

Methodology to study the three-dimensional spatial distribution of prostate cancer and their dependence on clinical parameters

Kristians Diaz Rojas,^a Maria L. Montero,^b Jorge Yao,^c Edward Messing,^d Anees Fazili,^d Jean Joseph,^d Yangming Ou,^e Deborah J. Rubens,^f Kevin J. Parker,^{g,*} Christos Davatzikos,^h and Benjamin Castaneda^a

^aPontificia Universidad Católica del Perú, Department of Engineering, Section in Electrical and Electronic, Laboratory Medical Images, Av. Universitaria 1801, San Miguel Lima 32, Perú

^bPontificia Universidad Católica del Perú, Department of Science, Section of Mathematics, Laboratory Statistics, Av. Universitaria 1801, San Miguel Lima 32, Perú

^cUniversity of Rochester Medical Center, Department of Pathology and Laboratory Medicine, 601 Elmwood Avenue, Box 648, Rochester, New York 14642, United States

^dUniversity of Rochester Medical Center, Department of Urology, 601 Elmwood Avenue, Box 648, Rochester, New York 14642, United States

^eAthinoula A. Martinos Center for Biomedical Imaging, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts 02129, United States

^fUniversity of Rochester Medical Center, Department of Imaging Sciences, 601 Elmwood Avenue, Box 648, Rochester, New York 14642, United States

^gUniversity of Rochester, Department of Electrical and Computer Engineering, Hopeman Engineering Building 203, Box 270126, Rochester, New York 14627, United States

^hUniversity of Pennsylvania, Departments of Radiology and Electrical and Computer Engineering, 3600 Market Street, Suite 380, Philadelphia, Pennsylvania 19104, United States

Abstract. A methodology to study the relationship between clinical variables [e.g., prostate specific antigen (PSA) or Gleason score] and cancer spatial distribution is described. Three-dimensional (3-D) models of 216 glands are reconstructed from digital images of whole mount histopathological slices. The models are deformed into one prostate model selected as an atlas using a combination of rigid, affine, and B-spline deformable registration techniques. Spatial cancer distribution is assessed by counting the number of tumor occurrences among all glands in a given position of the 3-D registered atlas. Finally, a difference between proportions is used to compare different spatial distributions. As a proof of concept, we compare spatial distributions from patients with PSA greater and less than 5 ng/ml and from patients older and younger than 60 years. Results suggest that prostate cancer has a significant difference in the right zone of the prostate between populations with PSA greater and less than 5 ng/ml. Age does not have any impact in the spatial distribution of the disease. The proposed methodology can help to comprehend prostate cancer by understanding its spatial distribution and how it changes according to clinical parameters. Finally, this methodology can be easily adapted to other organs and pathologies. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.2.3.037502](https://doi.org/10.1117/1.JMI.2.3.037502)]

Keywords: prostate cancer; ultrasound; prostate specific antigen; image processing; registration; spatial distribution.

Paper 15042PRR received Mar. 2, 2015; accepted for publication Jun. 29, 2015; published online Jul. 29, 2015.

1 Introduction

Prostate diseases occur commonly in men after age 40. The prostate gland tends to increase in size with age, which can cause the urethra to become narrower and decrease urine flow, causing pain to patients and affecting their quality of life. Benign prostatic hyperplasia, prostatism, prostatitis, and prostatodynia are among the more common benign, noncancerous prostatic pathologies.

According to the American Cancer Society, prostate adenocarcinoma remains a disease that affects most men. The estimated incidence and mortality numbers during 2014 due to this disease in the USA were 233,000 (28% estimated new cancer cases) and 29,480¹ (10% estimated deaths), respectively. The incidence and mortality rates are similar in developing countries as well. In South America as a whole, prostate cancer

remains the most common cancer among men² (26.4% of new cancer cases and 14.6% of cancer deaths). Worldwide, the incidence and mortality numbers are ~899,000 and 258,000, respectively.³

Standard prostate cancer screening tests include a digital rectal examination (DRE) and prostate specific antigen (PSA) test. The DRE has long been used to diagnose prostate cancer due to its positive predictive value (from 33% to 83%) when the PSA value is high (3 to 9.9 ng/ml).⁴ However, few tumors are detected only with an abnormal DRE (11%) when the PSA value is low (0 to 2.9 ng/ml).⁴ Further, ~17% of men with a normal DRE are diagnosed with prostate cancer after a biopsy.

PSA is a glycoprotein produced naturally by prostate gland epithelial cells. High levels of PSA (>4 ng/ml) may be an indication of prostate infection, inflammation, enlargement, or cancer. PSA was adopted for cancer screening by the early

*Address all correspondence to: Kevin J. Parker, E-mail: kevin.parker@rochester.edu

1990s. However, this screening methodology has become somewhat controversial since randomized trials do not show a clear benefit. The European Randomized Study of Screening for Prostate Cancer (ERSPC) reported only a small benefit of PSA screening after 13 years of follow-ups.⁵ A report from the large United States trial, the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, published concurrently with the European trial, found no benefit for annual PSA and DRE screening after 13 years of follow-ups.⁶ Nevertheless, PSA combined with DRE is still used to identify prostate cancer in early stages.

After the initial screening, the ultimate diagnosis of prostate cancer is currently made by biopsy-guided transrectal ultrasound (TRUS) imaging. However, TRUS images exhibit low signal-to-noise ratio and low contrast. Therefore, biopsies are obtained in a random fashion with the ultrasound being used to localize areas of the prostate for sampling, rather than having the ability to target suspicious lesions directly. Flanigan et al.⁷ demonstrated that biopsies have a success rate of ~30% using the gold standard, six-core biopsy protocol. Subsequent studies indicate that to increase the level of success, more cores should be tested, and the standard of care today includes at least a 12-core sample.

In an effort to improve this situation, several researchers have explored the possibility of building a probability map containing the tumor distribution within the prostate gland based on histopathological images. This distribution map can be used to study the efficiency of existing biopsy schemes or to design optimal biopsy strategies.^{8–14} An estimation of the tumor distribution has a significant impact on both clinical and academic perspectives, allowing a better understanding of the disease.

Previous studies have focused on optimizing the positioning of biopsy cores based solely on the cancer distribution without considering additional relevant information such as cancer volume and Gleason score, two of the most important surrogate markers to determine the staging of the disease. In addition, the PSA level and location of the tumor in the gland can be included to improve clinical decisions (where to perform the biopsy). Studying the influence of variables such as PSA, Gleason score, age, and capsule penetration can lead to a better understanding of the disease and its treatment.

In this work, a software tool to study the correlation between these surrogate markers and the cancer distribution in the gland is presented. We illustrate the image processing methodology to develop several probabilistic maps (spatial distribution) of the

tumor distribution from a database of whole-mount histology sections. We also provide statistical analysis to compare two spatial distributions based on differences of proportions.

The manuscript is organized as follows: in Sec. 2 we present the methodology to develop the spatial distribution. Section 3 describes the statistical method to differentiate spatial distribution. Results are presented in Sec. 4, and discussion and conclusions are presented in Sec. 5.

2 Methodology for Three-Dimensional Reconstruction

Health Insurance Portability and Accountability Act recommendations were followed to guarantee the security and privacy of patient information.

Information from two different clinical sites was used (two databases): 58 patients from the University of Rochester Medical Center (URMC), Rochester, New York, United States, and 158 patients from the Section of Biomedical Image Analysis (SBIA), University of Pennsylvania, Philadelphia, Pennsylvania, United States.

All the patients were scheduled to undergo a radical prostatectomy. Their age range was 44 to 73 years old, with an average age of 60.51 ± 6.28 years. Their PSA values at the time of surgery ranged from 0.7 to 138 ng/ml, with a mean PSA value of 9.93 ± 14.13 ng/ml. Table 1 provides pathological variables for both databases (URMC and SBIA).

2.1 University of Rochester Medical Center Database

In the following sections, we will explain the framework followed for registering the URMC database.

2.1.1 Image acquisition

Each gland was weighed and measured. The gland was inked and a device,¹⁵ which consists of two sets of four 3-mm diameter mating metal prongs, was used to create markers for three-dimensional (3-D) reconstruction. Each set was attached to a metal base in a square grid of 13×13 mm². The prostate was inserted through the apex and the base symmetrically around the urethra. The specimen then was fixed in 4% formalin for 24 h.

After fixation, the device was removed and the gland was measured again to assess the shrinkage factor. The prostate was then sectioned into 4 mm slices from the apex to the

Table 1 Clinical or pathological variables of the University of Rochester Medical Center (URMC) and Section of Biomedical Image Analysis (SBIA) databases.

Pathological variable	URMC		SBIA		Combined databases (all patients)	
	Mean	Range	Mean	Range	Mean	Range
Age (years)	59.8 ± 5.4	48.5–72.8	60.6 ± 6.2	44.0–73.0	60.5 ± 6.2	44.0–73.0
PSA at time of surgery (ng/ml)	6.9 ± 8.2	1.0–61.1	11.2 ± 18.6	0.7–138.0	9.9 ± 14.1	0.7–138.0
Gleason score	6.8 ± 0.7	6.0–9.0	5.4 ± 1.0	3.0–8.0	6.2 ± 1.1	3.0–9.0
Cancer percentage	20.7	5.0–60.0	8.6	0.3–68.0	12.3	0.3–68.0

Note: PSA, prostate specific antigen.

base. Depending on the size of the individual prostate, 6 to 12 gross pathology whole-mount slices (3 to 5 mm thick), were obtained. The processed whole-section slices were embedded in Paraplast™, using large metal base molds (Surgipath Medical). Sections 4 to 5 μm thick were cut from the surface of each block. The slides were baked for 1 h at 60°C and stained using routine hematoxylin–eosin protocol. The result was one histology slide for each gross slab. The microscope whole-mount sections were examined by an expert pathologist in order to delineate the tumor in every slice.

Both gross pathology slices and histology slices were photographed after preparation, so that 3-D reconstruction could be done. The color images for both gross pathology and histological slices were captured using a digital camera (Diagnostic Spot RT digital color camera, Diagnostic Instruments Inc., Sterling Heights, Michigan, United States). The sheets were mounted on a flat table, parallel to the focal plane of the camera. The distance between the camera and the sheets was about 30 cm. The images were stored on a computer using the Advanced Spot camera software (Diagnostic Instruments). For the gross specimens, a caliper was used to measure the thickness of each slice. This measurement was used for 3-D reconstruction. Additional information on the acquisition protocol is given in the article by Taylor et al.¹⁵

2.1.2 Three-dimensional reconstruction

The methodology used to create the 3-D spatial distribution was as follows: (1) alignment of the histological and pathological image of the same prostate; (2) segmentation of the gland and the tumor in order to determine the outlines; and (3) interpolation to determine an isotropic volume. For interpolation, we used the distance between the apex and the base of the fresh gland when it was extracted; that information can be found on the patient info sheet. Figure 1 provides a diagram of the methodology that was used to build the 3-D model.

Alignment. Since pathological and histological images are positioned differently during image acquisition, it is necessary to have an alignment step. First, all histological and pathological slices are registered to their respective central slices (fixed images) since they are more likely to show four holes caused by the marking device. Each hole is labeled clockwise with

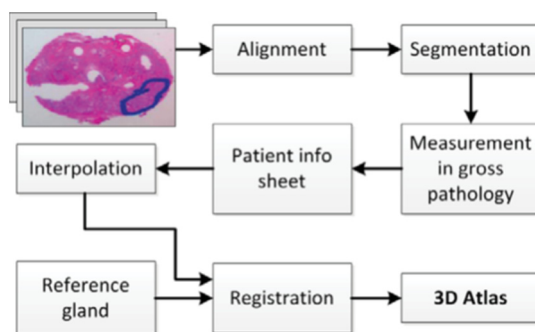


Fig. 1 Methodology to build the three-dimensional (3-D) spatial distribution. First, alignment the histological and pathological image of the same prostate; then segmentation the gland and the tumor; followed by interpolation to have an isotropic volume using the patient information sheet in order to get the distance between slices and finally the registration step.

a number [see Fig. 2(a)] in all slices. This step is performed manually.

The device used in the acquisition protocol provides landmarks, which are 13 mm apart in a square configuration. This information is helpful to correct any deformations which the tissue underwent as part of the pathological processing. Figure 2(a) shows the four holes from the landmark device [Fig. 2(b)] in a histology slice. Figure 2(c) shows the next slice from the same prostate. Note that the holes are not positioned in the same place. Finally, Fig. 2(d) shows the result after applying the alignment step to the second image.

For the alignment step, we used a warping algorithm,¹⁶ a pair of corresponding lines in the source, and destination images to define a coordinate mapping from the pixel X in the destination image to the pixel X' in the source image. Let us consider that the landmarks 1 and 4 form the line PQ in the destination image; similarly the landmarks 1 and 4 in the source image form the line $P'Q'$, as we can see in Fig. 3.

We calculate the distance of a pixel X to the line PQ to obtain v and the distance from P to the projection of X onto the line PQ to obtain u . With these values, we determine the location of X' in the source image, maintaining the proportions. This procedure is repeated for all image points.

$$u = \frac{\langle (X - P), (Q - P) \rangle}{\|Q - P\|^2}, \quad (1)$$

$$v = \frac{(X - P) \perp (Q - P)}{\|Q - P\|}, \quad (2)$$

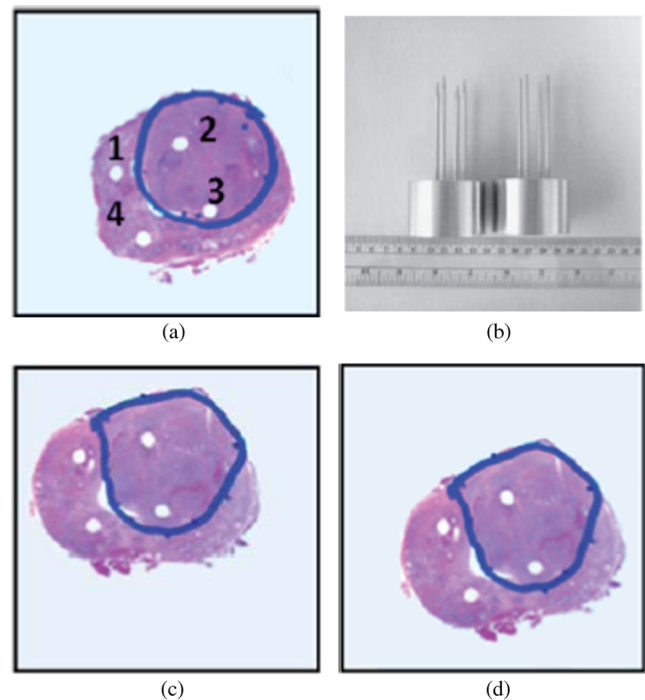


Fig. 2 (a) Label of the four holes in the central histology slice. (b) Device used in the image acquisition protocol. (c) Next, histology slice of the same prostate gland not aligned with the base image. (d) Result of image (c) aligned with the base image (a). The four holes in the prostate gland are used as landmarks in the alignment step.

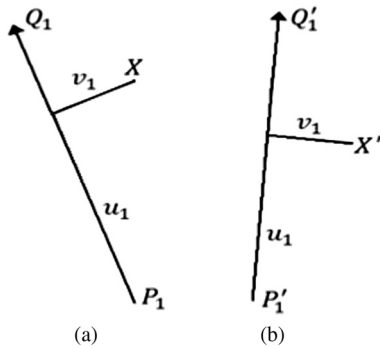


Fig. 3 (a) Positions of a single line pair in the destination image and (b) in the current source image.

$$X' = P' + u(Q' - P') + \frac{\langle v, \perp (Q' - P') \rangle}{\|Q' - P'\|}, \quad (3)$$

where $\perp$ represents the perpendicular vector operator and $\langle \cdot \rangle$ represents the dot products. For details, see Ref. 16.

Due to cutting and tearing in the acquisition protocol, some histological images will have missing parts [see Fig. 4(a)]. In order to determine a possible border, we propose to use information from the corresponding gross pathology section. Following the same procedure to align the set of histology images, the holes in the corresponding pathology images are labeled and the warping transform is applied.

Pathology images are used when there is missing information in the boundary of the histological image (see segmentation section). In all cases, the central slice of the prostate is selected as the fixed image and the rest of the cuts are aligned to it.

Histological cuts of the base and apex do not always show all four holes. Since at least two holes are needed for the warping transformation, a translation transform is applied when only one hole is present.

Segmentation. The purpose of the segmentation was to identify the outline of the prostate and the shape of the tumor previously delimited by an expert pathologist. This segmentation step is required to generate a 3-D reconstruction of the prostate and the tumor.

Since the images were not acquired in a controlled environment, there were several factors that made this step difficult (lighting and unwanted shadows in images). Therefore, a

preprocessing step was required. The histological images have a characteristic red to pink color with a white background. Thus, the intensity levels corresponding to the image are located on the first mode of its histogram. The other mode corresponds to the background (highest intensity values—nearly white). The histogram was formed as the sum of each histogram layer. We detected the peak of the first mount to establish an upper threshold, whereas the lower threshold was established at the 10% intensity of the peak. Finally, an image-stretching transformation was applied with these two thresholds.

After applying the preprocessing step to the histological images, they showed high contrast between the prostate and the background. The outline of the prostate gland was located in the red and blue layers.

The tumor was delineated by an expert pathologist on the histology cut using a blue or black marker. To segment it from the image, an automated thresholding based on Otsu's algorithm¹⁷ was performed on the red layer. After this step, we only considered the objects which were within the previously segmented prostate.

Some histological images had missing parts [see Fig. 4(a)]. We used information from its pathological counterpart to correct the segmented boundary. To do this, we developed a semiautomatic algorithm. First, the pathological and histological images were aligned as explained in the alignment section [Fig. 4(c)]. Then the border of the segmented prostate was decomposed into a set of vertices which are overlaid on the pathology image. A few vertices are dragged on the desired boundary (pathology image) by the user.

Then discrete dynamic contour¹⁸ was applied to move the rest of the outline closer to the correct boundary. The vertices inserted by the user (two or three) are clamped; this ensures that adjacent vertices generally deform toward the clamped vertices. This semiautomatic algorithm allows the user to edit and guide the segmentation process. In our experiments, the weights were $w_i^{\text{img}} = 0.4$ and $w_i^{\text{int}} = 0.6$. The image force is responsible for moving the vertex to the nearest and strongest edge as long as it is within the influence area, which is determined by a two-dimensional (2-D) Gaussian function. Internal forces are computed based on neighboring vertices and constrain the vertex to form a smooth contour. The damping force provides stability in the iteration process.

In addition to this problem, in some images, the pathologist did not completely delineate the tumor with a continuous solid line but with spaced dots. To obtain the complete outline, cubic interpolation was applied.

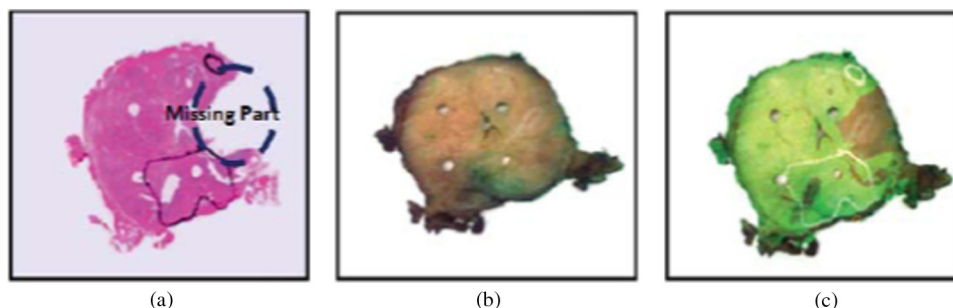


Fig. 4 (a) The dashed circle in the histology slide shows the missing boundary; the thin, black outline is the localization of the cancer (defined by an expert pathologist). (b) The counterpart gross pathology image of image (a). (c) Gross image aligned to the histology slide, in green the overlaid histology slide with cancers outlined and registered.

Measurement in gross pathology. To build the 3-D reconstruction of the prostate gland, it is necessary to quantify the separation between histology and gross pathology slices. The thickness of each gross pathology slab was measured using a caliper. Eight measurements were performed, two measurements per hole (one using the outside jaws and the other using the depth probe of the caliper). This task was repeated by two observers and the average measurement is used. The result was the measured distance between adjacent histology cuts.

Handling during the acquisition process caused the gland to shrink or deform. Therefore, the distance between the apex and the base of the gland was measured (with a caliper) when the prostate was removed. The distance before deformation was available on the patient information sheet and it was compared with the distance obtained postdeformation. Thickness values were then scaled to make the sum of the thicknesses of all slices equal to the apex-to-base distance measured in the fresh prostate. Next, an interpolation step was required in order to have a coherent 3-D reconstruction.

Interpolation. The distance between adjacent image elements (pixels) within a slice is smaller than the distance between adjacent image elements in two neighboring slices. Histology images had a thickness of 5 μm but were obtained at spacings of 4 mm. Typically, there were eight histology slices per gland.

There are several approaches for reconstructing and displaying 3-D objects from serial cross sections.^{19,20} The major difference in these approaches lies in the interpolation method employed. The two main methods of interpolation techniques for reconstructing objects are gray-level and shape-based interpolations. Gray-level methods employ nearest neighbor, linear, or polynomial and splines interpolation. On the other hand, shape-based interpolation methods consider shape features extracted from the image. Shape-based interpolation converts binary images into distance maps using distance transformation functions such as the Manhattan or city-block distance template^{21,22} to approximate the Euclidean distance between the pixel and the contour of the object. Shape-based interpolation is quite simple to understand, easy to implement, and computationally fast. However, this method fails to interpolate slices when there is no overlapping area between the two objects.

We used the method proposed by Lee and Wang²³ to overcome this limitation. A morphology-based interpolation method uses dilation and erosion morphological operators to create distance maps and perform the interpolation. After interpolating the prostate, the same procedure is repeated for each tumor present in the gland. However, there are additional considerations to be taken in order to avoid ambiguities in the reconstruction. For example, Figs. 5(a)–5(c) show the location of a tumor in three slices and Figs. 5(d) and 5(e) show their possible 3-D reconstructions. As we can see, there is more than one option, since the tumor can grow in any direction and sense.

First, we determined tumor-matching between slices (since multiple cancers may exist on two input slices) by measuring the distance between the contours of the tumor. If the result was less than 5 mm, they were considered the same cancer. This is a heuristic based on the experience of the expert pathologist.^{24,25}

If a slice contains a small region of cancer with no corresponding regions in an adjacent slice—for example, the upper lesion in Fig. 5(b), compared to Fig. 5(c)—then the

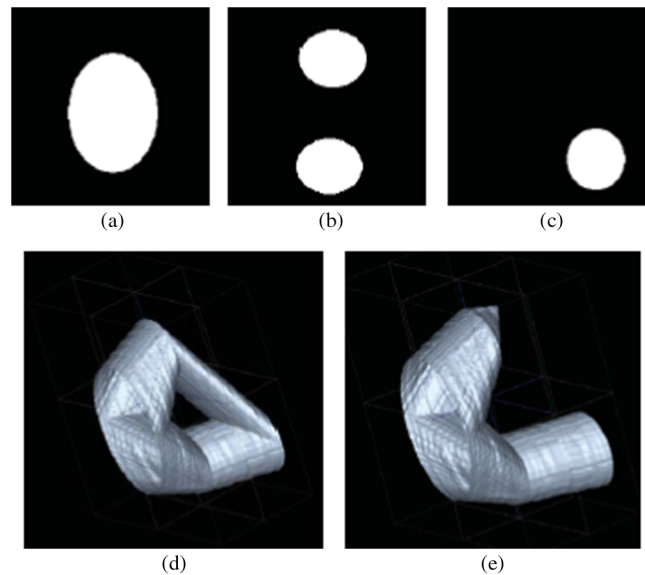


Fig. 5 (a), (b), and (c) show the location of a tumor in three consecutive slices. If there is not a decision rule, there will be two possible 3-D reconstructions: (d) and (e).

confirmed cancer is extended into 3-D using morphological operators. However, the interpolated cancer volume is constrained by the adjacent cancer-negative slice.

Registration. A key challenge of this study is the registration of 3-D surface models. All of the prostate models studied need to be reshaped such that tumors are placed in a uniform prostate atlas. Frimmel et al.¹⁴ divided all prostate models into three groups by size (small, medium, and large). Prostate models were reshaped into one of these three models using a 2-D reshaping algorithm. Shen et al.^{10,11,13} presented the use of a deformation-based registration approach using an adaptive focus deformable model (AFDM)^{26,27} to spatially normalize the external and internal structures of the prostate samples.

In this work, the registration process was composed of four steps: (1) a rigid registration was used to align the volumes, (2) an affine transformation allowed compensation for shear and scale, (3) a deformable B-spline registration²⁸ with a coarse grid was applied, and (4) another B-spline registration with a finer grid. Rigid and affine transformations bring the registration process close to its global minimum, and B-spline transformations were used to compensate for deformations in the gland due to mechanical and chemical procedures in the histological processing. This approach was developed by Castaneda et al.²⁹

Since our database consists of whole prostatectomy glands subsequent to cancer diagnosis, we do not have a healthy prostate to use as a model for the development of the atlas. Therefore, the model used for SBIA (Sec. 2.2) was selected to be the atlas model and the 3-D reconstructions from all other glands were warped against it.

2.2 Section of Biomedical Image Analysis Database

The SBIA database was built using the method proposed by Shen and Davatzikos²⁶ and reported in previous publications.^{27,30} The spatial normalization is done using a 3-D surface-based, nonrigid image registration, as demonstrated in Zhan et al.³¹ The spatial normalization preserves the zonal

anatomy of the prostate so that after normalization, the same spatial coordinates will correspond to approximately the same anatomic locations. This framework is based on an AFDM. This deformable model includes information about the statistical variation of prostate structures within a given population.

The database from SBIA was combined with the database from UPMC to improve the statistical analysis and to compare two different databases.

3 Statistical Analysis

Once all of the prostate glands have been registered into a unique atlas model, the spatial distribution is developed by counting the number of occurrences in a given position in the 3-D space from each registered prostate reconstruction. Our clinicians (E. Messing and A. Fazili) recommended the comparative analysis of subgroups based on available data and clinical significance:

- PSA at the time of surgery less than, or greater than, or equal to 5 ng/ml.
- Age less than, or greater than, or equal to 60 years old.
- Gleason score greater than, or less than, or equal to 6.

A particular way to illustrate differences between subgroups is to apply the Z-test^{32,33} on pairs of voxels. This approach is capable of highlighting voxels where the cancer distributions have a meaningful difference, although it does not consider the joint probability of having a significant difference (SD) in a region (see Sec. 5 for detailed explanation).

The proportions are defined as the number of occurrences divided by total number of patients that correspond to the population of study and analysis (e.g., 216 if we analyze the entire population, see Fig. 6).

The difference between proportions was applied to each voxel and was evaluated with a Z-test in order to ensure that

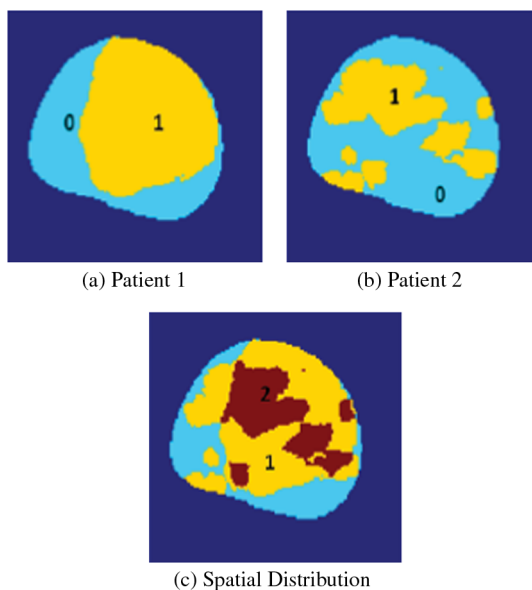


Fig. 6 (a and b) Spatial distribution results from two patients; for each patient, the colors yellow and cyan represent the zones with and without tumor, respectively. The values of 1 and 0 are assigned to the yellow and cyan zone, respectively. (c) To generate the spatial distribution, we count the number of occurrences after summing the two images. The color red corresponds to two occurrences.

the voxel was significantly different between the two spatial distributions. This process was repeated for all the voxels corresponding to the prostate model. We chose the critical Z value for two tails 1.65 (equivalent to $p = 0.1$) in order to find SDs while producing clearly defined regions within different subgroups, as shown in Sec. 4. The value used for PSA analysis is 5 ng/ml in order to obtain representative statistical analysis. There is no report in the literature about the exact cutoff values for PSA. Table 2 shows the number of patients by pathological variable with the threshold used in this research.

The values chosen as threshold in each parameter are a compromise between commonly accepted clinical thresholds separating high and low values, and the number of patients in each subgroup required to make a good statistical evaluation.

4 Results

The database consists of 215 prostate glands (58 patients from UPMC and 157 patients from SBIA). For each gland, histopathological images as well as pathological variables (PSA at time of the surgery, Gleason score, patient age, and percentage of cancer) were available. Figure 7 shows the cancer probability distribution of the complete database, going through the base to the apex (from top left to bottom right). We will use the same sequence, from base to apex, in later Figs. 8–12. The results show that the prostate cancer is more likely to occur in the middle right zone of the prostate (represented in red in Fig. 8). We used a pseudocolor map in order to identify the zones with more level of occurrences. The color scale indicates the number of occurrences for a given population.

Figure 8 shows the location of the largest number of occurrences corresponding to the right and left sides of the gland for the entire database (UPMC and SBIA). Table 3 details the maximum number of occurrences for each database.

Figure 9 shows the cancer probability distribution within the two studies (UPMC and SBIA). Both distributions have a higher number of occurrences in the anterior zone (reddish color). There is an SD in the base and apex of the prostate, but not in the middle zone (p -value > 0.1), as shown in Fig. 9(c), where a p -value < 0.1 is outlined in white.

Next, we combined the two data bases to create a larger ($N = 215$) group, which can be subdivided by different clinical parameters. Statistical analysis shows that for a population with PSA greater and less than 5 ng/ml, there is a spatial SD in the right zone of the prostate. Figure 10 shows the two spatial distributions and their SD.

Table 2 Group of patients by pathological variable used in this research.

		Number of patients
PSA at time of surgery	<5 ng/ml	134
	≥ 5 ng/ml	55
Age	<60 years old	77
	≥ 60 years old	117
Gleason score	> 6	49
	≤ 6	51

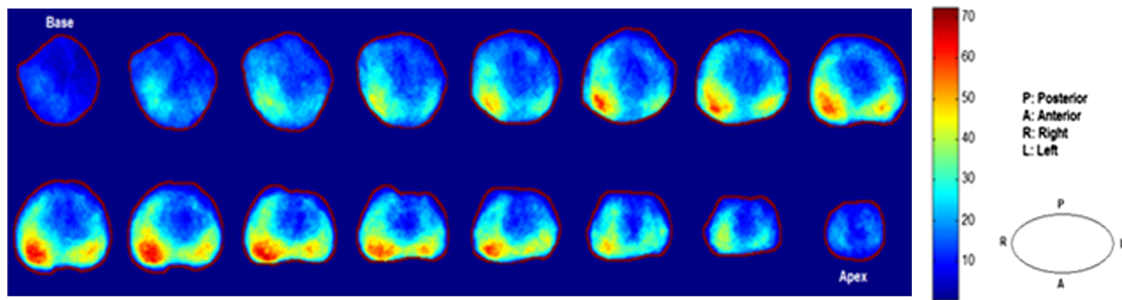


Fig. 7 Spatial distribution of all the patients (216 total). The color scale indicates the number of occurrences; the right zone shows to have the most level of occurrences. The labels—right, left, apex, and posterior—are related to the patient's position.

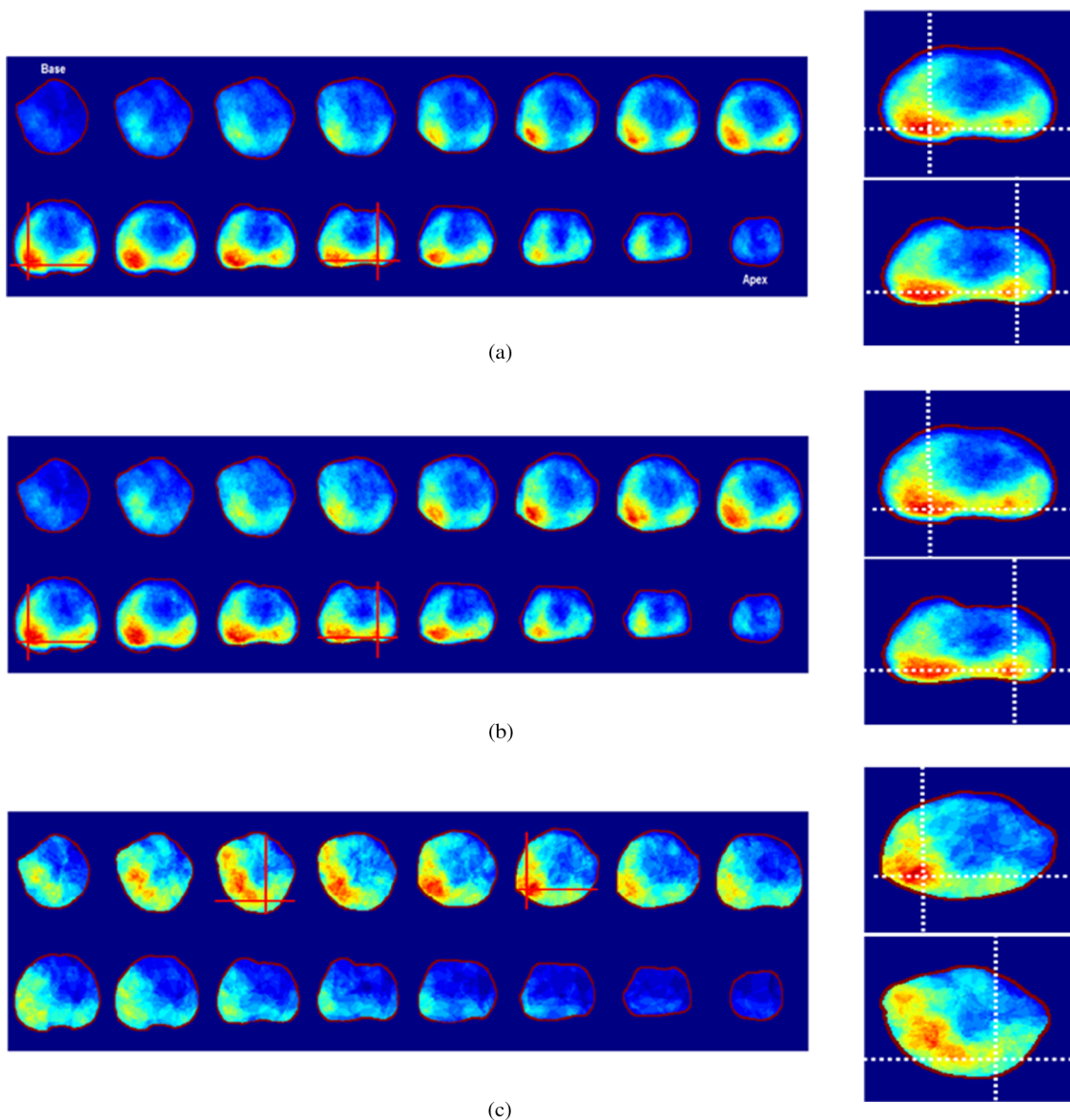


Fig. 8 The figures on the left show the 3-D location (red crosses) of the largest number of cancer occurrences for right and left sides of the prostate. The figures on the right show the corresponding slices with the largest number of occurrences. (a) Entire database, (b) Section of Biomedical Image Analysis (SBIA) database, and (c) University of Rochester Medical Center (URMC) database.

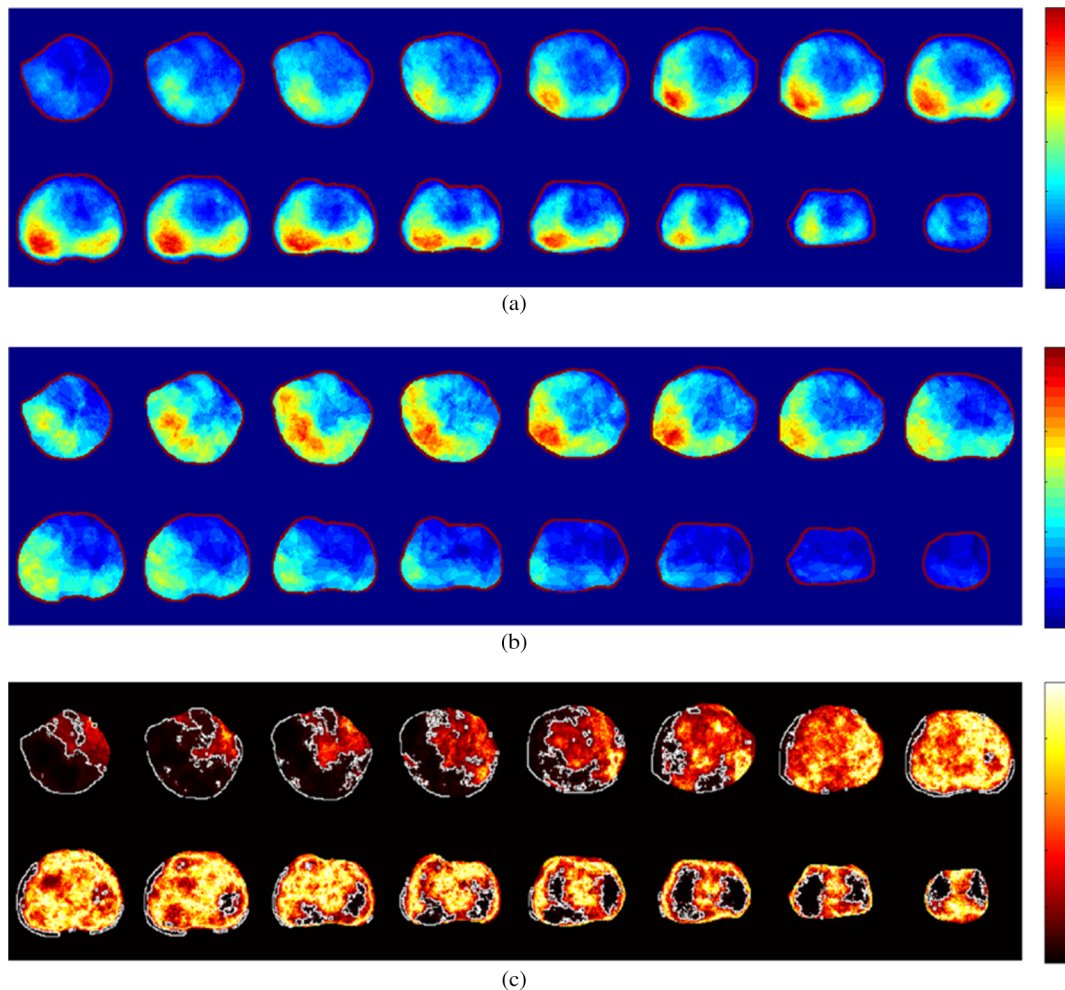


Fig. 9 Spatial distribution of (a) SBIA and (b) URM. (c) Spatial p -value; the area inside the white outline has a p -value <0.1 .

Similar results for the population with Gleason score greater and less or equal to 6 were obtained (Fig. 11).

In the case of populations older and younger than 60 years old, statistical analysis shows that age does not have any impact on the spatial distribution of the disease, see Fig. 12. In all cases, we used a Z-test for two independent proportions; the critical value is $Z = 1.65$. The red color indicates SD.

Table 4 quantifies the percentage of the whole prostate voxels that exceeded the Z-test threshold, for each comparison of two subgroups. Even though there are significant spatial differences of prostate cancer within the prostate gland when analyzed according to our clinical-pathological variables, the extent of these differences are highly variable, as seen in the table below.

5 Discussion and Conclusions

Probabilistic maps are used to investigate where cancer is more likely to develop within the prostate gland. The goal of this study was to develop a computational tool, which extends the use of 3-D probabilistic maps for a statistical analysis of prostate cancer between different populations in relation to the cancer's clinical parameters: PSA, Gleason score, patient age, and cancer percentage. This analysis provides spatial SDs between maps, which would increase our understanding of this disease in relation to its clinical parameters. This tool would help create a starting point for future work to comprehend

the relationship between tumor distribution and different surrogate markers. Therefore, this study could have a significant impact.

Several computer-assisted approaches^{9-13,34} developed in this field have focused on the optimal targeting of biopsy needles so as to maximize the detection of cancer but without considering other clinical parameters. Our approach can estimate the spatial distribution of the cancer with respect to clinically significant parameters, and this can help create a probabilistic map customized to the patient.

The database used in this study is composed of prostate glands obtained after radical prostatectomy for prostate cancer. Consequently, the database does not include normal prostates, and the probabilistic model has some deformation because of the presence of a tumor. An expert pathologist manually selected one prostate from the database with an average size and shape to be the atlas model. This process is similar to other reported studies^{10,11,13,30} and does not affect the statistical analysis used since the same prostate is used for the registration step.

Most of the statistical analyses are based on the hypothesis about two independent population medians (Mann-Whitney U test and Kolmogorov-Smirnov test for two independent samples), which perform a comprehensive analysis of the distributions indicating the similarity between populations. However, these analyses do not provide the spatial difference between

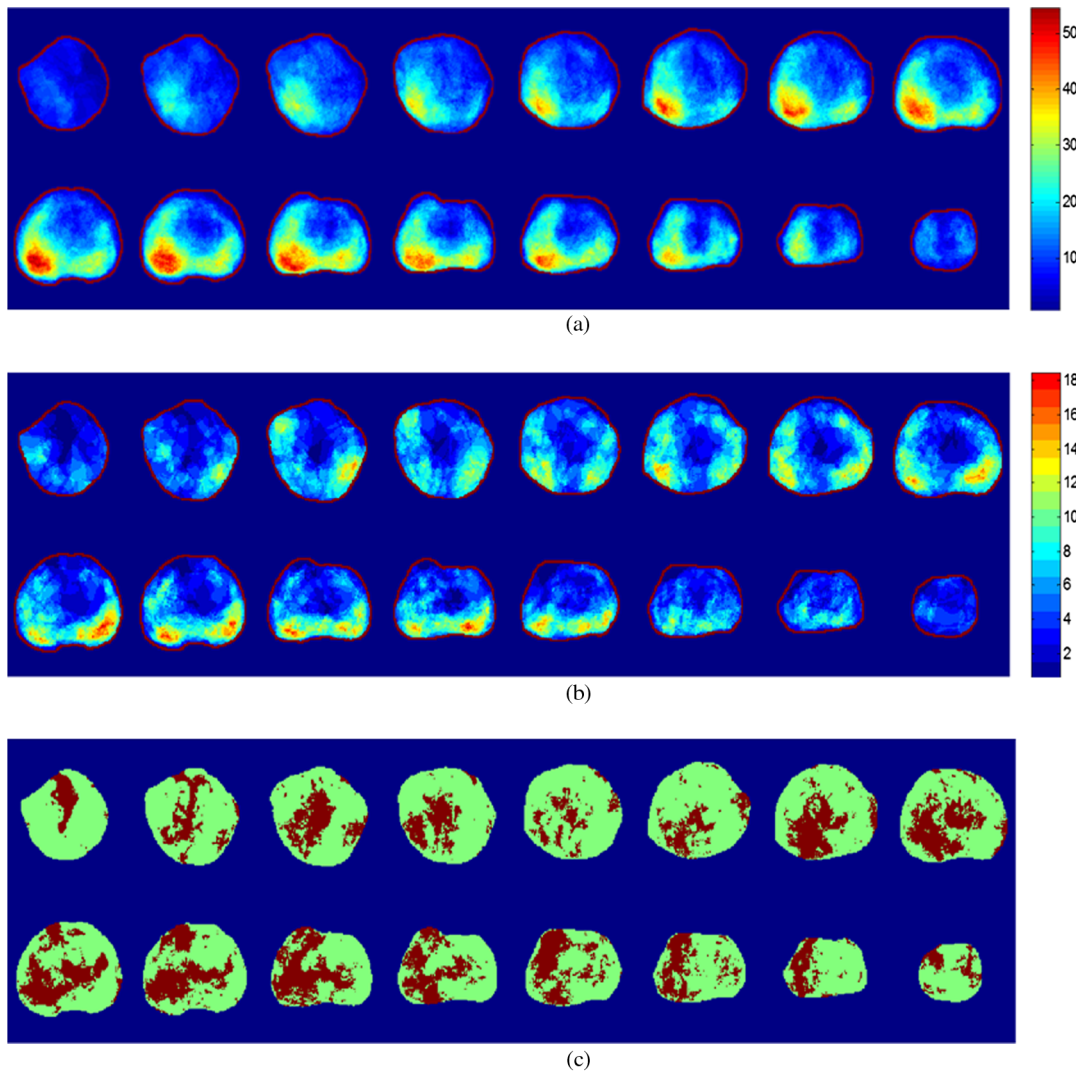


Fig. 10 Cancer probability distribution for prostate specific antigen (PSA) at the time of surgery. (a) Greater than 5 ng/ml/number of patients = 134/189. (b) PSA ≤ 5 ng/ml/number of patients = 55/189, where color scale indicates number of occurrences. (c) significant difference (SD) using Z-test for two independent proportions. Red color indicates SD.

two populations. The Z-test for two independent proportions was proposed in order to determine the most SD, voxel by voxel. We analyze differences in distribution between two sub-populations separated by a clinical or pathological variable. One limitation of this approach is that the Z-test as applied does not account for correlation or the conditional probability of adjacent voxels within an individual prostate. In our approach, we simply display the list of voxels, which individually have an SD between the independent groups being analyzed. If the joint probability of all the highlighted voxels is required, repeated use of the Z-test on adjacent pixel pairs will increase the overall probability of type 1 errors. Therefore, a difference test, such as the Holm–Bonferroni method, for multiple comparisons would be required for more rigorous statistical analysis.³⁵

Previous studies on 3-D spatial distribution of prostate cancer^{8–12,36} have identified some zonal differences in the occurrence of prostate cancer—specifically, a higher incidence in posterior (versus anterior) and mid-apex (versus base) based on biopsy samples.⁸ In our combined database, the region with the highest number of occurrences is the right zone, more

precisely in the right posterior peripheral zone near the base of the prostate. However, a biological explanation for this has not yet been given in the literature.

According to our results, there is little difference in the spatial distribution of cancer between patients greater or less than 60 years old. The effect of age on the distribution of prostate cancer may have been confounded by the larger differences seen when comparing groups based on PSA, Gleason score, and tumor volume, all of which are expected to be higher with advanced age.

In regards to PSA, the value of 5 ng/ml was used as a cutoff in our analysis to develop two spatial distributions to compare two populations and find SDs. According to our results, prostate cancer has an SD in the right zone of the prostate in patients with a PSA greater than and less than 5 ng/ml.

As seen in Table 3 as well as Figs. 7 and 8, the right zone of the prostate shows a greater number of occurrences of prostate cancer than the left, and this result was seen in both individual institutional databases as well as the combined database of all cases. Although there is no clear biological explanation for this

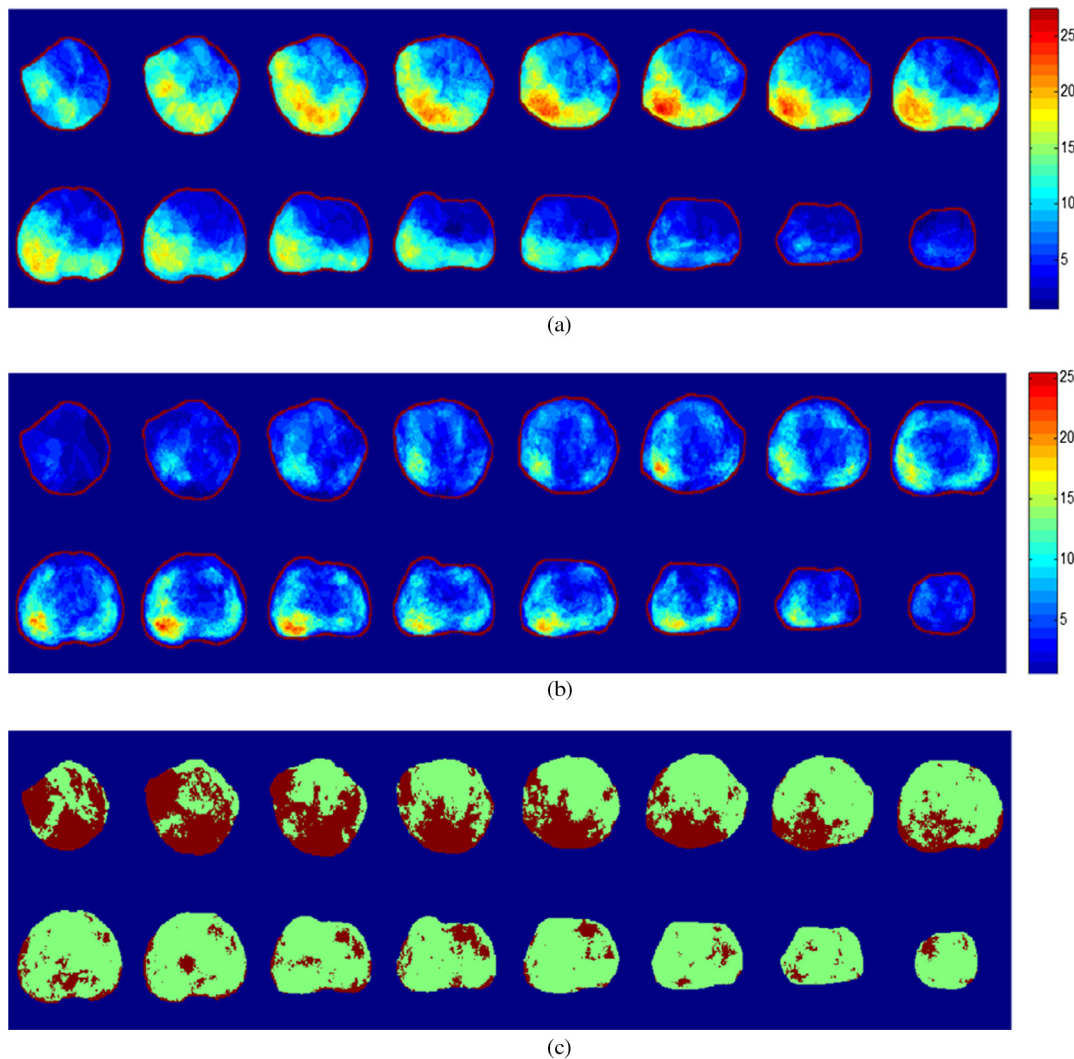


Fig. 11 Cancer probability distribution for Gleason score. (a) Greater than 6/number of patients = 49/100. (b) $\text{PSA} \leq 6$ /number of patients = 51/100, where color scale indicates number of occurrences. (c) SD using Z-test for two independent proportions. Red color indicates SD.

finding, a right to left asymmetry is not unique in urologic diseases and can be seen in other pathologies such as congenital ureter pelvic junction obstruction, which is more common on the left. Hence, the lack of any explanation for the asymmetry at the present time does not detract from the observed result. It also brings into question whether equal sampling of the prostate gland in each sextant, as is the current standard of care, should be altered to include a greater number of biopsies from the patient's right side, although this issue requires further study.

Ultimately, there is a clear disparity between the incidence and cancer-specific mortality of prostate cancer, as discussed previously. Given this discrepancy, there has been an increasing reliance on “active surveillance” of what is considered to be “low risk” prostate cancer in recent years. This entails the serial monitoring of PSA, DRE, and Gleason score on repeat biopsy after the initial diagnosis of prostate cancer has already been confirmed via an initial biopsy. There are various active surveillance criteria and protocols in existence, but the cornerstone of all of these protocols is repeat prostate biopsy, since an increase in tumor grade or volume is the most likely variable to move a patient from active surveillance of the disease to direct therapy for their cancer.^{37–39}

Hence, the importance of accurate biopsy for guiding subsequent treatment decisions becomes even more imperative, and it is in this context that a patient-specific biopsy template based on parameters such as PSA, Gleason score, and tumor volume may be of even more benefit. Accordingly, since all active surveillance protocols have an exclusion criteria for Gleason scores >6 , the increased prevalence of prostate cancer seen at the right base of the gland in our database, when the Gleason score was >6 (Fig. 11), may suggest that physicians should increase sampling in this zone of the prostate to help diagnose clinically significant prostate cancer.

It should be noted that there has been a move toward magnetic resonance imaging (MRI)-guided prostate biopsy in recent years, which signals a move away from random prostate biopsies—currently done via TRUS—and represents a move toward actual directed lesion targeting.⁴⁰

Despite much progress in the development of prostatic MRI, including MRI spectroscopy and more recently ultrasound–MRI fusion biopsy devices, the higher costs associated with this technique as well as its still-investigational status has meant that the standard of care for prostate cancer biopsy and diagnosis currently remains a TRUS-guided, 12-core prostate biopsy.

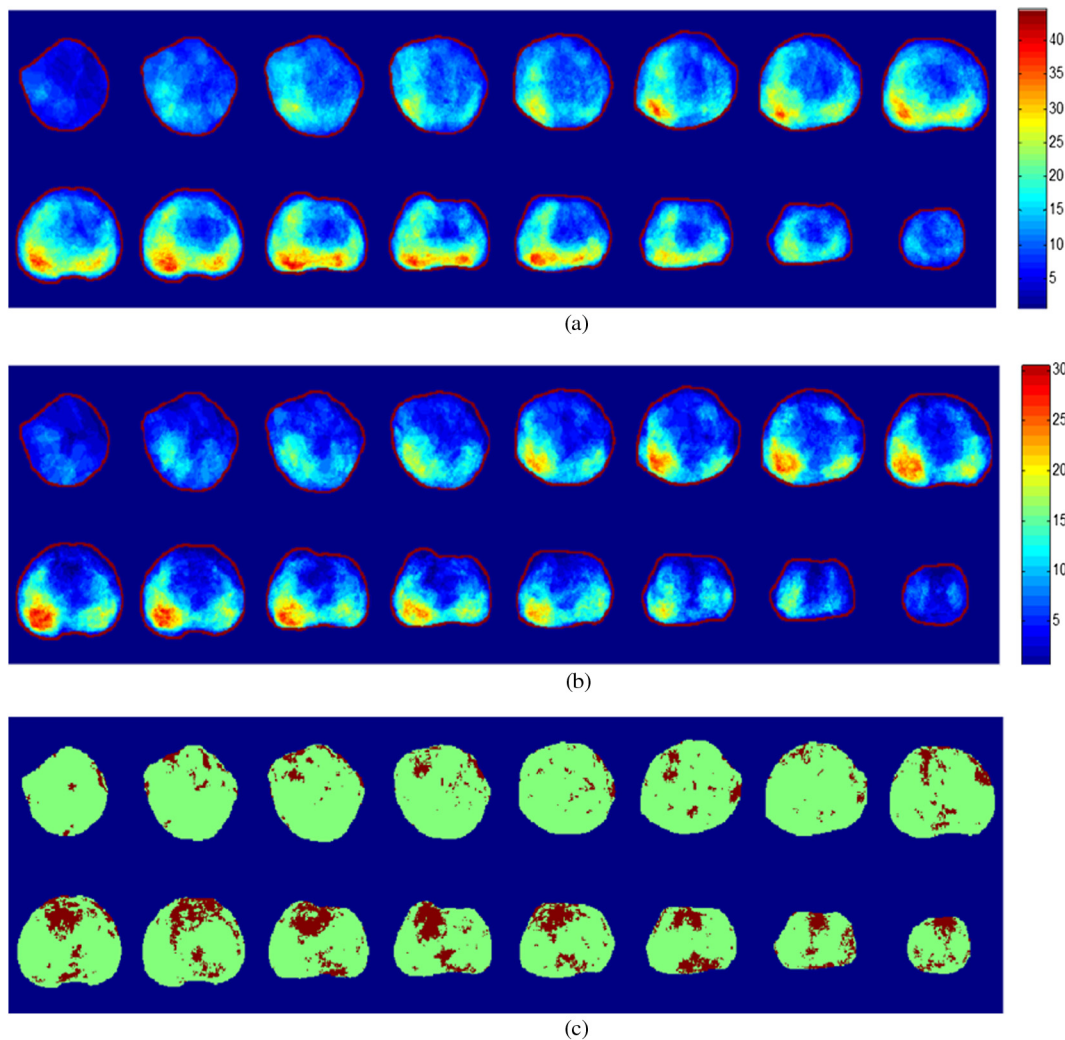


Fig. 12 Cancer probability distribution according to age. (a) Older than 60 years old/number of patients = 117/194. (b) Younger or equal to 60 years old/number of patients = 25/58, where color scale indicates number of occurrences. (c) SD using Z-test for two independent proportions. Red color indicates SD.

Table 3 Maximum number of occurrences found on left and right slides.

Databases	Number of occurrences	
URMC	19/58	28/58
SBIA	47/157	55/157
URMC + SBIA	51/215	70/215

Table 4 Significant difference (SD) in subgroups divided by clinical parameters.

Clinical parameter	SD (percent of voxels within the reference atlas prostate)
PSA at time of surgery: < and ≥ 5 ng/ml	28.77%
Age: over and less than 60 years old	12.27%
Gleason score: > and ≤ 6	29.81%

To summarize, a statistical atlas based on histologic images of whole-mount prostatectomy specimens has been created in this study. The methodology enables us to better understand the differences in the spatial occurrence of prostate cancer and test for patterns associated with clinical–pathological variables.

Acknowledgments

This work was supported by NIH Grants 5R01AG029804 and R01CA104976, and by DGI-PUCP Grant 2010-0208. The authors are grateful to Ollivier Hyrien, PhD, for his advice on biostatistics.

References

1. American Cancer Society, *Cancer Facts and Figures*, American Cancer Society, Atlanta, Georgia (2014).
2. O. G. de Planteamiento y Presupuesto, *Instituto nacional de enfermedades neoplasicas organismo publico ejecutor: Plan operativo institucional*, http://www.inen.sld.pe/portal/documentos/pdf/institucional/Plan_Estrategico/22112011_PEL_2011_2015.pdf (2011).
3. J. Ferlay et al., “Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008,” *Int. J. Cancer* **127**, 2893–2917 (2010).

4. F. H. Schröder et al., "Evaluation of the digital rectal examination as a screening test for prostate cancer. Rotterdam section of the European randomized study of screening for prostate cancer," *J. Natl. Cancer Inst.* **90**(23), 1817–1823 (1998).
5. European Randomized Study of Screening for Prostate Cancer, *Greater Absolute Benefit from PSA Screening at 13 years of Follow-Up in the ERSPC Trial Not Sufficient to Justify Population-Based Screening*, European Randomized Study of Screening for Prostate Cancer (2014).
6. G. L. Andriole et al., "Prostate cancer screening in the randomized prostate, lung, colorectal, and ovarian cancer screening trial: mortality results after 13 years of follow-up," *J. Natl. Cancer Inst.* **104**(2), 125–132 (2012).
7. R. C. Flanigan et al., "Accuracy of digital rectal examination and trans-rectal ultrasonography in localizing prostate cancer," *J. Urol.* **152**(5 Pt 1), 1506–1509 (1994).
8. J. Zeng, J. J. Bauer, and S. K. Mun, "Modeling and mapping of prostate cancer," *Comput. Graph.* **24**(5), 683–694 (2000).
9. J. Zeng et al., "Distribution of prostate cancer for optimized biopsy protocols," in *Medical Image Computing and Computer-Assisted Intervention*, S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds., pp. 287–296, Springer-Verlag, Berlin Heidelberg (2000).
10. D. Shen et al., "A statistical atlas of prostate cancer for optimal biopsy," in *Medical Image Computing and Computer-Assisted Intervention*, W. Niessen and M. Viergever, Eds., pp. 414–424, Springer-Verlag, Berlin Heidelberg (2001).
11. D. Shen et al., "Statistically optimized biopsy strategy for the diagnosis of prostate cancer," in *Proc. of 14th IEEE Symposium on Computer-Based Medical Systems*, pp. 433–438 (2001).
12. M. B. Opell et al., "Investigating the distribution of prostate cancer using three-dimensional computer simulation," *Prostate Cancer Prostatic Dis* **5**(3), 204–208 (2002).
13. D. Shen et al., "Optimized prostate biopsy via a statistical atlas of cancer spatial distribution," *Med. Image Anal.* **8**(2), 139–150 (2004).
14. H. Frimmel et al., "Modeling prostate cancer distributions," *Urology* **54**(6), 1028–1034 (1999).
15. L. S. Taylor et al., "Three-dimensional registration of prostate images from histology and ultrasound," *Ultrasound Med. Biol.* **30**(2), 161–168 (2004).
16. T. Beier and S. Neely, "Feature-based image metamorphosis," *SIGGRAPH Comput. Graph.* **26**(2), 35–42 (1992).
17. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
18. S. Lobregt and M. A. Viergever, "A discrete dynamic contour model," *IEEE Trans. Med. Imaging* **14**(1), 12–24 (1995).
19. S. Kels and N. Dyn, "Reconstruction of 3D objects from 2D cross-sections with the 4-point subdivision scheme adapted to sets," *Comput. Graph.* **35**(3), 741–746 (2011).
20. A. B. Albu, T. Beugeling, and D. Laurendeau, "A morphology-based approach for interslice interpolation of anatomical slices from volumetric images," *IEEE Trans. Biomed. Eng.* **55**(8), 2022–2038 (2008).
21. S. P. Raya and J. K. Udupa, "Shape-based interpolation of multidimensional objects," *IEEE Trans. Med. Imaging* **9**(1), 32–42 (1990).
22. G. T. Herman, J. Zheng, and C. A. Bucholtz, "Shape-based interpolation," *IEEE Comput. Graphics Appl.* **12**(3), 69–79 (1992).
23. T.-Y. Lee and W.-H. Wang, "Morphology-based three-dimensional interpolation," *IEEE Trans. Med. Imaging* **19**(7), 711–721 (2000).
24. J. E. McNeal et al., "Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread," *Am. J. Surg. Pathol.* **12**(12), 897–906 (1988).
25. A. Erbersdobler et al., "Pathological and clinical characteristics of large prostate cancers predominantly located in the transition zone," *Prostate Cancer Prostatic Dis.* **5**(4), 279–284 (2002).
26. D. Shen and C. Davatzikos, "An adaptive-focus deformable model using statistical and geometric information," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 906–913 (2000).
27. D. Shen, E. H. Herskovits, and C. Davatzikos, "An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures," *IEEE Trans. Med. Imaging* **20**(4), 257–270 (2001).
28. T. S. Yoo, *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis*, A K Peters, Wellesley, Massachusetts (2004).
29. B. Castaneda et al., "Application of non-rigid registration to validate detection of prostate cancer using three-dimensional sonoelastography," in *7th Int. Conf. on the Ultrasonic Measurement and Imaging of Tissue Elasticity*, p. 62 (2008).
30. Y. Ou et al., "Sampling the spatial patterns of cancer: optimized biopsy procedures for estimating prostate cancer volume and Gleason score," *Med. Image Anal.* **13**(4), 609–620 (2009).
31. Y. Zhan et al., "Registering histological and MR images of prostate for image-based cancer detection," in *Medical Image Computing and Computer-Assisted Intervention*, R. Larsen, M. Nielsen, and J. Sporring, Eds., pp. 414–424, Springer-Verlag, Berlin Heidelberg (2006).
32. D. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, 2nd ed., Chapman & Hall/CRC, Boca Raton (2000).
33. R. E. Walpole et al., *Probability & Statistics for Engineers & Scientists, Section 9.11*, 9th ed., Prentice Hall, Boston (2012).
34. J. J. Bauer et al., "Three-dimensional computer-simulated prostate models: lateral prostate biopsies increase the detection rate of prostate cancer," *Urology* **53**(5), 961–967 (1999).
35. S. Holm, "A simple sequentially rejective multiple test procedure," *Scand. J. Statist.* **6**(2), 65–70 (1979).
36. J. Zeng et al., "Prostate biopsy protocols: 3D visualization-based evaluation and clinical correlation," *Comput. Aided Surg.* **6**(1), 14–21 (2001).
37. H. B. Carter et al., "Expectant management of prostate cancer with curative intent: an update of the Johns Hopkins experience," *J. Urol.* **178**(6), 2359–2364; discussion 2364–2355 (2007).
38. T. L. Jang et al., "Low risk prostate cancer in men under age 65: the case for definitive treatment," *Urol. Oncol.* **25**(6), 510–514 (2007).
39. L. Klotz, "Active surveillance for prostate cancer: for whom?," *J. Clin. Oncol.* **23**(32), 8165–8169 (2005).
40. M. A. Bjurlin et al., "Optimization of prostate biopsy: the role of magnetic resonance imaging targeted biopsy in detection, localization and risk assessment," *J. Urol.* **192**(3), 648–658 (2014).

Kristians Diaz Rojas is a PhD candidate at Technische Universität München. He received his BS and MS degrees in electronic engineering and digital signal and image processing from Pontificia Universidad Católica del Perú in 2003 and 2013, respectively.

Maria L. Montero, PhD, is an adjunct professor in the Department of Science, Section of Mathematics, Laboratory Statistics at the Pontificia Universidad Católica del Perú.

Jorge Yao, MD, MBA, was with the Department of Pathology and Laboratory Medicine at the University of Rochester Medical Center. He is currently assistant laboratory director (histology, cytology) at Pathline Labs.

Edward Messing, MD, is the Winfield W. Scott Chair of Urology at the University of Rochester Medical Center. He is also a professor in the Departments of Urology and Pathology and Laboratory Medicine, and the Cancer Center.

Anees Fazili, MD, is a chief resident in the Department of Urology at the University of Rochester Medical Center.

Jean Joseph, MD, MBA, is a professor in the Department of Urology and the Cancer Center at the University of Rochester Medical Center.

Yangming Ou, PhD, is a research fellow in the Athinoula A. Martinos Center for Medical Imaging, Harvard Medical School, Massachusetts General Hospital.

Deborah J. Rubens, MD, is an associate chair of imaging sciences and professor of imaging sciences at the University of Rochester Medical Center, and professor of biomedical engineering at the University of Rochester.

Kevin J. Parker, PhD, is the William F. May Professor of Engineering in the Departments of Electrical & Computer Engineering, Biomedical Engineering, and Imaging Sciences and dean emeritus of the Hajim School of Engineering and Applied Sciences at the University of Rochester.

Christos Davatzikos, PhD, is the Wallace T. Miller Senior Professor of Radiology and Electrical & Computer Engineering, director of the Center for Biomedical Image Computing and Analytics, and director of the Section of Biomedical Image Analysis at the University of Pennsylvania.

Benjamin Castaneda, PhD, is an associate professor in the Department of Engineering, Section in Electrical & Electronics, Laboratory of Medical Imaging, at the Pontificia Universidad Católica del Perú.