Classification of incidental carcinoma of the prostate using learning vector quantization and support vector machines

Torsten Mattfeldt a,*, Danilo Trijic a, Hans-Werner Gottfried b and Hans A. Kestler c,d

a Department of Pathology, University of Ulm, Ulm, Germany
b Department of Urology, University of Ulm, Ulm, Germany
c Department of Neuroinformatics, University of Ulm, Ulm, Germany
d Department of Internal Medicine I, University of Ulm, Ulm, Germany

Received 7 July 2003
Accepted 29 November 2003

Abstract. The subclassification of incidental prostatic carcinoma into the categories T1a and T1b is of major prognostic and therapeutic relevance. In this paper an attempt was made to find out which properties mainly predispose to these two tumor categories, and whether it is possible to predict the category from a battery of clinical and histopathological variables using newer methods of multivariate data analysis. The incidental prostatic carcinomas of the decade 1990–99 diagnosed at our department were reexamined. Besides acquisition of routine clinical and pathological data, the tumours were scored by immunohistochemistry for proliferative activity and p53-overexpression. Tumour vascularization (angiogenesis) and epithelial texture were investigated by quantitative stereology. Learning vector quantization (LVQ) and support vector machines (SVM) were used for the purpose of prediction of tumour category from a set of 10 input variables (age, Gleason score, preoperative PSA value, immunohistochemical scores for proliferation and p53-overexpression, 3 stereological parameters of angiogenesis, 2 stereological parameters of epithelial texture). In a stepwise logistic regression analysis with the tumour categories T1a and T1b as dependent variables, only the Gleason score and the volume fraction of epithelial cells proved to be significant as independent predictor variables of the tumour category. Using LVQ and SVM with the information from all 10 input variables, more than 80 of the cases could be correctly predicted as T1a or T1b category with specificity, sensitivity, negative and positive predictive value from 74–92%. Using only the two significant input variables Gleason score and epithelial volume fraction, the accuracy of prediction was not worse. Thus, descriptive and quantitative texture parameters of tumour cells are of major importance for the extent of propagation in the prostate gland in incidental prostatic adenocarcinomas. Classical statistical tools and neuronal approaches led to consistent conclusions.

Keywords: Artificial neural networks, bioinformatics, classification, immunohistochemistry, incidental carcinoma, learning vector quantization, logistic regression, pathology, pattern recognition, prostatic cancer, stereology, support vector machine

1. Introduction

Incidental prostatic cancer is prostatic adenocarcinoma discovered by chance. In contrast to the categories pT2–pT4, it is not a pathologically defined tumour stage but includes a heterogeneous population of cancers with different extent of invasion, and is only characterized by the mode of its clinical presentation. In most cases such tumours are found at the occasion of transurethral resections (TUR) or surgical adenomectomies of the prostate after a diagnosis of benign prostatic hyperplasia (BPH) [2]. Moreover, incidental carcinomas are found in radical cystoprostatectomy specimens removed because of bladder cancer [29]. In the course of time, many clinical data have been accumulated for the T1 tumour [2,5]. However, the number of studies in which modern methods of structural cell biology, quantitative stereology and genet-

*Corresponding author: Prof. Dr. T. Mattfeldt, Department of Pathology, Oberer Eselsberg M23, D-89081 Ulm, Germany. Fax: +49 731 58738; E-mail: torsten.mattfeldt@medizin.uni-ulm.de.

1570-5870/04/$17.00 © 2004 – IOS Press and the authors. All rights reserved
ics have been applied to this collective is scarce. Also these techniques have rarely been applied in a multivariate approach in order to look for the relative importance of the factors. Recently, the first findings on T1 prostate cancer using comparative genomic hybridization (CGH) have been reported [52].

We attempted to study all cases of incidental prostatic cancer that were diagnosed in our department in one decade (1990–1999) as a retrospective clinical and histopathological investigation. Additionally, immunohistochemical and stereological studies on the tumor texture, proliferation and angiogenesis were performed, and the data were studied statistically by a stepwise logistic regression analysis. Specifically, it was attempted to find out which variables mainly predispose to the subcategories T1a and T1b, as T1b is known to be associated with worse prognosis, and may lead to more aggressive therapies such as radical prostatectomy or radiation especially in younger patients [10,40]. We follow the definitions of T1a and T1b indicating $\leq 5\%$ and $>5\%$ area fraction of resected prostate occupied by prostate cancer tissue according to the current tumour classification of the UICC [36].

For the purpose of classification (pattern recognition) we used learning vector quantization (LVQ), which has been applied for classification of prostatic cancer before [23–25]. Additionally support vector machines (SVM) were used for classification [3,4,17,45].

2. Materials and methods

2.1. Pathology

We proceeded from the total number of prostate specimens of the Department of Urology of the University of Ulm, which had been examined histopathologically at the Department of Pathology at the University of Ulm in the years 1990–1999. This material included prostate tissue, that had been removed as TUR because of BPH, and prostate specimens surgically operated because of BPH (adenomectomies). Cystoprostatectomies with incidental prostatic carcinomas were not included into the study. On the whole, we found 59 incidental carcinomas in the TUR specimens and 7 incidental carcinomas in the adenomectomies, when strict criteria of inclusion were applied. Especially, the original documents of each case were studied again, and a search was performed for previous histological findings of that patient. A case was only included when there was no prior diagnosis of PCA by biopsy or antecedent TUR, and when in the written clinical diagnosis only benign changes of the prostate gland (adenoma, BPH) were mentioned. From all cases with incidental carcinomas, the preoperative PSA-value in the serum estimated according to the Hybritech kit (Hybritech Inc; Beckman Coulter Fullerton, CA; see [23]) was evaluated.

2.2. Histology and immunohistochemistry

The Gleason score was reevaluated microscopically for all cases by the first author, without knowledge of the other variables. Additionally, immunohistochemical stains were performed from the technically best suitable paraffin block of each case. We used monoclonal antibodies against the human Ki-67 antigen (Mib-1) and a monoclonal antibody against human p53-protein. For the study of vascularization an antibody against factor VIII-related antigen (FVIII-ra, von Willebrand factor) was used [43]. All antibodies were obtained from DAKO, Glostrup, Denmark. The reactions were judged semiquantitatively. In the evaluation of the Mib-1 reaction, a subjective estimate of the fraction of positively labelled nucleus profiles was obtained in intervals of 10% (0, $>0–10\%$, $>10–20\%$, ..., $>90–100\%$; $>0.1–0.2$, ..., $>0.90–1.00$). The p53-reaction was judged in the same manner.

2.3. Image analysis and quantitative stereology

Stereology means that data obtained from sections are extrapolated to the true three-dimensional properties of a structure using unbiased mathematical methods [8]. Stereological principles were used to study the epithelial texture and the capillary vascularization. For both evaluations, paraffin sections with an immunohistochemical stain using an antibody against FVIII-ra were used. Such sections show simultaneously the epithelial cells, the stroma, the glandular lumina and the capillaries (Fig. 1). To characterize the epithelial tissue texture, systematic series of visual fields containing tumour tissue were evaluated with a random startpoint [21,22,24]. Using the well-known fundamental stereological equations:

\[
V_V = A_A \quad (1)
\]

\[
S_V = (4/\pi)B_A \quad (2)
\]

we estimated the volume fraction $V_V$ of epithelium from the area fraction of epithelial tissue per unit reference area, $A_A$ of epithelium, whereas the surface area of
epithelium per unit tissue volume, $S_V(\text{epi})$, was estimated from the mean boundary length of epithelial tissue per unit reference area, $B_A(\text{epi})$ (see, e.g., [8]). Note that $V_V(\text{epi})$ is different from the fraction of tumor tissue in the whole specimen that determines the categories T1a and T1b: $V_V(\text{epi})$ refers to the volume fraction of tumour cells within the tumour, whereas the latter is the fraction of tumour tissue in the total sample. Image analysis was performed by means of PC programs after interactively segmenting the images with the Kontron system IBAS 2000 using a CCD camera connected to the microscope and converting the images to the TIF-format [21,22,24]. The final magnification corresponded to a width of 0.4 mm of the quadratic visual field at the scale of the tissue (Fig. 1).

To characterize tumour vascularization, the volume fraction and the surface area of vessels per unit tissue volume, $V_V(\text{cap})$ and $S_V(\text{cap})$, were estimated by using Eqs (1), (2) as above. A further important parameter in the context of vascularization is the length of the capillary network per unit tissue volume, $L_V(\text{cap})$ [18, 20]. It was estimated from the fundamental stereological equation:

$$L_V = 2Q_A$$

where $Q_A$ denotes the mean number of vessel profiles per unit area, a planar parameter sometimes denoted as capillary density or microvessel density [8,18, 20,50,51]. A case was completed with the evaluation of that field in which the 200th capillary profile was found.

2.4. Logistic regression analysis

Logistic regression is a mathematical modeling approach that can be used to describe the relationship of several independent variables to a dichotomous (binary) variable $f(z)$. Such a variable assumes only two values, here it is the tumor category, i.e. T1a or T1b, which was coded as 0 and 1. We consider now the linear function $z = f(X)$ of a $k$-dimensional input vector $X = (x_1, \ldots, x_k)$, where the $x_i$ are the $k$ independent variables: $z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k = \alpha + \sum_{i=1}^{k} \beta_i x_i$. The function $z = f(X)$ can be interpreted as an index of combined risk factors. The logistic model can be written as

$$f(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(\alpha + \sum_{i=1}^{k} \beta_i x_i)}}$$
3. Methods of data classification

3.1. Learning vector quantization (LVQ)

LVQ is an artificial neural network with a supervised learning rule [12,14]. It can be seen as the supervised counterpart to self-organizing maps (SOM) (for references on SOMs see [9,13–15,26,27,38]), which function according to an unsupervised learning rule. While SOMs are fed only with input variables, e.g., gene expression data or CGH data, LVQ obtains not only the input vectors but also an output vector. If this is a single variable with two values only, e.g., 0 and 1 as here, we enter the domain of binary pattern recognition. When one has decided on a set of \( k \) variables to evaluate, the data of a single case can be represented as a vector or point in the corresponding \( k \)-dimensional input space. Using LVQ, further vectors are randomly placed into this space, which are moved until the distance of these vectors to the input vectors is minimized [12,14]. In this manner, the algorithm attempts to spread the new vectors as uniformly as possible across the input vectors, very similar as in SOMs. The new vectors are called codebook vectors and are marked as 0 or 1 for the classes. The input vectors are classified from the system according to the mark of their nearest neighbour among the codebook vectors. On the whole, we classified the data with LVQ networks with 1–20 codebook vectors, and for each number of codebook vectors 46 different combinations of LVQ parameters were applied. More details about LVQ can be found in [12,14] and in our previous publications [23–25]. LVQ can be downloaded with documentation as academic software by internet under http://www.cis.hut.fi/research/som-research/nmr-c-programs.shtml as a set of source files for Unix or DOS, and as binary (executable) files for Windows. For our study the source files were compiled and executed under Linux.

3.2. Support vector machines

Support vector machines (SVM) were developed by Vapnik since the end of the seventies and are a further well-known supervised method for data classification [3,4,45]. In the simplest case, a SVM in binary pattern recognition mode has the task to classify a separable dataset of vectors by a linear decision function (hyperplane) into two halfspaces. Each input vector, which is marked in two classes as \(-1\) or \(+1\), is classified according to the halfspace where its coordinates are located. The hyperplane is constructed by the SVM in such a manner that the width of the margin of the separating hyperplane to the nearest neighbours among the input vectors of both classes becomes a minimum. Only these input vectors have an influence on the definite equation of the hyperplane, and the vectors of this subset of input vectors are denoted as “support vectors”. For more complicated tasks, such as classifying overlapping vector sets from two classes, or in case of nonlinearity of the data, variations of the aforementioned fundamental algorithm are available. Instead of complete separation an incomplete separation can be accepted, where overlapping is allowed but penalized. Thus, it is possible to select a kernel function \( K(x,y) \), and parameters which determine the penalty function for misclassification. We have used 5 types of kernels: the function \( K(x,y) = xy \), i.e. the simple dot product of the vectors (= inner or scalar product of vectors), radial basis functions with \( K(x,y) = \exp(-\gamma|x-y|^2) \), the regularized Fourier kernel \( K(x,y) = \prod_{k=1}^{d} K_k(x^k,y^k) \), and the polynomial kernels \( K(x,y) = (xy + 1)^d \) (Vapnik polynomial, simple polynomial) and \( K(x,y) = [(xy/a) + b]^d \) (full polynomial) [34]. This is only a very limited selection of the full range of kernel functions available for SVM, see [34]. As implementation we used the program package _svm_ under Linux, which can be downloaded by internet under http://svm.dcs.rhbnc.ac.uk/dist/index.shtml as binary program or as source code [34].
3.3. Evaluation with the leave-one-out method

In the two data sets outlined below, output data were predicted from input data on an individual basis by the leave-one-out principle (synonyms: jack-knife, round-robin) [41,45]. This means that the total group of \( n \) cases is partitioned into a subgroup of \( n - 1 \) cases (the training cases) and another subgroup which consists only of the single remaining case (the test case). In the training phase, the network “learns” to estimate the output variable from the input variables within the training group. In the test phase thereafter, the output variable of the test case is estimated from the input variables of the test case making use of the information learnt previously from the training group (here: 65 cases). This strategy simulates a confrontation of the network with a new case, and by this manner tests its ability to generalize. Thus the leave-one-out approach simulates the clinical situation where the diagnosis of a new patient is made on the basis of the doctor’s previous experience with similar cases. The prediction is repeated cyclically for every patient as test case with the complementary set of cases serving as its training group. For each cyclic evaluation, the following statistics were computed: accuracy – overall percentage of correctly classified cases, sensitivity – ratio of number of cases correctly classified as positive to the total number of positive cases, specificity – ratio of number of cases correctly classified as negative to total number of negative cases, positive predictive value – ratio of number of cases classified as positive, negative predictive value – ratio of number of cases classified as negative, sensitivity – ratio of number of cases correctly classified as positive to total number of cases classified as positive, negative predictive value – ratio of number of cases correctly classified as negative to total number of cases classified as negative.

It is well known that leave-one-out error estimation can be performed in two fundamentally different ways. The first approach is to decide on feature selection (i.e. the choice of input variables for classification) on the complete data set beforehand, and thereafter perform the leave-one-out cross-validation. The second approach is to perform feature selection as well as optimizing the classifier parameters for each cycle of the leave-one-out simultaneously. The latter method results in a lower bias, but a higher variance of the error estimate [35]. Here the former of the two aforementioned approaches was selected, using a stepwise logistic regression procedure for feature selection; see also our previous papers [17,23–25].

4. Results

4.1. Group comparisons

In Table 1 we find the mean values and standard deviations of the variables in the T1a and T1b carci-

| Variable          | T1a tumours | T1b tumours | Level of significance |
|-------------------|-------------|-------------|-----------------------|
| Age (years)       | 70.23       | 72.25       | N.S.                  |
| Mib-1 (%)         | 2.40        | 3.47        | N.S.                  |
| p53 (%)           | 0.39        | 2.91        | N.S.                  |
| PSA (ng/ml)       | 8.82        | 24.60       | \( p < 0.05 \)        |
| \( L_{V}(\text{cap}) \) [\text{mm}^{3}]/\text{unit tissue volume} | 147.63 | 182.88 | N.S. |
| \( S_{V}(\text{cap}) \) [\text{mm}^{2}]/\text{unit tissue volume} | 7.88 | 9.73 | N.S. |
| \( V_{V}(\text{cap}) \) | 0.02 | 0.03 | N.S. |
| \( S_{V}(\text{epi}) \) [\text{mm}^{2}]/\text{unit tissue volume} | 33.93 | 42.47 | \( p < 0.01 \) |
| \( V_{V}(\text{epi}) \) | 0.28 | 0.35 | \( p < 0.05 \) |
| Gleason score     | 4.10        | 6.18        | \( p < 0.0001 \)     |

Mean values with standard deviations are presented for the T1a and T1b group. The right column shows the results of group comparisons using a Wilcoxon rank sum test. The difference between the Gleason scores of T1a and T1b tumours is highly significant. Also the preoperative PSA value, and volume and surface area of epithelial cells per unit tissue volume were significantly higher in the group of T1b-tumours. Abbreviations: Mib-1, p53: semiquantitative score of Mib-1 and p53 immunohistochemistry, GS: Gleason score, \( V_{V}(\text{cap}) \): volume fraction of capillaries, \( L_{V}(\text{cap}) \): length of capillaries per unit tissue volume, \( S_{V}(\text{cap}) \): surface area of capillaries per unit tissue volume, \( V_{V}(\text{epi}) \): volume fraction of epithelial cells per unit tissue volume, \( S_{V}(\text{epi}) \): surface area of epithelial cells per unit tissue volume.
nomal groups. For group comparisons, the nonparametric Wilcoxon rank sum test (Mann–Whitney U test) was used [1]. There were 39 cases in category T1a and 27 cases in category T1b. The Gleason score and the preoperative PSA-values were significantly higher in the T1b-group. With respect to the quantitative stereological findings of tumour texture, we found significantly higher estimates of epithelial volume fraction and epithelial surface area per unit volume in the T1b group. There was a trend towards a denser tumour vascularization in terms of length, surface area and volume of capillaries per unit tissue volume in the T1b group, but this trend was statistically not significant at the 5%-level. The same statement holds for nuclear Mib-1 expression and p53-overexpression.

4.2. Logistic regression

Applying the aforementioned stepwise algorithm for logistic regression to our data, we used age, Mib-1, p53, the PSA value, $L_V$(cap), $S_V$(cap), $V_V$(cap), $S_V$(epi), $V_V$(epi) and the Gleason score as $k = 10$ independent variables $x_i$, and the tumor categories T1a and T1b as dependent variable $f(z) \in [0, 1]$. In the stepwise logistic regression procedure, only the two variables Gleason score and epithelial volume fraction met the 5% significance level for entry into the model. The corresponding values were $\chi^2 = 9.05$ for Gleason score ($p < 0.01$) and $\chi^2 = 7.05$ for the volume fraction of epithelial cells ($p < 0.01$). The other variables were eliminated from the model as non-contributory.

4.3. Classification of cases by LVQ and SVM

In this part of the study, four data sets of the same 66 cases with different sets of input variables were examined. The single output variable in all 4 settings was the tumor category, which was coded as 0 or 1 for LVQ, and as −1 or 1 for SVM, respectively, corresponding to T1a and T1b. For the purpose of classification, missing values were estimated from the mean values of that variable of the cases with available data [42], and all input values were scaled (normalized) to the interval [0, 1]. The first dataset consisted of all 10 input variables (age, Mib-1, p53, PSA, Gleason score, $L_V$(cap), $S_V$(cap), $V_V$(cap), $S_V$(epi) and $V_V$(epi)) and the output variable T1a/T1b. For the second data set only the two variables Gleason score and epithelial volume fraction were considered. For the third and fourth dataset, merely one significant input variable was included: only Gleason score or only epithelial volume fraction, respectively. Accuracy, sensitivity, specificity, positive and negative predictive values of the classification were tested by cross-validation according to the leave-one-out method [23–25,41,45].

Table 2 shows the results for these 4 data sets obtained by LVQ and SVM in comparison. In the table, only the best results obtained by the two algorithms have been shown. These are based on an optimization of the algorithm parameters, which is performed partially on the basis of experience but also to a large extent by trial and error. It is fundamentally impossible to

| Table 2 | Classifications |
|----------|-----------------|
| Input variables | Algorithm | Accuracy | Sensitivity | Specificity | PPV | NPV | Algorithm properties |
| Def. Number | LVQ | SVM | LVQ | SVM | LVQ | SVM | LVQ | SVM |
| All 10 | 0.8485 | 0.8333 | 0.8519 | 0.7407 | 0.8462 | 0.8571 | 0.7931 | 0.8333 | 0.8919 | 0.8333 | 7 CV |
| GS 2 | 0.8485 | 0.8636 | 0.8519 | 0.8888 | 0.8462 | 0.8462 | 0.8000 | 0.8919 | 0.9167 | RF |
| $V_V$(epi) | 0.7424 | 0.7037 | 0.7692 | 0.6786 | 0.7076 | 0.7333 | 0.7895 | 0.7895 | 0.7895 | VP, RBF |
| GS 1 | 0.7273 | 0.5556 | 0.7556 | 0.5556 | 0.8462 | 0.7143 | 0.7143 | 0.7333 | 0.7333 | RBF |

Classification (prediction) of category T1a versus T1b on the basis of all 10 input variables, and on reduced data sets with 2 or only 1 input variables. For the complete data set, the 10 variables age, Mib-1, p53, Gleason score, $V_V$(cap), $L_V$(cap), $V_V$(epi), $S_V$(epi), PSA were used. For the other data sets, the input variables are indicated in the left column. LVQ: learning vector quantization, SVM: support vector machine, NPV: negative predictive value, PPV: positive predictive value. In the right column more detailed specifications of the algorithms are given: number of codebook vectors (CV) for LVQ; type of kernel function for SVM (RF: regularized Fourier kernel, VP: Vapnik polynomial, RBF: radial basis function). For the data set based on GS only, identical results were obtained by SVMs with VP and RBF kernels. Note identical predictions by LVQ and SVM on the basis of the one-dimensional datasets in the two lower rows.
explore all theoretically conceivable parameter combinations, because various algorithm parameters are numerical quantities which possess an infinite range of variation. Dependent on the algorithms and the set of variables, the accuracy ranged between 73% and 86% when new test cases were presented to the system after training. The number of codebook vectors necessary for the complete data set and for the reduced data set with two input variables with a regularized Fourier kernel. For the reduced data sets with only one input variable, the best results were obtained with polynomials and radial basis functions as kernels. Note that neither for LVQ nor for SVM, the accuracy of prediction could be improved by rising the number of input variables from 2 to the full set of 10 variables. In general, the classification accuracies by SVM and LVQ were very similar, sometimes even identical results were obtained (Table 2). As an increase of the number of variables beyond 2 did not augment the precision, the input variables in addition to GS and $V_V$ (epi) must be considered as non-contributory also from the viewpoint of a purely neural approach to the data. On the other hand, a further reduction to a single input variable (only GS, only $V_V$ (epi)) leads to a loss of accuracy. This shows that GS and $V_V$ (epi) possess some independent information. Summarizing, a correct prediction of category T1a versus T1b was feasible in a high percentage of incidental prostatic cancers (>80%) on the basis of all 10 variables or using only the 2 significant variables. These predictions were obtained by LVQ as well as by SVM and were accompanied by similar results for sensitivity, specificity, positive and negative predictive value.

5. Discussion

The vascular supply and proliferation of prostatic carcinomas have been investigated in various studies on prostatic cancer in advanced tumour stages [28, 46,50,51]. In two previous studies capillary vascularization was also studied in incidental prostatic carcinomas [47,48]. Also in other studies on T1 prostatic carcinoma, overexpression of p53-protein and the expression of proliferation-associated antigens such as Ki-67 (Mib-1) have been explored [6,44]. We found a trend towards an increased capillary vascularization in T1b tumours as compared to T1a tumours. In general, higher pathological stage, higher histological grade and metastasis have been shown to be correlated with increased tumour vascularization; our results are compatible with this view. While some previous data on these aspects of the quantitative pathology of incidental prostatic cancer are available, the stereological study of epithelial texture has only rarely been applied in general [17,19,23,24]. The present data for volume fraction and surface area of epithelial cells per unit tissue volume are generally in the same size range for prostatic carcinomas as in the previous communications [23,24]. The study of epithelial texture is laborious as long as the images are interactively segmented. However, in principle automatic image analysis after contrast enhancement of epithelial structures by appropriate immunohistochemical stains (e.g., with antibodies against cytokeratin or PSA) could strongly increase the efficiency of this method, which is presently still restricted to scientific studies.

It was attempted to characterize incidental prostatic carcinomas by quantitative and multivariate techniques. Stepwise logistic regression was used to find out which of these variables remain significant when their interactions are taken into account. Thus, an approach of classical statistics was applied for the purpose of feature selection for the neural paradigms. It could be shown that the dependent variable, T1a vs. T1b, could be explained largely by two independent significant variables, the Gleason score and the epithelial volume fraction of the tumour tissue. Clearly this observation is consistent with the view that the Gleason score is correlated to, but not fully determined by the epithelial volume fraction. A look at the well-known schematic diagrams of the Gleason score [7] shows that in general the volume fraction of the epithelia rises with increasing Gleason score, but not monotonously and in addition to volume fraction changes there are also complex pattern alterations from tubular to cribriform to solid architecture (Fig. 2). This aspect is not entirely represented by first-order parameters such as volume and surface area, but can be quantified only by special techniques of stereology and stochastic geometry, e.g., pattern analysis on the basis of second-order properties or by comparison of empirical structures to reference models of random set theory [17,19,22,24,30,37]. Both Gleason score and epithelial texture stereology have the advantage that they can be obtained from every specimen also if these should be uninformative for immunohistochemical stains. They are also available independently from any clinical data. Presently the Gleason score remains a highly efficient and indispensable tool for grading epithelial texture.
Fig. 2. Central areas of a reference image for Gleason grading were digitized and segmented (quoted from [7]). Here the components are epithelial cells (white), glandular lumina (gray), and stroma (black). The numbers indicate the Gleason grade of individual tumour patterns; the Gleason score is the sum of the grades of the two major patterns (e.g., $3 + 4 = 7$). The patterns related to the grades differ by volume fraction and surface area of tumour cells per unit volume, but in addition there are complex shape changes which transcend quantification by simple first-order parameters.

In the present investigation newer tools of data analysis were applied for the first time to cases of incidental prostatic cancer. There exists already a number of studies with artificial neural networks on prostatic carcinoma, in which different predictions were made, e.g., preoperative prediction of tumour stage from clinical data and biopsy findings, prediction of tumour relapse after prostatectomy, and other study designs (see [17,23,25,39,49] and references therein). In these investigations, the object of the study was clinically manifest carcinoma, and in the large majority of studies multilayer feedforward networks with backpropagation (multilayer perceptrons, MLP) were used. In the present study, the neuronal paradigms LVQ and SVM were applied for the first time to prostatic cancer data in comparison. Previous data of our group have shown that LVQ may be advantageous with respect to accuracy and other characteristics of the quality of prediction in comparison to MLPs [23–25]. Superior results by LVQ in classification of urological and other tumours have also been reported by other groups [16, 31,32]. On the whole, the results obtained by LVQ and SVM were similar in this study. One cannot exclude that even better results could have been obtained by SVM if still more variations of the kernel function and other SVM parameters had been performed; here we restricted ourselves to 5 basic types of kernel functions in SVM [34]. We suggest that alternative techniques of
data analysis such as LVQ and SVM, which are also often faster and no more difficult to use than MLPs, should be tried in the general context of prediction in clinical studies. Clearly SVMs can give additional information by identifying a subset of the cases as support vectors, but in the present context we have not exploited this option any further. The reader might wonder why the availability of more information (here: 10 input variables as compared to only 2 input variables) did not improve the accuracy of the classification. To understand this phenomenon, one must take into account that the networks learn on preclassified training data sets and thereafter must generalize to new, unknown test cases. For the ability to generalize, it is necessary to find a good compromise between too little learning and too much learning; in the latter situation the system “learns the patterns by heart”, but extrapolates less well to new patterns. In the present situation we see this effect (albeit at a very low level) for SVM when switching from 2 to 10 variables, which implies a considerable increase of model complexity. A similar behaviour is well-known as “overtraining” (overfitting), e.g., for multilayer perceptrons [23,24], where greatly enhanced numbers of training epochs may rather lead to worse than improved results.

With respect to practical applications, we have restricted our study to techniques applicable to diagnostic routine material. It would also have been possible to include genetic data, e.g., from CGH-studies [26,27,52], but this technique is more laborious and will possibly remain restricted to research laboratories. Anyway, the conventional distinction between T1a and T1b rests on an estimation whether the tumour cells occupy 5% or less, or more than 5% of the resected prostatic tissue. Such a value can be highly biased by the localization where the tissue has been removed. In contrast, properties such as histological grade, texture parameters or estimates of proliferation and angiogenesis are more homogeneous and should depend less on the site of the removed histological specimen. Clearly our cases were not reclassified after the study. The diagnosis T1a or T1b case was kept unchanged on the basis of the volume fraction of tumour per total tissue as defined according to the UICC. Nevertheless it is plausible to conceive hypothetical cases where a computer-based classification of a case would be desirable, e.g., because a reliable determination of the tumor fraction per tissue is not feasible due to paucity of resected material. Also it is of general pathobiological interest which factors enable a clinically silent prostatic neoplasm to occupy more than 5% of the gland. It is planned to investigate the predictive value of quantitative histological texture variables in further prospective studies. If our best network (SVM with regularized Fourier kernel and 2 input variables) were used for this purpose, this would mean that 86% of the cases would be correctly classified into two classes, which could be defined as “group with low risk of T1b” and “group with high risk of T1b”. In the first group, the risk of having T1b cancer would be only 8.3%, whereas in the second group, this risk would be 80%. This observation confirms that artificial neural networks can be useful for the identification of individual patients in low and high risk categories [23,25,26].

Acknowledgements

This study was supported by the IZKF (interdisciplinary centre for clinical research) of the University of Ulm, grant number C8, and by the Stifterverband für die Deutsche Wissenschaft (H.A.K.). Thanks are due to Katrin Schierle for excellent technical help with the illustrations.

References

[1] J. Bortz, G. Lienert and K. Boehnke, Verteilungsfreie Methoden in der Biostatistik, Springer, Berlin, 2001.
[2] D.G. Bostwick, The pathology of incidental carcinoma, Cancer Surveys 23 (1995), 7–18.
[3] M.P.S. Brown, W.N. Grundy, D. Lin, N. Cristianini, C.W. Sugnet, T.S. Furey, M. Ares and D. Haussler, Knowledge-based analysis of microarray gene expression data by using support vector machines, Proc. Natl. Acad. Sci. USA 97 (2000), 262–267.
[4] J.C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining Knowl. Discov. 2 (1998), 121–167.
[5] L. Cheng, E.J. Bergstrath, B.G. Scherer, R.M. Neumann, M.L. Blute, H. Zancke and D.G. Bostwick, Predictors of cancer progression in T1a prostate adenocarcinoma, Cancer 85 (1999), 1300–1304.
[6] M.R. Feneley, M. Young, C. Chinayama, R.S. Kirby and M.C. Parkinson, Ki-67 expression in early prostate cancer and associated pathological lesions, J. Clin. Pathol. 49 (1996), 741–748.
[7] D.F. Gleason, Histologic grading of prostate cancer: a perspective, Hum. Pathol. 23 (1992), 273–279.
[8] C.V. Howard and M.G. Reed, Unbiased Stereology. Three-Dimensional Measurement in Microscopy, Bios Scientific Publishers, Oxford, 1998.
[9] A.K. Jain and R.C. Dubes, Algorithms for Clustering Data, Prentice Hall, Englewood Cliffs, 1988.
T. Mattfeldt et al. / Classification of incidental prostatic cancer

[10] D. Keyser, P. Kupelian, C.D. Zippe, H.S. Levin and E.A. Klein, Stage T1–2 prostate cancer with pretreatment prostate-specific antigen level < or = 10 ng/ml: radiation therapy or surgery?, Int. J. Radiat. Oncol. Biol. Phys. 38 (1997), 723–729.

[11] D.G. Kleinbaum, Logistic Regression: A Self-Learning Text, Springer, New York, 1994.

[12] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen and K. Torkkola, LVQ_PAK: The learning vector quantization program package, Technical Report A30, Helsinki University of Technology, Laboratory of Computer and Information Science, Otaniemi, Finland, 1996.

[13] T. Kohonen, J. Hynninen, J. Kangas and J. Laaksonen, SOM_PAK: The Self-Organizing Map Program Package, Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 1996.

[14] P. Mangiameli, S.K. Chen and D. West, A comparison of SOM neural network and hierarchical clustering, Eur. J. Operat. Res. 93 (1996), 402–417.

[15] C. Markopoulos, P. Karakitsos, E. Botsoli-Stergiou, A. Poulakis, A. Iokim-Liossi, K. Kyrkou and J. Gogas, Application of the learning vector quantizer to the classification of breast lesions, Analyt. Quant. Cytol. Histol. 19 (1997), 453–460.

[16] T. Mattfeldt, Classification of binary spatial textures using stochastic geometry, nonlinear deterministic analysis and artificial neural networks, Int. J. Pattern Recogn. Artif. Intell. 17 (2003), 275–300.

[17] T. Mattfeldt and G. Mall, Estimation of length and surface of anisotropic capillaries, J. Microsc. 135 (1984), 181–190.

[18] T. Mattfeldt and D. Stoyan, Improved estimation of the pair correlation function, J. Microsc. 200 (2000), 158–173.

[19] T. Mattfeldt, G. Mall, H. Gharehbaghi and P. Möller, Estimation of surface area and length with the orientator, J. Microsc. 159 (1990), 301–317.

[20] T. Mattfeldt, H. Frey and C. Rose, Second-order stereology of benign and malignant alterations of the human mammary gland, J. Microsc. 171 (1993), 143–151.

[21] T. Mattfeldt, V. Schmidt, D. Reepschläger, C. Rose and H. Frey, Centred contact density functions for the statistical analysis of random sets, J. Microsc. 183 (1996), 158–169.

[22] T. Mattfeldt, H.A. Kestler, R. Hautmann and W. Gottfried, Prediction of prostatic cancer progression after radical prostatectomy using artificial neural networks: a feasibility study, BJU Int. 84 (1999), 316–323.

[23] T. Mattfeldt, H.W. Gottfried, V. Schmidt and H.A. Kestler, Classification of spatial textures in benign and cancerous glandular tissues by stereology and stochastic geometry using artificial neural networks, J. Microsc. 198 (2000), 143–158.

[24] T. Mattfeldt, H.A. Kestler, R. Hautmann and H.W. Gottfried, Prediction of postoperative prostatic cancer stage on the basis of systematic biopsies using two types of artificial neural networks, Eur. Urol. 39 (2001), 530–537.

[25] T. Mattfeldt, H. Wolter, R. Kemmerling, H.W. Gottfried and H.A. Kestler, Cluster analysis of comparative genomic hybridization (CGH) data using self-organizing maps: application to prostate carcinomas, Anal. Cell. Pathol. 23 (2001), 29–37.

[26] T. Mattfeldt, H. Wolter, D. Trijic, H.W. Gottfried and H.A. Kestler, Chromosomal regions in prostatic carcinomas studied by comparative genomic hybridization, hierarchical cluster analysis and self-organizing feature maps, Anal. Cell. Pathol. 24 (2002), 167–179.

[27] T. Mattfeldt, R. Mazzucchelli, R. Montironi, A. Santinelli, G. Lucarini, A. Pugnaloni and G. Biagini, Vascular endothelial growth factor expression and capillary architecture in high-grade PIN and prostate cancer in untreated and androgen-ablated patients, Prostate 45 (2000), 72–79.

[28] G. Moutzouris, C. Barbatis, D. Plastiras, N. Mertzios, C. Katsofotis, V. Plessovels and C. Theodorou, Incidence and histological findings of unsuspected prostatic adenocarcinoma in radical cystoprostatectomy for transitional cell carcinoma of the bladder, Scand. J. Urol. Nephrol. 33 (1999), 27–30.

[29] J. Ohser and F. Mücklich, Statistical Analysis of Microstructures in Materials Sciences, Wiley, Chichester, 2000.

[30] D. Pantazopoulos, P. Karakitsos, A. Iokim-Liossi, A. Poulakis and K. Dimopoulos, Comparing neural networks in the discrimination of benign from malignant lower urinary tract lesions, Br. J. Urol. 81 (1998), 574–579.

[31] D. Pantazopoulos, P. Karakitsos, A. Poulakis, A. Iokim-Liossi and M.A. Dimopoulos, Static cytometry and neural networks in the discrimination of lower urinary system lesions, Urology 51 (1998), 946–950.

[32] SAS Institute, SAS/STAT User’s Guide, Release 6.03 Edition, SAS Institute Inc., Cary, NC, 1988.

[33] R. Saunders, M.O. Stitson, J. Weston, L. Botton, B. Schölkopf and A. Smola, Support vector machine reference manual, Technical Report, Royal Holloway, University of London, 1998.

[34] J. Weston, M. Sahbi, A. Smola, D. Schölkopf and M. Jordan, Index of the Baire theorem and automatic relevance determination, Proc. Natl. Acad. Sci. USA 96 (1999), 15063–15068.

[35] A. Tewari, Artificial intelligence and neural networks: concept, applications and future in oncology, Br. J. Urol. 88 (Suppl. 3), 1996, 53–58.

[36] T. Toh, P.H. Tan and C.W. Cheng, Six-year follow-up of untreated T1 carcinoma of prostate, Ann. Acad. Med. Singapore 29 (2000), 201–206.

[37] G.D. Tourassi and C.E. Floyd, The effect of data sampling on the performance evaluation of artificial neural networks in medical diagnosis, Med. Decis. Making 17 (1997), 186–192.

[38] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein and R.B. Altman, Missing value estimation methods for DNA estimation, Bioinformatics 17 (2001), 520–525.
[43] R.E. Unger, J.B. Oltrogge, H. von Briesen, B. Engelhardt, U. Woelki, W. Schlote, R. Lorenz, H. Bratzke and C.J. Kirkpatrick, Isolation and molecular characterization of brain microvascular endothelial cells from human brain tumors, In Vitro Cell. Dev. Biol. Anim. 38 (2002), 273–281.

[44] P.J. van Veldhuizen, R. Sadasivan, R. Cherian, T. Dwyer and R.L. Stephens, p53 expression in incidental prostatic cancer, Am. J. Med. Sci. 305 (1993), 275–279.

[45] V.N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.

[46] R.K. Vartanian and N. Weidner, Endothelial cell proliferation in prostatic carcinoma and prostatic hyperplasia: correlation with Gleason’s score, microvessel density, and epithelial cell proliferation, Lab. Invest. 73 (1995), 844–850.

[47] S. Vesalainen, P. Lipponen, M. Talja, E. Alhava and K. Syrjänen, Tumor vascularity and basement membrane structure as prognostic factors in T1–2M0 prostatic adenocarcinoma, Anticancer Res. 14 (1994), 709–714.

[48] M. Volavsek, A. Masera and Z. Ovcak, Incidental prostatic carcinoma. A predictive role of neoangiogenesis and comparison with other prognostic factors, Pathol. Oncol. Res. 6 (2000), 191–196.

[49] J.T. Wei, Z. Zhang, S.D. Barnhill, K.R. Madyastha, H. Zhang and J.E. Oesterling, Understanding artificial neural networks and exploring their potential applications for the practicing urologist, Urology 52 (1998), 161–172.

[50] N. Weidner, Tumor angiogenesis: review of current applications in tumor prognostication, Semin. Diagn. Pathol. 10 (1993), 302–313.

[51] N. Weidner, P.R. Carroll, J. Flax, W. Blumenfeld and J. Folkman, Tumor angiogenesis correlates with metastasis in invasive prostate carcinoma, Am. J. Pathol. 143 (1993), 401–409.

[52] H. Wolter, D. Trijic, H.-W. Gottfried and T. Mattfeldt, Chromosomal changes in incidental prostatic carcinomas detected by comparative genomic hybridization, Eur. Urol. 41 (2002), 328–334.