Purpose: The reliability and objectivity of computer assisted transrectal ultrasound are examined.

Materials and Methods: Pathological examination of radical prostatectomy specimens was compared prospectively to automated cancer detection in corresponding transrectal ultrasound images.

Results: For automated cancer detection, a sensitivity of 0.75 and a specificity of 0.78 were obtained. Moreover, 74% of human interpretation of the percentage of malignancy in the analyzed images was equal to the actual calculated percentage (Pearson's product moment correlation coefficient 0.85).

Conclusions: Comparing these results to those obtained with normal transrectal ultrasound, automated analysis provides additional information in the interpretation of transrectal ultrasound images by color coding them in an objective manner according to the probability of malignancy.

Key Words: ultrasonography, prostatectomy, prostatic neoplasms, computers

In our department a system for automated analysis of ultrasonographic prostate images has been developed. With this system, structures seen on transrectal ultrasonographic images, used by many urologists for evaluation of the prostate gland, were quantified and correlated with the histopathological results of the corresponding tissue. This relationship provided visualization of image texture related to pathological structures, which are often difficult to discern or are not perceptible to the human eye. With image processing algorithms, transrectal ultrasound images are analyzed and color coded according to the probability of malignancy. A complete description of the conditions and methods used has been reported previously.

The reliability of this tissue discrimination was proved retrospectively and prospectively. For training purposes, 381 images of needle biopsies were used (229 benign and 102 malignant tissues). The results obtained for cancer detection showed a sensitivity of 0.77 and a specificity of 0.73. Prospective evaluation (using cross-validation) resulted in a sensitivity of 0.70 and a specificity of 0.74. Moreover, the reliability of the histopathological analysis of the needle biopsies used in the training phase was tested by reviewing the pathological specimens.

In the training phase images of needle biopsies were used, more than half of which were recorded from locations with no ultrasonographic suspicion for malignancy (iso-echoic). Most biopsies were taken from the outer gland (peripheral and central zones). It is unknown to what extent the ultrasonographic properties of the biopsy specimens used are a significant reflection of the characteristics of the total prostate. Because of histological and biological discrepancies between malignancies in the outer and inner (transition zone) gland, cancer in the different zones of the gland will be reflected differently on transrectal ultrasound images. The distribution of cancer was described by McNeal et al as 68% in the peripheral zone, 8% in the central zone and 24% in the transition zone. Only when the recorded needle biopsy images from the inner gland were representative for ultrasonographic reflection of malignancy in this area was computer analysis reliable in this region. To investigate the reliability of the automated analysis for the entire gland, a study was conducted for prospective comparison of computer analysis (color coding) of transrectal ultrasound images recorded before radical prostatectomy with corresponding histopathological analysis of the tissue specimens. Moreover, interpretative differences of the color coded images among several observers and the objectivity of these interpretations were investigated. Finally, the results of the computer analysis were compared to the diagnostic results using transrectal ultrasound only.

Material and Methods

The study population comprised 12 men 58 to 72 years old (mean age 64 years) who underwent radical prostatectomy. Preoperatively, a series of transrectal ultrasound images was recorded using a Kretz Combison 330 ultrasound scanner in combination with a 7.5 MHz, multiplane transrectal transducer (VRW77AK). The probe was fixed in a stepwise unit. The first image was recorded by digitization of the video signal at the base of the prostate. After retracting the probe 4 mm., the next image was recorded. This process was repeated until the apex was reached. The gland was completely imaged in an average of 12 transverse cross sections within approximately 1 minute.

The prostatectomy specimens, fixed overnight in 5% buffered formalin, were correspondingly step-sectioned at 4 mm. intervals perpendicular to the rectal surface of the prostate. The apical and basal sections were cut radially to determine the status of the surgical margins. After embedding in paraffin, a 5 μm thin tissue section was cut from the top of each 4 mm. section, and stained with hematoxylin and eosin for histological examination. From each 4 mm. section a photograph of the macroscopic image was taken on which the locations of the cancerous areas were outlined after histological examination by an experienced pathologist. These photo-
The color coded transrectal ultrasound images and outlined photographs of the prostatectomy specimens were compared prospectively. Interpretation of these images only involved discrimination between different colors, so that no particular knowledge was necessary. They were compared by 5 examiners with different backgrounds (1 experienced urologist, 3 persons familiar with transrectal ultrasound images and 1 with no ultrasonographic experience at all). For this comparison the prostate was divided into 8 segments. The left and right sides of the gland were split into apical dorsal and apical ventral segments, as well as dorsal and ventral segments at the base (fig. 1). For each segment on the computer images and histopathologically marked photographs an observer had to score prospectively the relative volume of the malignancy (that is the ratio between cancer and non-cancerous tissue): 0—0% relative volume of malignancy, 1—0 to 10%, 2—10 to 25%, 3—25 to 50% and 4—50 to 100%.

To investigate the difference between interpretation of the automated analyses and the outlined photographs, for each segment a dependent t test was performed to compare both interpretations, which resulted in 8 analyses for every 12 prostate x 5 observers = 60 observations. The categories on the malignancy scale were also combined according to the presence or absence of malignancy. This combination is performed using 2 different cutoff values for marking a region as cancerous in the computer images (relative volume of malignancy of more than 0% and more than 10%). The cutoff value of more than 10% was used to eliminate influences of artifacts in the transrectal ultrasound images, which were often manifested in small areas and marked as malignant on the computer images. For the histopathologically marked photographs, combination is performed using a relative volume of malignancy of more than 0%. In this manner the true negative, false-positive, true positive and false-negative interpretations of classifications were calculated separately for the different segments of the prostate. By combining these values the specificity and sensitivity for automated cancer detection were calculated.

To test the inter-observer variability for the interpretation of the automated analysis, the standard deviation for the individual interpretation of each observation was computed. Moreover, the observations (5 observers x 12 prostates x 8 segments = 480 total) were compared with the actually computed relative volume of malignancy in the images. To investigate the objectivity, Pearson's product moment correlation coefficient (R) and the distribution of the differences between the interpreted and actual relative volumes of malignancy were computed.

The transrectal ultrasound images used for computer analysis were interpreted by our experienced urologist (J. de la R.). To evaluate the presence or absence of carcinoma, the criteria used were seminal vesicle asymmetry, capsular irregularity, hypoechoic lesions and clustered hyperechoic foci within a hypoechoic lesion.

**RESULTS**

An example of a transrectal ultrasound cross section with corresponding automated analysis and histological section on which the tumor area was outlined is presented in figure 2. The computer image is color coded according to the computed probability for malignancy. These probabilities are translated to a color scale ranging from blue (0% probability for malignancy) to red (100% probability). The histopathologically marked photographs and color coded transrectal ultrasound images were compared using a t test for paired samples. For this test the data were labeled according to the category number on the malignancy scale. Table 1 provides the mean difference (computer analysis minus histological examination) in combination with the standard error for each octant of the prostate. Moreover, the 95% confidence intervals for the differences between both interpretations were computed. In 6 of the 8 segments computer analysis predicted a higher relative volume of malignancy than pathological examination (category range 0.12 to 0.87). In the other 2 cases computer analysis underestimated the relative volume of malignancy (categories 0.22 and 0.25, respectively). However, regarding the confidence intervals, only in 2 cases (base right dorsal and apex right ventral) were the differences between interpretations of the computer images and photographs significant (α = 0.05).

The true negative, false-positive, true positive and false-negative results, computed for each octant and both combination methods, are presented in table 2. The resulting specificity and sensitivity were 0.30 and 0.95, respectively, for relative volume of malignancy more than 0%, and 0.78 and 0.75, respectively, for relative volume of malignancy more than 10%. Comparing the results of both combination methods, a large increase of the specificity can be observed using relative volume of malignancy more than 10% instead of more than 0%, which can be explained by the fact that most of the false-positive interpretations were caused by artifacts in the images resulting in small areas marked as malignant. Using relative volume of malignancy more than 10%, these artifacts were ignored and the specificity increased markedly.

The inter-observer variability for the interpretation of the computer analysis and the histopathologically outlined photographs was computed by the average of the standard deviations among the interpretations of the 5 observers. These deviations ranged from 0.00 to 1.00 (mean 0.32) for the photographs and from 0.00 to 1.14 (mean 0.30) for the computer analysis. For both analyses the maximum deviation occurred in the right dorsal segment at the base of the gland.

Concerning the objectivity of interpretation of the computer images, in 386 of the 480 cases (74.0%) the human interpretation and the actual relative volume of malignancy were equal, while the difference between interpreted and actual relative volume of malignancy was −1 category in 54 (11.3%), 1 category in 14 (2.9%) and 2 categories in 1 (0.2%). Moreover, a Pearson correlation coefficient of 0.85 (p < 0.001) was computed for the relationship between interpreted and actual relative volume of malignancy. Use of transrectal ultrasound alone for the detection of carcinoma resulted in a specificity of 0.83 (automated analysis 0.60) and a sensitivity of 0.74 (automated analysis 0.84) for the dorsal segments, 0.77 (automated analysis 0.83) and 0.12 (automated analysis 0.63), respectively, for the ventral segments and 0.68 (automated analysis 0.78) and 0.50 (automated analysis 0.75), respectively, for the entire prostate gland.

**DISCUSSION**

Although only material from 12 radical prostatectomies was used to compare the results of computer analysis with
histopathological examination, this study provides good insight into the possibility of analyzing whole prostate images. Because the accuracy of the system was already proved in a study using 331 images of needle biopsies, our study was conducted to provide an assessment of the reliability of the tissue discrimination for clinical use. The analysis of whole prostate images and the objectivity as well as variability of the interpretation of the color coded images by the different observers were investigated.

The sensitivity and specificity for automated cancer detection obtained in our study were 0.75 and 0.78, respectively. For these results a cutoff value of 10% was used to mark a segment malignant, which eliminated influences of artifacts but also introduced false-negative classifications. However, we realize that small but poorly differentiated foci of adenocarcinoma, which are not innocuous, could also be missed by computer analysis. Because the classification results were in the same range as those obtained in the biopsy study (sensitivity 0.77 and specificity 0.73), we conclude that these results are reliable for analysis of whole prostate images in clinical use. Moreover, no difference was found concerning the dorsal and ventral sides of the prostate. In both regions there was no significant difference between the interpretation of the computer images and histopathological examination in 3 of the 4 segments, although most of the needle biopsies used for training were from the dorsal area.

Concerning the inter-observer variability for the computer images, a mean deviation was found of 0.30 category, which is approximately the same deviation as found for interpretation of the histopathologically marked photographs. When combining this finding with the correlation of 0.85 between the interpreted and actual relative volumes of malignancy, we conclude that automated tissue discrimination provides a reasonable objective interpretation of transrectal ultrasonography. However, no significant difference was found (2-tailed test, p = 0.218) for interpretation of the computer images between the experienced urologist and the person with no ultrasonographic experience at all, which confirms that no specific knowledge is necessary to interpret the analyzed transrectal ultrasound images.

Malignancy was found in 30 of the 48 ventral segments (62.5%) and 43 of the 48 dorsal segments (89.6%, table 2). Translating these values to the distribution of prostate cancer, 41.1% of the tumors were in the ventral and 58.9% were in the dorsal segments. When we compare these percentages to those presented by McNeal et al., cancer was more frequently located in the transition (ventral) zone in our study. These carcinomas are known to be difficult to detect because they are nonpalpable and often not visible on transrectal ultrasonography. Therefore, especially for this region computer analysis can provide important additional information.

The automated computer analysis only predicts the probability for the presence of malignancy regardless of the grade differentiation of the tumor. Therefore, in the prostatectomy specimens and computer images the relative volume of malignancy was scored without using a grading system for tumor differentiation. Moreover, the estimated relative volumes of malignancy were divided into nonlinear categories, which were used for analysis of the data. With this approximately logarithmic categorization, differences among smaller volumes of malignancy are more emphasized than the same differences among larger volumes. If the categorization shifts 1 category the relative volume of malignancy will be doubled.

The transrectal ultrasound images used in our study were recorded in the transverse plane in contrast to the images used during training on the system. For this purpose the images of needle biopsies used were recorded longitudinally. Indeed, a good correlation was found between texture descriptions from longitudinally and transversely recorded ultrasonographic prostate images. Therefore, the computed correlation between the image texture and histopathology using the needle biopsy images was also applicable on the transverse images recorded before radical prostatectomy.

Transrectal ultrasound and pathology specimens from radical prostatectomy have been compared in several studies to determine the sonographic pattern in prostatic cancer. The methods used for comparison varied from segmentation of the prostate in several regions, for example 2 used by Hardeeman et al. (left and right sides) and 8 (as in our study) used by Jansen et al. to comparison of each cross section. The classification of malignancy varied from the presence or
absence of carcinoma, to use of the Gleason grading system. Most studies presented no sensitivity and specificity results for the use of whole prostate transrectal ultrasound images. Besides, in contrast to our study, most of the above-mentioned studies were performed retrospectively. Therefore, division into 8 segments was used. Consequently, the diagnostic value of transrectal ultrasound in our study was determined using the images recorded for computer analysis. However, in clinical practice transrectal ultrasound is a dynamic process and, therefore, the diagnostic accuracy can be greater.

The division of the gland using proper zonal anatomy is preferable to division into 8 segments. However, it is not always easy or possible to detect zonal anatomy on transrectal ultrasound and corresponding macroscopic images. Therefore, division into 8 segments was used. Consequently, some (peripheral zone) tumors that were identified in the "apex right and left ventral" regions were misclassified as transition or ventral zone tumors.

CONCLUSIONS

Comparing the diagnostic results of transrectal ultrasound to those obtained with computer analysis, the latter can provide additional information in the interpretation of transrectal ultrasound images by indicating regions with an increased probability for malignancy. Certainly, in combination with digital rectal examination and prostate specific antigen levels, computer analysis extends the diagnostic accuracy of ultrasonographic examinations of the prostate for the detection of prostate cancer. Based on our results, we conducted a large clinical study to compare transrectal ultrasound, computer assisted transrectal ultrasound and magnetic resonance imaging using the histopathological examination of radical prostatectomy specimens. In that study tumor volume and grade, the difference between the primary tumor and additional foci, and clinical diagnostic factors, such as digital rectal examination and prostate specific antigen level, will also be incorporated.

REFERENCES