine deformation yielded the lowest mean TRE (0.5 mm), although the difference between the affine and similarity models was 0.06 mm and our analysis failed to show a statistically significant difference between the affine and TPS models. The 0.5 mm TRE under the affine and TPS deformation model is larger than the 0.2 mm FLE, suggesting

Table 3: Descriptive statistics for the orientation, minimum depth, mean depth and maximum depth of histology sections relative to the tissue blocks from which they were cut

Statistic	Standard deviation	95% CI of the mean	Sensitivity to FLE
Orientation (°)	0.9	**(1.1, 1.9)**	0.30
Minimum depth (mm)	0.4	(0.4, 0.7)	0.12
Mean depth (mm)	0.4	**(0.9, 1.3)**	0.05
Maximum depth (mm)	0.6	(1.4, 1.9)	0.13

CI: Confidence interval, FLE: Fiducial localization error. Key findings are shown in bold. Sensitivity to FLE was quantified as the mean of the standard deviation of measurements in a Monte Carlo simulation with perturbed landmark positions

Table 4: *Post hoc* analyses comparing TRE/FRE after intrinsic landmark reconstruction under varying deformation models: 95% CI of mean TRE/FRE for model A-model B

Deformation model A Deformation model B	Rigid Similarity	Similarity Affine	Affine TPS
95% CI of mean TRE (mm) with front face assumption	**(0.54, 0.70)**	**(0.01, 0.06)**	**(−0.01, 0.01)**
95% CI of mean TRE (mm) without front face assumption	**(0.78, 0.98)**	**(0.03, 0.10)**	**(−0.02, 0.02)**
95% CI of mean FRE (mm) with front face assumption	(0.47, 0.62)	(0.03, 0.06)	(0.04, 0.07)
95% CI of mean FRE (mm) without front face assumption	(0.75, 0.94)	(0.07, 0.13)	(0.29, 0.35)[1]

[1]The FRE after an unconstrained thin plate spline transformation is 0 by construction. TRE: Target registration error, FRE: Fiducial registration error, CI: Confidence interval, TPS: Thin-plate-spline. Key findings are shown in bold

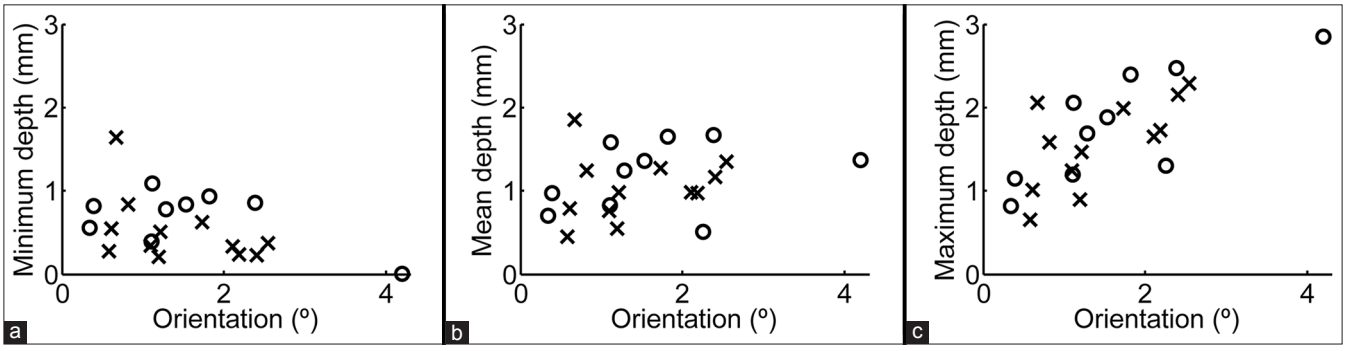

Figure 6: Correlation of minimum, mean and maximum histology section depths with orientations. Tissue slices corresponding to sections marked with circles are shown in Figure 7

Table 5: *Post hoc* analyses comparing TRE/FRE after intrinsic landmark reconstruction under varying deformation models: 95% CI for reconstruction with-without front face assumption

Deformation model	Rigid	Similarity	Affine	TPS
95% CI of mean TRE (mm)	(0.36, 0.46)	**(0.60, 0.75)**	(0.63, 0.77)	(0.62, 0.77)
95% CI of mean FRE (mm)	(0.40, 0.51)	(0.68, 0.83)	(0.74, 0.89)	(1.00, 1.15)[1]

[1]The FRE after an unconstrained thin plate spline transformation is 0 by construction. TRE: Target registration error, FRE: Fiducial registration error, CI: Confidence interval, TPS: Thin-plate-spline. Key findings are shown in bold

Table 6: Mean (SD) TRE after intrinsic landmark reconstruction under varying constraints

Deformation model	Rigid	Similarity	Affine	Thin plate spline
With front face assumption	1.85 (0.71)	1.23 (0.48)	1.20 (0.47)	— 1.20 (0.47)
Without front face assumption	1.44 (0.73)	0.56 (0.31)	0.50 (0.28)	— 0.50 (0.28)

TRE: Target registration error, SD: Standard deviation. Statistical comparisons (performed between adjacent columns and rows) where the statistical tests failed to detect a significant difference are connected by lines

Table 7: Mean (SD) FRE after intrinsic landmark reconstruction under varying constraints

Deformation model	Rigid	Similarity	Affine	Thin plate spline
With front face assumption	1.72 (0.66)	1.18 (0.48)	1.13 (0.48)	1.08 (0.50)
Without front face assumption	1.27 (0.66)	0.42 (0.25)	0.32 (0.19)	0 (0)[1]

[1]The FRE after an unconstrained thin plate spline transformation is 0 by construction. FRE: Fiducial registration error, SD: Standard deviation. All statistical comparisons (performed between adjacent columns and rows) showed significant differences

that there is some sub-millimeter-scale non-affine deformation that occurs, but that is not well-captured by the interpolation of the TPS deformation model with the landmark configurations identified in this work.

The analysis of deformation from tissue slice MR to paraffin images suggests that most of the deformation during the paraffin processing is characterized by isotropic scaling, which is consistent with the dehydration that occurs during this process. The analysis of deformation from paraffin to histology images suggests that some further affine deformation occurs during histological sectioning, which is consistent with anisotropic cutting forces that are applied during sectioning. The larger

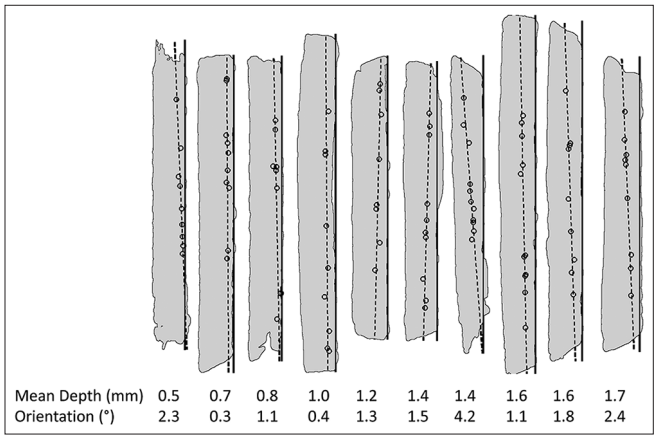

Figure 7: Renderings of the spatial relationships between tissue slices, histology-visible landmarks, front face plane and histology section planes for 10 tissue slices, ordered by increasing mean depth from left to right. Each tissue slice is shown as a silhouette projected along $n(F_j) \times n(H_j)$, the cross-product of the front face and histology section plane normals. With this projection, the front face plane F_j and the histology section plane H_j can be represented as solid and dashed lines respectively. The projected histology-visible landmarks $m_{i,j}$ are shown as circles

mean TRE across all deformation models and the larger change in TRE with isotropic scaling for the paraffin processing compared to the microtoming suggests that paraffin processing is the source of most of the observed deformation. Notably, the mean TRE of the combined processes under the rigid deformation model is less than that of the paraffin processing alone, which is consistent with expansion on the water bath partially cancelling out contraction due to dehydration during paraffin processing.

To perform these analyses, at least five sets of homologous landmarks were required; in particular, four sets of landmarks are required for a 2D-3D TPS transformation to define a non-affine transformation and a fifth is needed to enable the leave-one-out evaluation. Three histology sections were omitted from the analysis because fewer than five sets of landmarks were identifiable.

Spatial Misalignment of Tissue Sections Induced by Microtome Cutting (Question 2)

The histology sections were taken at a mean depth of 1.1 mm and were taken at an average angle of 1.5° relative to the front face. To illustrate these values, a histology section cut with the mean depth and the mean orientation from a hypothetical tissue slice 30 mm in diameter (typical for our sample of tissue slices), would be 0.7 mm from the front face at the closest point and 1.5 mm from the front face at the furthest point.

The SDs of the orientation (0.9°) and depth measurements (0.4 mm) are greater than would be expected due to the 0.16 mm FLE alone, suggesting that there is operator variability in the alignment of

the tissue block face with the microtome blade and in the depth of cutting. The variability in the minimum section depth suggests that the variability in depth is not directly caused by variability in tissue block alignment; if the variability in depth of cutting were the result of variability in tissue block alignment followed by consistently cutting until a full cross-section of tissue were barely reached, we would expect the minimum depth to have low variability. We speculate that the continued cutting beyond the best-fit front face plane could be due to concavity of the tissue front face that can be introduced during paraffin embedding, which would require a deeper cut to achieve a full face. This continued cutting could also be due to the practice of removing the paraffin block from the microtome to cool the cutting surface with ice, leading to variability in the orientation when the block is replaced.

The impact of the observed variability in depth and orientation on the relative spatial relationship of histology sections in a 3D reconstruction can be seen in Figure 8, where the tissue slices were sliced to be parallel at an even spacing (by embedding the specimen in an agar gel and cutting it on a rotary slicer), but after an alignment of tissue slice MR images (and accompanying registered histology sections) with an MR image of the intact *ex vivo* specimen with a 0.5 mm TRE, the non-parallel and uneven spacing of the three midgland histology images can be seen.

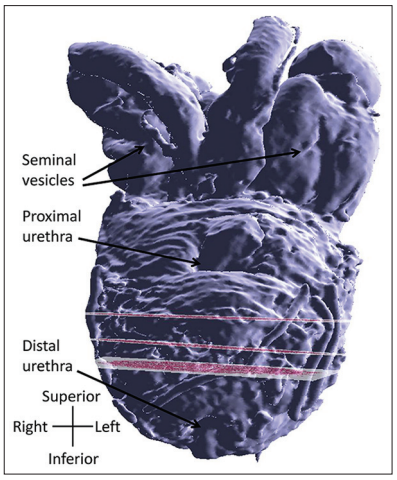

Figure 8: 3D reconstruction of three histology sections aligned (with a mean TRE of 0.5 mm) to an anterior view of 3D surface rendering of the corresponding intact *ex vivo* prostate gland with seminal vesicles, illustrating the potential for non-parallel, non-evenly spaced histological tissue sections

Impact of Reconstruction Model on 3D Reconstruction Error (Question 3)

If a reconstruction algorithm used a reconstruction model wherein histology sections corresponded to the front faces of tissue slices, the lower bound of the achievable mean TRE for any of the tested deformation models would be 1.1 mm (for the affine and TPS deformation models), suggesting that to achieve sub-millimeter reconstruction error, the front face assumption should not be made. For a reconstruction model unconstrained by the front face assumption using an affine deformation model, the lower bound of the mean TRE is 0.3 mm (the corresponding lower bound of the mean TRE for the TPS model is 0 mm by construction and therefore does not provide for an informative comparison). The improvement in TRE for modeling isotropic scaling (from rigid to similarity deformation models) was 0.78-0.98 mm when the front face assumption was not made, but the improvement from the similarity deformation model to the more flexible affine and TPS deformation models was less than 0.1 mm, suggesting that a similarity transformation may be sufficient. Note that this reconstruction error is only a component of the overall registration error from the histology images to *in vivo* images; Groenendaal *et al.*[12] and Orczyk *et al.*,[30] presented two methods for the registration of reconstructed *ex vivo* images to *in vivo* images with reported errors of 2.1 mm and 1.6 mm, which would, under the assumption that these errors are independent, be added in quadrature with reconstruction error.

It is important to interpret these reconstruction errors in the context of the application in which the reconstructions could be used. A recent model quantifying the impact of registration error on the statistical power (and thus the required sample size) of imaging validation studies[21] can be used to relate the differences in reconstruction error in these experiments to an appropriate application. This can be illustrated through the scenario of an imaging validation study testing for differences between the mean imaging signal of tumors and normal tissue regions, under the assumptions that the tumors are spherical foci of the smallest clinically significant volume (0.2 cm^3)[31] and that reconstruction error can be modeled as an isotropic Gaussian. For this scenario, our reconstruction is combined with a registration of reconstructed histology to *in vivo* images with mean TRE 2.1 mm (as reported by Groenendaal *et al.*[12]) In this scenario, we can compare the required sample sizes for the imaging validation study over a range of reconstruction errors as compared to an arbitrarily chosen baseline. In this illustration, we use landmark-based reconstruction under the similarity deformation model and the front face assumption with a mean TRE of 1.23 mm as the baseline. The relative required sample sizes for reconstruction errors ranging from 0.3 to 2.0 mm mean TRE are shown in Figure 9, with each of the assessed combinations of assumptions marked. Between the worst and best performing sets of assumptions, there is a 1.5-fold difference in required sample size. Based on per-patient costs of $10,000, from an ongoing imaging validation study in our center with 66 subjects, a 1.5-fold reduction in sample size could yield savings of $220,000 for the same statistical power.

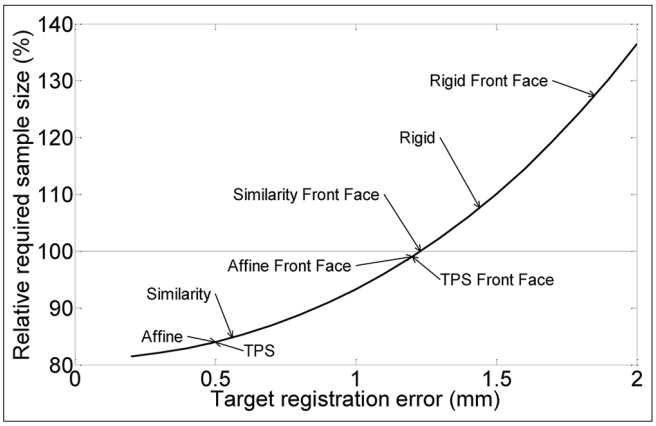

Figure 9: Sample sizes, relative to an arbitrarily chosen baseline, for imaging validation studies of the image signal differences between cancerous and background tissue for 0.2 cm^3 cancer foci, under assumptions that foci are spherical and reconstruction error can be modeled as a translation error distributed as a 3D Gaussian and is combined in quadrature with a 2.1 mm TRE due to registration to *in vivo* imaging. Reconstructions under differing deformation assumptions and with or without the front face assumption are indicated, with the reconstruction using a similarity transform and the front face assumption arbitrarily chosen as the 100% baseline reference

Limitations of Cutting Measurement

The conclusions of this work should be considered in the context of the limitations of the performed experiments. This study had four notable limitations. First, the process of sectioning tissue for histological examination varies between laboratories and many aspects of the processing can affect the sectioning distortion and possibly the amount of trimming before a full cross-section is successfully cut. Examples include tissue type, embedding medium, water bath duration and temperature,[28] knife quality and angle,[32] and possibly operator skill. It is also unclear how these results would generalize when prostate tissue slices are cut into quarters before paraffin processing, an approach adopted in many clinical laboratories. Although this study used histological sections sectioned by multiple histotechnologists, the use of a consistent processing protocol in a single clinical laboratory prevented us from assessing the impact of these other factors on the identified spatial relationships. A second limitation of the study is that the fidelity of the TPS deformation to the true underlying deformation is limited in part by the number of homologous landmarks identifiable on tissue slice MR, histology and paraffin block face images. It is not clear if the 7-15 landmarks identified per section are sufficient to characterize the unknown underlying deformation. If 7-15 landmarks are too few, then the TPS model may not capture the underlying deformation (resulting in a higher reported TRE) even if the deformation could be well-described by a TPS model. A third limitation of the study is that only one non-linear deformation model was assessed in this study, although there are an infinite number of such models. This work does, however, suggest an upper bound of less than 0.5 mm for the possible improvement that could be derived from better non-linear deformation models. Fourth, our assessment of the impact of the reconstruction errors on the statistical power of imaging validation studies examines the histology image reconstruction in isolation; if the reconstruction were followed by additional processing, such as 3D image registration, the impact of the reconstruction errors would be challenging to isolate and was not assessed by this study.

CONCLUSIONS

This work addressed three questions, as follows. (1) How does prostate tissue deform during histology processing? (2) What spatial misalignment of the tissue sections is induced by microtome cutting? (3) How does the choice of reconstruction model affect the accuracy of histology reconstructions? The key conclusion from these investigations is that for accurate 3D reconstruction of whole-mount histology, the reconstruction model should not assume that histology corresponds to the front face of the tissue slices from which it was cut because such an assumption yields a higher mean TRE by 0.6 to 0.8 mm and should use a similarity deformation model because the mean TRE under this model is 0.5 to 0.7 mm lower than that of a rigid deformation model and within 0.1 mm of the affine and TPS deformation models with more degrees of freedom. The mean TRE of 0.56 mm was measured for the least-squares best fit fiducial-based reconstruction using a similarity deformation model without the front face assumption. In addition, our characterization of the misalignment of histology sections revealed a mean section depth of 1.1 mm (with maximum depths as high as 2.8 mm) and a mean section orientation of 1.5° (with orientations as high as 4.2°), which may support commensurate heuristics in 3D reconstruction. Finally, in the context of imaging validation studies testing for imaging signal differences between cancerous and background tissue for the smallest clinically significant prostate cancer foci by correlation with reconstructed histology images, the range of reconstruction errors seen in this work would result in a 1.5-fold difference in the required sample size for such a study under our modeling assumptions, potentially translating to a difference of hundreds of thousands of dollars.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada, Cancer Care Ontario, the Ontario Graduate Scholarship Program, the Ontario Institute for Cancer Research, and the Canadian Institutes of Health Research [CTP 87515]. A. Fenster holds a Canada Research Chair in Biomedical Engineering and acknowledges the support of the Canada Research Chair Program. A. D. Ward holds a Cancer

Table A1: Our hospital's standard clinical pathology laboratory protocol for large specimens

Solution	Duration (h)	Temperature (°C)	Vacuum (inches)	Pressure (PSI)
80% ethanol, 20% formalin	1	40	15	7
95% ethanol, 5% formalin	2	40	15	7
100% ethanol	1	40	15	7
100% ethanol	2	40	15	7
100% ethanol	2	40	15	7
100% xylene	1	40	15	7
100% xylene	1	40	15	7
100% xylene	2	40	15	7
100% paraplast	1	60	15	7
100% paraplast	1	60	15	7
100% paraplast	1	60	15	7
100% paraplast	1	60	15	7

PSI: Pounds per square inch

Care Ontario Research Chair in Cancer Imaging.

Appendix A

Our hospital's standard clinical pathology laboratory protocol for large specimens consists of processing tissue slices through baths of graded alcohols, xylene and paraplast on a Tissue-Tek vacuum infiltration tissue processor (Sakura Finetek USA, Inc., Torrance, USA) following the schedule shown in Table A1.

REFERENCES

1. Turkbey B, Albert PS, Kurdziel K, Choyke PL. Imaging localized prostate cancer: Current approaches and new developments. AJR Am J Roentgenol 2009;192:1471-80.
2. Sciarra A, Barentsz J, Bjartell A, Eastham J, Hricak H, Panebianco V, *et al*. Advances in magnetic resonance imaging: How they are changing the management of prostate cancer. Eur Urol 2011;59:962-77.
3. Vargas HA, Akin O, Franiel T, Mazaheri Y, Zheng J, Moskowitz C, *et al*. Diffusion-weighted endorectal MR imaging at 3 T for prostate cancer: Tumor detection and assessment of aggressiveness. Radiology 2011;259:775-84.
4. Engelbrecht MR, Barentsz JO, Jager GJ, van der Graaf M, Heerschap A, Sedelaar JP, *et al*. Prostate cancer staging using imaging. BJU Int 2000;86 Suppl 1:123-34.
5. Ellis RJ, Kaminsky DA. Fused radioimmunoscintigraphy for treatment planning. Rev Urol 2006;8 Suppl 1:S11-9.
6. Kelloff GJ, Choyke P, Coffey DS, Prostate Cancer Imaging Working Group. Challenges in clinical prostate cancer: Role of imaging. AJR Am J Roentgenol 2009;192:1455-70.
7. Epstein JI. An update of the Gleason grading system. J Urol 2010;183:433-40.
8. Hughes C, Rouvière O, Mege-Lechevallier F, Souchon R, Prost R. Robust alignment of prostate histology slices with quantified accuracy. IEEE Trans Biomed Eng 2013;60:281-91.
9. Gibson E, Crukley C, Gaed M, Gómez JA, Moussa M, Chin JL, *et al*. Registration of prostate histology images to *ex vivo* MR images via strand-shaped fiducials. J Magn Reson Imaging 2012;36:1402-12.
10. Ward AD, Crukley C, McKenzie CA, Montreuil J, Gibson E, Romagnoli C, *et al*. Prostate: Registration of digital histopathologic images to *in vivo* MR images acquired by using endorectal receive coil. Radiology 2012;263:856-64.
11. Chappelow J, Bloch BN, Rofsky N, Genega E, Lenkinski R, DeWolf W, *et al*. Elastic registration of multimodal prostate MRI and histology via multiattribute combined mutual information. Med Phys 2011;38:2005-18.
12. Groenendaal G, Moman MR, Korporaal JG, van Diest PJ, van Vulpen M, Philippens ME, *et al*. Validation of functional imaging with pathology for tumor delineation in the prostate. Radiother Oncol 2010;94:145-50.
13. Castaneda B, Hoyt K, Zhang M, Pasternack D, Baxter L, Nigwekar P, *et al*. P1C-9 prostate cancer detection based on three dimensional sonoelastography. Ultrasonics Symposium. Piscataway: IEEE; 2007. p. 1353-6.
14. Taylor LS, Porter BC, Nadasdy G, di Sant'Agnese PA, Pasternack D, Wu Z, *et al*. Three-dimensional registration of prostate images from histology and ultrasound. Ultrasound Med Biol 2004;30:161-8.
15. Zhan Y, Ou Y, Feldman M, Tomaszeweski J, Davatzikos C, Shen D. Registering histologic and MR images of prostate for image-based cancer detection. Acad Radiol 2007;14:1367-81.
16. Park H, Piert MR, Khan A, Shah R, Hussain H, Siddiqui J, *et al*. Registration methodology for histological sections and *in vivo* imaging of human prostate. Acad Radiol 2008;15:1027-39.
17. Shah V, Pohida T, Turkbey B, Mani H, Merino M, Pinto PA, *et al*. A method for correlating *in vivo* prostate magnetic resonance imaging and histopathology using individualized magnetic resonance-based molds. Rev Sci Instrum 2009;80:104301.
18. Jackson AS, Reinsberg SA, Sohaib SA, Charles-Edwards EM, Jhavar S, Christmas TJ, *et al*. Dynamic contrast-enhanced MRI for prostate cancer localization. Br J Radiol 2009;82:148-56.
19. Chen LH, Ho H, Lazaro R, Thng CH, Yuen J, Ng WS, *et al*. Optimum slicing of radical prostatectomy specimens for correlation between histopathology and medical images. Int J Comput Assist Radiol Surg 2010;5:471-87.
20. Breen MS, Lancaster TL, Lazebnik RS, Nour SG, Lewin JS, Wilson DL. Three-dimensional method for comparing *in vivo* interventional MR images of thermally ablated tissue with tissue response. J Magn Reson Imaging 2003;18:90-102.
21. Gibson E, Fenster A, Ward AD. Registration accuracy: How good is good enough? A statistical power calculation incorporating image registration uncertainty. In: Ayache N, Delingette H, Golland P, Mori K, editors. Medical Image Computing and Computer Assisted Intervention. Vol. 7511. Berlin Heidelberg: Springer Verlag; 2012. p. 643-50.
22. Gibson E, Gomez JA, Moussa M, Crukley C, Bauman G, Fenster A, *et al*. 3D reconstruction of prostate histology based on quantified tissue cutting and deformation parameters. SPIE Med Imaging 2012;8317:83170N.
23. Fitzpatrick JM, West JB, Maurer CR Jr. Predicting error in rigid-body point-based registration. IEEE Trans Med Imaging 1998;17:694-702.
24. Bookstein FL. Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Trans Pattern Anal Mach Intell 1989;11:567-85.
25. Fitzpatrick JM, West JB. The distribution of target registration error in rigid-body point-based registration. IEEE Trans Med Imaging 2001;20:917-27.
26. Garcia-Parra R, Wood D, Shah RB, Siddiqui J, Hussain H, Park H, *et al*. Investigation on tumor hypoxia in resectable primary prostate cancer as demonstrated by 18F-FAZA PET/CT utilizing multimodality fusion techniques. Eur J Nucl Med Mol Imaging 2011;38:1816-23.
27. Mena E, Turkbey B, Mani H, Adler S, Valera VA, Bernardo M, *et al*. 11C-Acetate PET/CT in localized prostate cancer: A study with MRI and histopathologic correlation. J Nucl Med 2012;53:538-45.
28. Piert M, Park H, Khan A, Siddiqui J, Hussain H, Chenevert T, *et al*. Detection of aggressive primary prostate cancer with 11C-choline PET/CT using multimodality fusion techniques. J Nucl Med 2009;50:1585-93.
29. Groenendaal G, Borren A, Moman MR, Monninkhof E, van Diest PJ, Philippens ME, *et al*. Pathologic validation of a model based on diffusion-weighted imaging and dynamic contrast-enhanced magnetic resonance imaging for tumor delineation in the prostate peripheral zone. Int J Radiat Oncol Biol Phys 2012;82:E537-44.
30. Orczyk C, Mikheev A, Rosenkrantz A, Melamed J, Taneja SS, Rusinek H. Imaging of prostate cancer: A platform for 3D co-registration of *in-vivo* MRI *ex-vivo* MRI and pathology. SPIE Med Imaging 2012;8316:83162M.
31. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA 1994;271:368-74.
32. Yaegashi H, Takahashi T, Kawasaki M. Microcomputer-aided reconstruction: A system designed for the study of 3-D microstructure in histology and histopathology. J Microsc 1987;146:55-65.