Published in final edited form as:

Appl Immunohistochem Mol Morphol. 2009 October ; 17(5): 442–450. doi:10.1097/PAI.0b013e31819e6d65.

Computer-Aided Detection of Prostate Cancer on Tissue Sections

Yahui Peng, MS^{1,*}, Yulei Jiang, PhD¹, Shang-Tian Chuang, DO², and Ximing J. Yang, MD, PhD³

¹Department of Radiology, The University of Chicago, Chicago, IL 60637

²Department of Pathology, Moffitt Cancer Center, Tampa, FL 33612

³Department of Pathology, Northwestern Memorial Hospital, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611

Abstract

We report an automated computer technique for detection of prostate cancer in prostate tissue sections processed with immunohistochemistry. Two sets of color optical images were acquired from prostate tissue sections stained with a double-chromogen triple-antibody cocktail combining alpha-methylacyl-CoA racemase (AMACR), p63, and high-molecular-weight cytokeratin (HMWCK). The first set of images consisted of 20 training images (10 malignant) used for developing the computer technique and 15 test images (7 malignant) used for testing and optimizing the technique. The second set of images consisted of 299 images (114 malignant) used for evaluation of the performance of the computer technique. The computer technique identified image segments of AMACR-labeled malignant epithelial cells (red), p63-and HMWCK-labeled benign basal cells (brown), and secretory and stromal cells (blue) for identifying prostate cancer automatically. The sensitivity and specificity of the computer technique were 94% (16/17) and 94% (17/18), respectively, on the first (training and test) set of images, and 88% (79/90) and 97% (136/140), respectively, on the second (validation) set of images. If high-grade prostatic intraepithelial neoplasia (HGPIN), which is a precursor of cancer, and atypical cases were included, the sensitivity and specificity were 85% (97/114) and 89% (165/185), respectively. These results show that the novel automated computer technique can accurately identify prostatic adenocarcinoma in the triple-antibody cocktail-stained prostate sections.

Keywords

Prostate Cancer; Immunohistochemistry; Computer-Aided Diagnosis; Alpha-methylacyl-CoA Racemase (AMACR)

I. INTRODUCTION

Prostate cancer is the most frequently diagnosed non-skin cancer in American adult men. An estimated 186,320 new prostate cancer cases and 28,660 prostate cancer deaths will occur in 2008.¹ Screening with prostate-specific antigen (PSA) and digital rectal examination (DRE) helps detect prostate cancer, and histologic analysis provides the definitive diagnosis.^{2,3} In addition to the routine hematoxylin and eosin (H&E) staining, immunohistochemistry (IHC) has become a valuable tool in the clinical diagnosis of prostate cancer.⁴ Recently, a prostate cancer marker, alpha-methylacyl-CoA racemase (AMACR, also known as P504S), was found to be overexpressed in the cytoplasm of prostatic adenocarcinoma cells.^{5–7} A triple-antibody

*Address correspondence to: Yahui Peng, MS Department of Radiology, MC2026, The University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637 telephone: 773-834-5106 fax: 773-702-0371 yahuip@uchicago.edu.

cocktail combining AMACR, the basal cell markers p63, and high-molecular-weight cytokeratin (HMWCK) has been used clinically to improve prostate cancer diagnosis.^{8,9} The triple-antibody cocktail uses two chromogens—one of which stains malignant secretory cell cytoplasm red and the other stains benign basal cells brown—to provide a simple means for the detection of prostatic adenocarcinoma. The technique is used clinically for challenging or questionable cases in which the diagnosis is difficult based on H&E staining alone, especially in the diagnosis of small malignant foci in limited biopsy materials.

Computer-aided diagnosis (CAD), a technique in which the results of computer image analysis are provided to physicians as an aid for diagnostic decision-making, has been developed for the detection of breast, lung, and colon cancer in radiology practice.^{10–12} It has been shown that CAD can improve radiologists' performance in the interpretation of mammograms,¹³ while investigations of the clinical effects of CAD continue.^{14–17} Concepts similar to CAD have been proposed for quantitative analysis of pathology images,^{18–21} however, CAD has not been used clinically in prostate cancer diagnosis. Our purpose in this work was to evaluate the feasibility of developing an automated computer technique for detection of prostatic adenocarcinoma in tissue sections stained with the AMACR/p63/HMWCK triple-antibody cocktail.

II. MATERIALS AND METHODS

II.A. Case selection and immunohistochemistry

This study received Institutional Review Board approval from each participating institution. Two sets of prostate tissue images were collected. One genitourinary pathologist (X.J.Y.) collected the first set of images of 10 prostate cases (5 prostatectomy and 5 biopsy specimens) from clinical practice at Northwestern Memorial Hospital. Those images contained classic benign or malignant prostatic glands. The second set of images was collected retrospectively from the surgical pathology archives of prostate biopsy and prostatectomy at Feinberg School of Medicine, Northwestern University. A total of 164 cases (63 prostatectomy and 101 biopsy specimens) accessioned between July 2004 and December 2006 was identified. At least one section in each case was stained with the double-chromogen AMACR/p63/HMWCK cocktail. All patients were diagnosed with acinar adenocarcinoma except one with ductal adenocarcinoma. The mean age of patients was 64 year old (range 41–85). The mean PSA level was 5.6 ng/mL. Half of the patients (82/164) had Gleason score 6, 37.2% (61/164) had Gleason sums of 7, and 12.8% (21/164) had Gleason sum equal to or greater than 8. Data on the size of cancer involvement were not collected because size does not affect the diagnosis of prostate cancer. Two genitourinary pathologists (X.J.Y. and S.T.C.) reviewed the sections of all cases independently to confirm the diagnoses of either prostatic adenocarcinoma or non-malignant prostate tissue.

IHC staining for AMACR, p63, and HMWCK was performed prospectively for the purpose of making clinical diagnoses on 4- μ m formalin-fixed and paraffin-embedded tissue sections with the triple-antibody cocktail according to the protocol recommended by the manufacturer (Biocare, Concord, CA). Briefly, antigen retrieval was performed at 125°C and at a pressure of 125 PSI for approximately 20 minutes. Sections were then placed in Biocare's Sniper solution for 15 minutes for blocking of non-specific antibody binding. The prediluted PIN-4 triple-antibody cocktail (containing rabbit anti-AMACR monoclonal, mouse anti-p63 monoclonal, and mouse anti-HMWCK monoclonal antibodies) was applied at room temperature for 30 minutes. Anti-mouse horseradish peroxidase and anti-rabbit alkaline phosphatase were then applied for 30 minutes, followed by the application of the chromogens betazoid (labeling p63 and HMWCK brown) and Vulcan Fast Red (labeling AMACR red) for 5 and 8 minutes, respectively. Finally, sections were counterstained in hematoxylin for 3 minutes. A sample image with the triple-antibody-cocktail staining is shown in Fig. 1.

II.B. Image sets

We used an Olympus BX40 microscope and an Olympus DP70 V-TVO.5XC digital camera to acquire the two sets of images. The two pathologists recorded digital images of representative regions of interest in each case. Because prostate cancer is an inhomogeneous and multi-focal disease, it is common for tissue samples of a prostate cancer specimen to exhibit more than one focal diagnosis characteristic, and, therefore, independent analysis and diagnosis of multiple sections in each case are often necessary.

The first set of images (10 cases) consisted of a total of 35 images that depict glands with classical malignant and benign diagnoses. No high-grade prostatic intraepithelial neoplasia (HGPIN), a precursor of cancer, was identified in these images. We divided these images into two groups: 20 training images (10 malignant and 10 benign) that we used for development of the computer technique and 15 test images (7 malignant and 8 benign) that we used for testing and optimizing the computer technique. For the first set of images, digital color images were acquired at powers of 20x, 40x, or 60x and stored in a size of 1360 × 1024 pixels in bitmap format.

The second set of images (164 cases) was collected later and initially consisted of 302 images. In each case, 1-8 digital color images were acquired at 20x power and stored in the size of 1360 × 1024 pixels as TIFF (tagged-image-file-format) files. Three images were later excluded because of poor image quality caused by section movement during image acquisition, thus leaving 299 images in the second set of images.

We divided these 299 images into four groups based on the diagnoses given by the two pathologists: a consensus malignant group that consisted of images only containing prostatic adenocarcinoma (n = 90), a consensus HGPIN group that consisted of images not containing invasive glands (n = 42), a consensus benign group (n = 140), and a group of images that initially lacked a consensus diagnosis between the two pathologists (n = 27). Through subsequent discussion, the pathologists determined that the group of images that initially lacked a consensus diagnosis consisted of 2 benign, 1 HGPIN, and 24 malignant images.

II.C. Development of the computer technique

We developed a computer technique that identifies the following three types of cellular and glandular structures by color in the tissue section images: malignant secretory cell cytoplasm (stained red by AMACR), basal cells (stained brown by p63 and HMWCK), and stromal and secretory cell nuclei (counterstained blue by hematoxylin). Based on these structures, the computer identified the prostate epithelium, which consists of basal cells and secretory cells. We used the 20 training images of the first image set to develop the computer technique and used the 15 test images of the first image set to test it. Subsequently, after optimizing the computer technique based on the entire first set of 35 images, we validated the computer technique independently on the second set of 299 images. All tests of the computer image analysis were done in a blinded fashion: the images were provided by the pathologists for applying the computer image analysis without revealing the corresponding diagnosis, and subsequently the pathologists graded the results of the computer image analysis against their diagnoses.

II.C.1. Segmentation of epithelium—Figure 2 schematically illustrates the segmentation method. We decomposed each original color optical image into two color components—the Red and Yellow color channels—via the red-green-blue (RGB) and cyan-magenta-yellow-key (CMYK) color models,²² respectively. Each color channel is a gray-scale image in which the pixel value represents the intensity of the particular color component. After smoothing the two

color-channel images to suppress noise of pixel-to-pixel fluctuation,²³ we applied an automated global-thresholding method.²⁴

Because both the brown and blue colors appear in the Red channel with low intensity, and because both the brown and red colors appear in the Yellow channel with high intensity, we obtained two images from the Red and Yellow color-channel images: one that depicted only brown and blue colors, and the other that depicted only brown and red colors. After removing isolated pixels, we used logical operators to combine these two images into three separate images of the red, brown, and blue colors based on the following information: the brown color appears in both (of the parent) images, whereas the blue and red colors appear in only one of the two (parent) images. Independently, the Hue channel (via the hue-saturation-value [HSV] color model), in which pixel values correspond to the color spectrum, was also used for identifying the red, brown, and blue colors based on their definitions in the Hue channel.²² Only image regions that were identified of identical colors in both of these parallel analyses were retained for further analysis, and other pixels were considered as spurious noise and removed.

Subsequently, regions in all three of the color images that were spatially connected regardless of color were merged together to form regions that corresponded to glandular epithelium. Finally, a region-growing method was applied for enlarging candidate epithelium regions in the Green-channel image (via the RGB color model) for obtaining more accurate epithelium regions.

II.C.2. Malignant vs. non-malignant classification—The computer technique used the following rules to determine whether an image contained prostate cancer. Epithelial regions containing basal cells (identified by brown color) were classified as benign. Furthermore, all of the following, which lacked evidence of basal cells, were also classified as benign: isolated incidental AMACR staining of area smaller than $250 \mu\text{m}^2$ (about 5 times the size of a typical red blood cell nucleus); epithelium with AMACR staining less than 30% of the total area; and isolated regions containing no more than three secretory cell nuclei. All other segmented epithelial regions were classified as malignant. Specifically, malignant epithelium was identifiable by non-incidental AMACR staining of the cytoplasm (red) and the absence of basal cells (brown). An image was reported as containing prostatic adenocarcinoma if, and only if, it contained any region identified as malignant. These criteria are consistent with clinical diagnostic criteria used by pathologists. Because the first set of images used for developing the computer technique did not include any case of HGPIN, the computer technique was not designed to identify HGPIN specifically. Instead, the computer technique identified HGPIN cases as being non-malignant.

II.D. Statistical analysis

We calculated the sensitivity and specificity of the computer image analysis. Sensitivity was defined as the proportion of images that contained malignant tumor (determined by pathologists X.J.Y. and S.T.C.) that were correctly identified by the computer. Specificity was defined as the proportion of images that did not contain malignant tumor that were correctly identified by the computer. Sensitivity and specificity were calculated separately for the two sets of images. For the second set of images, sensitivity and specificity were calculated both including and excluding images of HGPIN, and both including and excluding images that initially lacked a consensus diagnosis. The exact binomial 95% confidence intervals (CIs) were calculated for all sensitivity and specificity values.²⁵

III. Results

Table 1 summarizes the results of the computer image analysis. On the first set of images (classic malignant and benign glands), the computer technique attained a perfect performance on the 20 training images and only a slightly reduced performance on the 15 test images. For the combined 35 training and test images, the computer technique had sensitivity and specificity values of 94% (16/17, CI: 71%-100%) and 94% (17/18, CI: 73%-100%), respectively.

On the second, validation, set of images, the performance of the computer technique remained high. It should be noted that HGPIN images were treated as “non-malignant” in Table 1. If HGPIN images were excluded from this analysis, the sensitivity would have remained unchanged at 85% (97/114, CI: 77%-91%), and the specificity would have been 97% (138/142, CI: 93%-99%). If, instead, images that initially lacked a consensus diagnosis were excluded from the analysis, the sensitivity and specificity would have been 88% (79/90, CI: 79%-94%) and 90% (163/182, CI: 84%-94%), respectively. Finally, if both HGPIN images and images that initially lacked a consensus diagnosis were excluded, the sensitivity and specificity would have been 88% (79/90, CI: 79%-94%) and 97% (136/140, CI: 93%-99%), respectively.

Figure 3 shows two cases and the corresponding results of the computer image analysis. In the case containing malignant glands, the computer correctly identified HGPIN as non-malignant in the vicinity of the malignant glands, which it also identified correctly (compare Fig. 3A and 3D). In the case containing only benign glands, the computer was able to classify the benign glands correctly even though a small amount of benign cytoplasm was positive for AMACR staining (compare Fig. 3E, 3F, and 3H).

IV. DISCUSSION

Computer-aided diagnosis is currently employed in radiology practice to help radiologists detect breast cancer in mammograms. The indications from experience and from studies are that CAD helps radiologists detect more cancers and incur commensurate additional false positives. With further development, with better-designed clinical studies, and with a better understanding of how radiologists can optimize their use of this technology, the benefit of CAD for clinical radiology likely will be further validated in the future.

Computer image analysis has entered clinical pathology, even though it is still limited. For example, computer image analysis can screen for atypical cells in the cytologic analysis of Pap smears. For histology, an Automated Cellular Imaging System (ACIS, Clariant, Aliso Viejo, CA) is used routinely in clinical laboratories for measurement of estrogen- and progesterone-receptor expression quantitatively in tumor cells. However, CAD has not been used clinically for prostate cancer diagnosis.

AMACR has been proven to be one of the few biomarkers that can distinguish cancer from benign cells with high sensitivity and high specificity for the diagnosis of prostatic adenocarcinoma.^{7,26} Using p63 and HMWCK individually or jointly has been shown to be reliable in distinguishing benign mimickers of prostate cancer from prostatic adenocarcinoma.^{4,27-30} Combining AMACR, p63, and HMWCK, the triple-antibody cocktail staining improves sensitivity and specificity, and is used widely in clinical practice as an adjunct to H&E staining in the histopathologic diagnosis of prostate cancer.^{8,9} The double-chromogen system of this cocktail, which colors malignant secretory cell cytoplasm red and benign prostatic basal cells brown, provides a simple and straightforward means for the visual identification of prostatic adenocarcinoma. The multiple-color appearance of the stained image also provides an excellent basis for testing the feasibility of developing computer image analysis methods for the detection of prostatic adenocarcinoma in tissue sections.

Sufficiently accurate computer analysis of histology images could aid pathologists in making clinical diagnoses in the future in a possible scenario shown in Fig. 4. Automated computer image analysis can provide pathologists with quantitative information at the time of tissue interpretation, either with concurrence or with the suggestion of a plausible alternative analysis. Such computer assistance will not subtract from the critical role of pathologists in tissue diagnosis, but likely can make pathologists work more efficiently and more accurately. This type of assistance could prove beneficial to practitioners, even for highly experienced experts. Automation or semi-automation of some of these steps, such as tissue processing, staining, and digitization, is already possible today. For example, current technology is capable of digitizing entire sections, but advances in the digitization process and in the handling of large amount of digital image data are necessary to make it practical for daily clinical practice. In this work, we focused on the computer image analysis of a specific type of IHC images, and we are currently developing computer image analysis of H&E images. Several of these computer techniques are necessary in combination for assisting pathologists clinically, and it is important for each of these techniques to be individually developed, optimized, and tested.

In this study, we have developed a computer technique to analyze triple-antibody cocktail-stained prostate tissue section images accurately. We developed the computer technique based on images of classic malignant and benign glands (the first set of images), and then evaluated the computer performance in a blinded fashion on images acquired independently from routine clinical practice (the second set of images). In the first set of images, the malignant glands had strong AMACR staining and well-defined boundaries, whereas the benign glands clearly had a layer of basal cells and had weak, if any, AMACR staining. The computer technique performed well on the classic malignant and benign glands in the first set of images, correctly classifying all but one malignant and one benign image (because of a limited field of view in these images).

The computer technique also performed well on the second set of validation images. These images were more difficult for the computer to analyze because there were large variations in the way the tissues were processed and stained, and how the digital images were acquired. These variations impose challenges for the computer image analysis in addition to the atypical appearance of the tissue sections. Although these images included 42 HGPIN images and 27 images that initially lacked a consensus diagnosis from two genitourinary pathologists, the computer technique was able to achieve a relatively high performance in sensitivity (85%) and specificity (89%).

If the HGPIN images and images that initially lacked consensus diagnosis were excluded, the sensitivity would have increased to 88% and the specificity would have increased to 97%. These encouraging results indicate that our computer technique is capable of detecting prostatic adenocarcinoma accurately in the triple-antibody cocktail-stained tissue section images.

The computer technique is capable of identifying a cluster of cancer glands in between benign or HGPIN glands. This infiltrative pattern is common for prostatic adenocarcinoma and can be identified in many cases in this study. The computer technique is also capable of identifying small foci of adenocarcinoma accurately (Fig. 5), even if there is only one single small malignant gland in the image (Fig. 6). These capabilities are important for the diagnosis of prostatic adenocarcinoma on needle biopsy because often a cancer focus shows only a limited number of malignant glands meandering in between non-malignant glands.

The criteria for the computer to identify cancer were designed to target the classic acinar adenocarcinoma in this initial study. While the computer succeeded in correctly identifying two ductal carcinoma images that were present in our images as malignant, the computer

performance on ductal adenocarcinoma cases needs to be tested in the future. Refinement of the criteria for identifying variant cancer types is potentially needed.

The main reason for missed malignant cases was due to errors in image segmentation. Because of color artifacts, the computer image segmentation method may not identify the microscopic structures correctly (Fig. 7). This is the reason for 12 of the 17 cancer images that the computer missed. The remaining five false-negative images were caused by weak AMACR staining (the computer was not able to identify the pale red color).

Reasons for benign cases that the computer confused with malignancy included: (a), missing basal cell layer in benign glands (8/20); (b), isolated prostatic epithelium separate from its basal cell layer (7/20; Fig. 8); and (c), both (a) and (b) above (5/20). All of these 20 benign glands had strong or intermediate AMACR expression, and resulted in an isolated piece of the epithelium showing AMACR expression without expression of the basal cell markers. The computer picked up the isolated epithelium and marked them as malignant.

We have developed a novel computer-aided detection technique that can be used for identifying prostatic adenocarcinoma in immunostained tissue sections. The performance of this computer technique was validated in a blinded fashion on a set of 299 images and achieved sensitivities of 85% to 88% and specificities of 89% to 97%. These preliminary results are encouraging and support the feasibility of computer image analysis of certain clinical prostate histology sections. Future studies need to test the computer performance on, and potentially refine the algorithm for, variant types of prostate cancer such as prostatic ductal adenocarcinoma, atypical small acinar proliferation, and other benign prostate cancer mimickers. With further development, computer image analysis could become a tool for aiding the pathologic diagnosis of prostate cancer in clinical practice.

Acknowledgments

This work was supported in part by the National Cancer Institute (through grant R21 CA97308) and the National Institute of Biomedical Imaging and Bioengineering (through grant R21 EB006466) of the National Institutes of Health.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58:71–96. [PubMed: 18287387]
2. Epstein, JI.; Yang, XJ. Prostate biopsy interpretation. 3rd ed.. Lippincott Williams & Wilkins; Philadelphia: 2002.
3. Bostwick, DG. Pathology of the Prostate. Churchill Livingstone; New York: 1990.
4. Wojno KJ, Epstein JI. The utility of basal cell-specific anti-cytokeratin antibody (34-beta-E12) in the diagnosis of prostate-cancer—a review of 228 cases. *Am J Surg Pathol* 1995;19:251–260. [PubMed: 7532918]
5. Jiang Z, Woda BA, Rock KL, et al. P504S: a new molecular marker for the detection of prostate carcinoma. *Am J Surg Pathol* 2001;25:1397–1440.
6. Rubin MA, Zhou M, Dhanasekaran SM, et al. α -Methylacyl coenzyme A racemase as a tissue biomarker for prostate cancer. *JAMA* 2002;287:1662–1670. [PubMed: 11926890]
7. Adley BP, Yang XJ. Application of alpha-methylacyl coenzyme A racemase immunohistochemistry in the diagnosis of prostate cancer—A review. *Anal Quant Cytol Histol* 2006;28:1–13. [PubMed: 16566275]
8. Ng VWL, Koh M, Tan SY, et al. Is triple immunostaining with HMWCK, p63, and Racemase in prostate cancer advantageous? A tissue microarray study. *Am J Clin Pathol* 2007;127:248–253. [PubMed: 17210521]

9. Jiang Z, Li CZ, Fischer A, et al. Using an AMACR(P504S)/HMWCK/p63 cocktail for the detection of small foci prostate carcinoma in needle biopsy specimens. *Am J Clin Pathol* 2005;123:231–236. [PubMed: 15842047]
10. Jiang Y, Nishikawa RM, Schmidt RA, et al. Improving breast cancer diagnosis with computer-aided diagnosis. *Acad Radiol* 1999;6:22–33. [PubMed: 9891149]
11. Kobayashi T, Xu XW, MacMahon H, et al. Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. *Radiology* 1996;199:843–848. [PubMed: 8638015]
12. Yoshida, Dachman AH. CAD techniques, challenges, and controversies in computed tomographic colonography. *Abdom Imaging* 2005;30:26–41. [PubMed: 15647868]
13. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology* 2001;220:781–786. [PubMed: 11526282]
14. Fenton JJ, Taplin SH, Carney PA, et al. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med* 2007;356:1399–1409. [PubMed: 17409321]
15. Nishikawa RM, Schmidt RA, Metz CE. Computer-aided screening mammography. *N Engl J Med* 2007;357:84. (author reply 85). [PubMed: 17615630]
16. Feig SA, Birdwell RL, Linver MN. Computer-aided screening mammography. *N Engl J Med* 2007;357:84. (author reply 85). [PubMed: 17615628]
17. Gromet M. Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. *AJR Am J Roentgenol* 2008;190:854–859. [PubMed: 18356428]
18. Thompson D, Bartels PH, Bartels HG, et al. Image segmentation of cribriform gland tissue. *Anal Quant Cytol Histol* 1995;17:314–322. [PubMed: 8534334]
19. Bartels PH, Thompson D, Bartels HG, et al. Machine vision-based histometry of premalignant and malignant prostatic lesions. *Pathol Res Pract* 1995;191:935–944. [PubMed: 8606876]
20. Diamond J, Anderson NH, Bartels PH, et al. The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. *Human Pathol* 2004;35:1121–1131. [PubMed: 15343515]
21. Saidi O, Cordon-Cardo C, Costa J. Technology Insight: will systems pathology replace the pathologist? *Nat Clin Pract Urol* 2007;4:39–45. [PubMed: 17211424]
22. Malacara, D. Color vision and colorimetry: theory and applications. SPIE Press; Washington: 2002.
23. Soille, P. Morphological image analysis: principles and applications. 2nd ed.. Springer; New York: 2003.
24. Otsu N. A threshold selection method from gray level histograms. *IEEE Trans Syst Man Cybern* 1979;9:62–66.
25. Zar, JH. Biostatistical analysis. 4th ed.. Prentice-Hall; New Jersey: 1999.
26. Carswell BM, Woda BA, Wang X, et al. Detection of prostate cancer by alpha-methylacyl CoA racemase (P504S) in needle biopsy specimens previously reported as negative for malignancy. *Histopathology* 2006;48:668–673. [PubMed: 16681682]
27. Shah RB, Kunju LP, Shen R, et al. Usefulness of basal cell cocktail (34betaE12 + p63) in the diagnosis of atypical prostate glandular proliferations. *Am J Clin Pathol* 2004;122:517–523. [PubMed: 15487448]
28. Zhou M, Shah R, Shen R, et al. Basal cell cocktail (34betaE12 + p63) improves the detection of prostate basal cells. *Am J Surg Pathol* 2003;27:365–371. [PubMed: 12604893]
29. Oliai BR, Kahane H, Epstein JI. Can basal cells be seen in adenocarcinoma of the prostate? An immunohistochemical study using high molecular weight cytokeratin (clone 34betaE12) antibody. *Am J Surg Pathol* 2002;26:1151–1160. [PubMed: 12218571]
30. Osunkoya AO, Hansel DE, Sun X, et al. Aberrant diffuse expression of p63 in adenocarcinoma of the prostate on needle biopsy and radical prostatectomy: report of 21 cases. *Am J Surg Pathol* 2008;32:461–467. [PubMed: 18300803]

Figure 1.

An image of prostate tissue processed with the triple-antibody-cocktail staining. Arrows indicate a number of small malignant glands with alpha-methylacyl-CoA-racemase (AMACR) staining in the secretory cell cytoplasm (red). Near the bottom of the image, arrowheads indicate a clearly visible basal-cell layer stained by p63 and high-molecular-weight cytokeratin (HMWCK) in two benign glands (brown). Stromal and secretory cell nuclei are stained by hematoxylin (blue).

Figure 2.

Flowchart that illustrates the computer-segmentation technique. Rounded rectangles with shadows represent single images (the dashed-line rectangle represents three separate images), which can be the original input color image, individual color channels, binary images after thresholding, or the output image. Other rectangles represent operations applied to the images. Two different thresholding methods were applied. See text for details.

Figure 3.

Examples of the computer image analysis in a malignant (**top**) and a benign (**bottom**) case. The results of the malignant case show (A) a section containing glands of prostatic adenocarcinoma; (B) computer results showing the segmentation of cytoplasm positive for alpha-methylacyl-CoA-racemase (AMACR) staining; (C) computer results showing the segmentation of basal cells positive for p63/HMWCK staining; and (D) final computer result showing the segmentation of malignant epithelial components. The results of the benign case show (E) a section containing benign glands; (F) computer results showing the segmentation of cytoplasm positive for AMACR staining; (G) computer results showing the segmentation of basal cells positive for p63/HMWCK staining; and (H) final computer result showing no malignant glands.

Figure 4.

Illustration of a possible scenario for future clinical application of computer-aided diagnosis of prostate cancer. With automation, computer image analysis can provide additional information to assist the pathologist in making a cancer diagnosis. The present work addresses computer image analysis of one type of IHC images (shown in the right-hand column); work on computer image analysis of H&E images is currently ongoing (denoted in grey in the left-hand column).

Figure 5.
An example of infiltrative malignant glands. **Left**, the original image. **Right**, the computer-identified malignant glands.

Figure 6.
An example of a single small cancerous gland identified by the computer. **Left**, the original image. Right, the computer reports the identification of cancer.

Figure 7.

A false-negative case of **(left)** original image and **(right)** computer segmentation result. The arrow indicates a focus of malignant glands. In the right image, dark regions correspond to the “brown” color and gray regions correspond to the “red” color of the original (left) image. The malignant glands were missed in this case because the computer over-estimated the “brown” color regions near the focus of malignant glands, incorrectly indicating the presence of a basal-cell layer.

Figure 8.

A false-positive case of **(left)** original image and **(right)** computer segmentation result. The arrow indicates a tissue fragment with alpha-methylacyl-CoA-racemase (AMACR) staining isolated from the gland. The computer mistakenly identified the isolated tissue fragment as malignant based on color analysis.

Table 1

Performance Summary of the Computer Technique

Cases	Sensitivity (%) (CI) [†]	Specificity (%) (CI)	Computer Results	Diagnostic Truth		Total
				Malignant	Benign [‡]	
Image Set 1a: Training Cases	100 (74-100)	100 (74-100)	Malignant	10	0	10
			Benign	0	10	10
			Total	10	10	20
Image Set 1b: Test Cases	86 (42-100)	88 (47-100)	Malignant	6	1	7
			Benign	1	7	8
			Total	7	8	15
Image Set 2: Validation Cases	85 (77-91)	89 (84-93)	Malignant	97	20	117
			Benign	17	165	182
			Total	114	185	299

[†] CI: 95% confidence interval.

[‡] Represents both benign and high-grade prostatic intraepithelial neoplasia cases.