AN ARTIFICIAL NEURAL NETWORK TO PREDICT THE OUTCOME OF REPEAT PROSTATE BIOPSIES

MESUT REMZI, THEODORE ANAGNOSTOU, VINCENT RAVERY, ALEXANDRE ZLOTTA, CARSTEN STEPHAN, MICHAEL MARBERGER, AND BOB DJAVAN, FOR THE EUROPEAN SOCIETY FOR ONCOLOGICAL UROLOGY (ESOU)

ABSTRACT

Objectives. To develop an advanced artificial neural network (ANN) to predict the presence of prostate cancer (PCa) and to predict the outcome of repeat prostate biopsies. The predictive accuracy was compared with the accuracy obtained using standard cutoffs for the free/total (f/t) prostate-specific antigen (PSA) ratio, PSA density (PSAD), PSA density of the transition zone (PSA-TZ), and the total and transition zone volumes. Clinical and biochemical diagnostic tests have been shown to improve PCa detection. When these tests are combined using an ANN, significant increases in specificity at high sensitivity are observed.

Methods. The Vienna-based multicenter European referral database for early PCa detection of 820 men with a PSA level between 4 and 10 ng/mL was used. The presence of PCa was determined using transrectal ultrasound-guided octant needle repeat biopsy. Variables in the database consisted of age, PSA, f/t PSA ratio, digital rectal examination findings, PSA velocity, and the transrectal ultrasound-guided variables of prostate volume, transition zone volume, PSAD, and PSA-TZ. The ANN used in the analysis was an advanced multilayer perceptron selected for accuracy by a genetic algorithm.

Results. The repeat biopsy PCa detection rate was 10% (n = 83). At 95% sensitivity, the specificity for ANN was 68% compared with 54%, 33.5%, 21.4%, 14.7%, and 8.3% for multivariate logistic regression analysis, f/t PSA ratio, PSA-TZ, PSAD, and total PSA, respectively. The ANN reduced unnecessary repeat biopsies by 68% in this study. The area under the curve was 83% for the ANN versus 79%, 74.5%, 69.1%, 61.8%, and 60.5% for multivariate analysis, f/t PSA ratio, PSA-TZ, PSAD, and total PSA, respectively.

Conclusions. The current ANN found a strong pattern predictive of PCa in patients with a negative initial biopsy. By combining the individual clinical and biochemical markers into the ANN, 68% specificity at 95% sensitivity was achieved. The ANN allows more accurate and individual counseling of patients with a negative initial biopsy.

Many investigators have suggested that patients with persistent prostate-specific antigen (PSA) levels of 4 to 10 ng/mL should undergo a repeat prostate needle biopsy; because a significant number of prostate cancers (PCa) can be found. Recent studies have showed prospectively,\(^1\) and retrospectively,\(^2-4\) a PCa detection rate of 10% to 27% on repeat biopsy. In a recent review from the Johns Hopkins Hospital,\(^5\) ANNs have been demonstrated to be superior to standard empirical methods of detecting, staging, and monitoring PCa.

Recently Djavan et al.,\(^6\) Finne et al.,\(^7\) Horninger et al.,\(^8\) and Babaian et al.\(^9\) showed that artificial neural networks (ANNs) increase the predictive accuracy for the initial prostate biopsy compared with PSA-related parameters for patients referred for early detection\(^6\) or screening\(^7,8\) of PCa in PSA ranges of 4 to 10 ng/mL\(^6-8\) and 2.5 to 4 ng/mL.\(^6,8,9\)

Little is known about the validity of ANNs in patients with negative initial biopsies versus standard statistical analysis and singular biochemical parameters. The objective of this study was to de-
velop an advanced ANN to predict the presence of PCa and to predict the outcome of repeated prostate biopsies. The predictive accuracy was compared with the accuracy obtained using standard cutoffs for the free/total (f/t) PSA ratio, PSA density (PSAD), PSA density of the transition zone (PSA-TZ), and the total and transition zone volumes. Individual counseling can be improved and costs reduced by avoiding unnecessary repeated biopsies. Clinical parameters,1,10 as well as biochemical parameters,1,11 have been shown to enhance the specificity of decision-making for repeat biopsy.

MATERIAL AND METHODS

SUBJECTS
In this subanalysis of the prospective European Prostate Cancer Detection Study, 820 patients (48 to 77 years old) were enrolled from January 1997 to January 2001 as consecutive referrals for early PCa detection. All had PSA levels between 4 and 10 ng/mL and a negative initial biopsy, as described previously.6

STUDY DESIGN
Serum levels of total PSA and free PSA were measured from deep frozen serum (−70°C) with the AxSYM total PSA and AxSYM free PSA assays (Abbott Laboratories, Abbott Park, Ill). Digital rectal examination, transrectal ultrasonography, and standard transrectal ultrasound-guided sextant biopsy as described by Hodg et al.12 plus two additional biopsies from the transition zone (TZ) were performed in all subjects for the first and repeat biopsy. Transrectal ultrasonography was performed using a biplanar 7.5-MHz ultrasound probe (Siemens, Germany or Bruel & Kjaer, Denmark). The total prostate and TZ volumes were calculated using the prolate ellipsoid formula (volume = 0.52 × length × width × height). Histopathologic findings from the biopsy samples were classified as negative (benign prostatic tissue, including prostate intraepithelial neoplasia) or positive for PCa. All subjects whose biopsy samples were negative underwent a repeat biopsy after 6 weeks.

STATISTICAL ANALYSIS
The ANN chosen for this analysis was a multilayer perceptron, as described previously.6 For each parameter, a receiver operating characteristic curve was generated. Additionally, the ANN model was compared with the conventional PSA and volume-related parameters. For all parameters, the sensitivity and specificity were calculated. Also a multivariate logistic regression analysis was performed and compared with the ANN findings. Areas under the curve (AUCs) for the receiver operating characteristic were compared with the McNemar test, modified by Bonferroni-Holm. Statistical analyses comparing continuous parameters were performed with the nonparametric Wilcoxon test (Statistical Analysis Systems, JMP, version 3.2.2).

RESULTS
The patient characteristics are shown in Table I. Of the 820 patients, 83 (10%) were diagnosed with PCa on repeat biopsy.

The 820 patients were randomly divided into three groups: group 1, the training group, including 40% of data (n = 320); group 2, the test group, including 30% of the data (n = 250); and group 3, the validation group, with 30% of the data (n = 250). The validation data were not used to develop the ANN and were only used to determine the predictive accuracy of the trained ANN. The only variables selected from the database for the ANN to predict PCa on repeat prostate biopsy were PSA, f/t PSA ratio, total prostate volume, TZ volume, and PSA-TZ.
PSAD, and PSA-TZ (Fig. 1). Age, digital rectal examination findings, and PSA velocity had no impact on the outcome of repeat prostate biopsies. For the multivariate logistic regression analysis, PSA-TZ and f/t PSA ratio were the only independent factors (Table II). The r value for the multivariate logistic regression analysis was 0.215 and the P = less than 0.001.

The cutoffs and specificities at 95% and 90% sensitivity, as well as the AUCs, are shown in Table III. The specificity at 95% sensitivity for the ANN was 68% compared with 54%, 33.5%, 21.4%, 14.7%, and 8.3% for multivariate analysis, f/t PSA ratio, PSA-TZ, PSAD, and total PSA, respectively. The ANN used here reduced unnecessary biopsies on repeat biopsy by 68%. The AUC for the ANN was 83% and was significantly greater (P < 0.001) than the AUCs for the f/t PSA ratio, PSA-TZ, PSAD, and the total PSA, respectively. The ANN used here reduced unnecessary biopsies on repeat biopsy by 68%. The AUC for the ANN was 83% and was significantly greater (P < 0.001) than the AUCs for the f/t PSA ratio, PSA-TZ, PSAD, and total PSA, respectively. The ANN used here reduced unnecessary biopsies on repeat biopsy by 68%. The AUC for the ANN was 83% and was significantly greater (P < 0.001) than the AUCs for the f/t PSA ratio, PSA-TZ, PSAD, and total PSA, respectively.

**COMMENT**

ANNs are computational methods that perform multifactorial analysis and were inspired by networks of biologic neurons. Like neural networks, ANNs also contain layers of simple points (nodes) of data that interact through carefully weighted connection lines. The weight-balance of these lines is accomplished by a training session of input data, to be used by the network as the means of adjusting its interconnections (Fig. 1). In the study by Babaian et al., a neural network-derived algorithm was developed using a retrospective data set that included the data from 151 men with PSA values from 2.5 to 4.0 ng/mL, who underwent an 11-core, multisided, directed biopsy. The ANN used variables such as age, total PSA, prostatic acid phosphatase, creatinine kinase, and f/t PSA ratio. It consisted of three individually trained networks that were developed with data from retrospective studies from three institutions. The ANN was significantly better in terms of specificity at 92% sensitivity. The AUC for the receiver operating characteristic, however, did not show any better results compared with the single parameters.

Recently, Djavan et al. developed two ANNs for patients referred for early detection of PCa with a total PSA range of 2.5 to 4 ng/mL and 4 to 10 ng/mL for an initial biopsy. They compared the ANNs with conventional statistical analysis (univariate and multivariate) of standard PSA parameters. The f/t PSA ratio, PSA-TZ, PSA velocity, free PSA, TZ volume, total PSA, and PSAD were selected as inputs for the ANN for PSA from 4 to 10 ng/mL, and the PSA-TZ, f/t PSA ratio, PSAD, total prostate volume were used for the ANN using PSA from 2.4 to 4 ng/mL. The specificity at 95% sensitivity and AUC for PCa detection on initial biopsy in the PSA range of 4 to 10 ng/mL was 67% and 91.3%, 60% and 90%, 40% and 81%, 34% and 77%, 30% and 73%, and 4% and 61% for the ANN, logistic regression analysis, f/t PSA ratio, TZ volume, PSA-TZ, and total PSA, respectively.

Similar results were reported by Finne et al. using an ANN and a logistic regression model compared with standard PSA-related parameters. At 95% sensitivity, the ANN, logistic regression model, and f/t PSA had a specificity of 33%, 24%, and 19%, respectively. At 80% to 99% sensitivity levels, the accuracy of the ANN and the logistic regression analysis was significantly greater than that of the f/t PSA ratio. At 89% to 99% sensitivity, the accuracy of the ANN was better than that of the logistic regression analysis (P < 0.001).

In addition, a multicenter study from Innsbruck and Baltimore showed that ANNs improve the specificity compared with standard cutoff levels by between 150% and 200% and were not affected by the presence of benign prostatic hyperplasia or prostatitis in a screening population of 3474 men.

Of the parameters used to decide whether to perform a repeat biopsy, the f/t PSA ratio was the most promising in larger multicenter studies, with a specificity of 20% to 33.5% at 95% sensitivity. A study by Stephan et al. evaluated the diagnostic usefulness of the f/t PSA ratio alone or in combination with an ANN in PSA ranges from 2 to 4 ng/mL, 4.1 to 10 ng/mL, and 10.1 to 20 ng/mL. At 90% sensitivity, the ANN, f/t PSA ratio, and total PSA showed a 63%, 38%, and 23% specificity for PSA levels from 2 to 4 ng/mL, a 57%, 37%, and 19% specificity for PSA levels from 4.1 to 10 ng/mL, and a 46%, 42%, and 14% specificity for PSA levels from 10.1 to 20 ng/mL, respectively. Later, the investigators confirmed the advantage of the ANN model compared with the f/t PSA ratio in a multicenter study of 1188 patients.

Recently, Remzi et al. showed that the total prostate volume and TZ volume alone were able to spare 7.1% and 10% of repeat biopsies using a cut-

**TABLE II. Multivariate logistic regression analysis for repeat prostate biopsy**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Chi-Square</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.093</td>
<td>10.3</td>
<td>0.0013</td>
</tr>
<tr>
<td>PSA-TZ (ng/mL/mL)</td>
<td>2.277</td>
<td>15.6</td>
<td>0.001</td>
</tr>
<tr>
<td>f/t PSA (%)</td>
<td>0.083</td>
<td>41.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Whole model test</td>
<td></td>
<td>90.5</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

(r = 0.215)

Abbreviations as in Table I.
off for total prostate volume of less than 20 mL and more than 80 mL and for the TZ volume of less than 9 mL and more than 41 mL. Additionally, Djavan et al.\textsuperscript{13} reviewed the usefulness of markers in deciding on whom, how, and when a repeat biopsy should be performed. The f/t PSA ratio and PSA-TZ were the best predictors for the outcome of repeat biopsies with a cutoff of 30% and 0.26 ng/mL/mL, respectively.

The ANN presented in this study is to our knowledge the first for decision making of repeat prostate biopsies, with a 68% specificity at 95% sensitivity, allowing the reduction of more than two thirds of repeat biopsies. All standard parameters failed to reduce repeat biopsies by such an amount so far. Clearly, the ANN could have been extended further by adding newer markers such as complexed PSA or human kallikrein 2 (hk2), or even other clinical parameters (ie, body mass index, serum creatinine). Currently, an improved ANN is under investigation and will be compared with the ANN presented here.

**CONCLUSIONS**

The ANN presented here has a high accuracy for the early detection of PCa on repeat biopsy and can reduce unnecessary biopsies by 68%. In addition, this ANN will allow more individual counseling of patients with a negative initial biopsy and thus reduce patients’ anxiety, while reducing unnecessary diagnostic and invasive steps. The additional inclusion of newer serum markers and clinical parameters may improve the performance of this ANN.

**REFERENCES**


**TABLE III. Results at 95% and (90%) sensitivity for 820 patients who underwent repeat prostate biopsy**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cutoff</th>
<th>Specificity (%)</th>
<th>AUC* (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>0.42 (0.44)</td>
<td>68 (78)</td>
<td>83</td>
</tr>
<tr>
<td>Multivariate logistic regression analysis</td>
<td>0.37 (0.41)</td>
<td>54 (63)</td>
<td>79</td>
</tr>
<tr>
<td>f/t PSA (%)</td>
<td>38 (50)</td>
<td>33.5 (50)</td>
<td>74.5</td>
</tr>
<tr>
<td>PSA-TZ (ng/mL/mL)</td>
<td>0.19 (0.23)</td>
<td>21.4 (26.7)</td>
<td>69.1</td>
</tr>
<tr>
<td>PSAD (ng/mL)</td>
<td>0.09 (0.11)</td>
<td>14.7 (22)</td>
<td>61.8</td>
</tr>
<tr>
<td>Total PSA (ng/mL)</td>
<td>4 (4.5)</td>
<td>8.3 (11.9)</td>
<td>60.5</td>
</tr>
</tbody>
</table>

K: AUC = area under receiver operating characteristic curve; ANN = artificial neural network; ROC = receiver operating characteristic; other abbreviations as in Table I.

* Comparison of AUC for ROC; ANN vs. others, \( P < 0.01 \) (Mc Nemar test modified by Bonferroni-Holm).


