Comparison of the predictive qualities of three prognostic models of colorectal cancer

Billie Anderson[1], J. Michael Hardin[2], Dominik D. Alexander[3], Sreelatha Meleth[4], William E. Grizzle[5], Upender Manne [5,6]

[1]SAS Institute, Cary, NC, [2]Department of Information Systems, Statistics, and Management, University of Alabama at Tuscaloosa, AL, [3]Exponent Health Sciences, Wood Dale, IL, [4] Division of Preventive Medicine, [5]Department of Pathology,[6] Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35294

TABLE CONTENTS

1. ABSTRACT

Most discoveries of cancer biomarkers involve construction of a single model to determine predictions of survival.. 'Data-mining' techniques, such as artificial neural networks (ANNs), perform better than traditional methods, such as logistic regression. In this study, the quality of multiple predictive models built on a molecular data set for colorectal cancer (CRC) was evaluated. Predictive models (logistic regressions, ANNs, and decision trees) were compared, and the effect of techniques for variable selection on the predictive quality of these models was investigated. The Kolmogorov-Smirnoff (KS) statistic was used to compare the models. Overall, the logistic regression and ANN methods outperformed use of a decision tree. In some instances (e.g., for a model that included 'all variables without tumor stage' and use of a decision tree for variable selection), the ANN marginally outperformed logistic regression, although the difference between the accuracy of the KS statistic was minimal (0.80 versus 0.82). Regardless of the variable(s) and the methods for variable selection, all three predictive models identified survivors and non-survivors with the same level of statistical accuracy.

2. INTRODUCTION

Researchers are now examining the methodology for predicting the survival or disease recurrence in cancer patients by use of data-mining techniques, such as artificial neural networks (ANNs), decision trees, and k-nearest neighbor (k-NN). In particular, this is being done to predict the clinical outcome for patients with colorectal cancer (CRC) (1-4).

Most of these studies have compared the prediction of ANNs to other methods, such as survival analysis or logistic regression. For example, Burke et al. (5) demonstrated that use of the TNM components of tumor staging variables by themselves in an ANN significantly increased the predictive accuracy of the variables when compared to a model for survival analysis with the same variables. The predictive accuracy of the variables used in the models was measured by the area-under-the-ROC curve. The ANN increased the predictive accuracy of the model by 44-74%. Also, ANNs were used by this group to build predictive models for breast cancer survival; they found that, compared to the TNM staging, the ANN provided better predictive accuracy (5, 6). Therefore, it is

not clear whether the improved predictive capacity was a reflection of the ANN method, or whether the variables such as positive lymph nodes and p53 status contributed to improving the predictability of a model. Furthermore, the predictive accuracy of a model is a factor of the variables in the model, and the technique for variable selection determines the quality of a model (7).

ANNs and other data mining tools are called 'black-box' techniques, since the logic used to determine the final model is not transparent. Although "black box" techniques such as ANNs perform better than traditional methods, the results are not uniform. In some cases, within the same study, the superiority of a classifier seems to be dependent on the variable (8). As described above, data-mining techniques such as ANNs provide higher predictive accuracy than familiar, traditional models, such as logistic regression (9-13). If the same predictive accuracy can be obtained from logistic regression, which is generally understood by statisticians, basic researchers, and clinicians, it is difficult to justify the expense, in terms of computing time, and clarity of process for the "black-box" methods. As suggested by some of the studies above, if the accuracy of prediction is variable-dependent, which in turn makes it dependent on techniques for variable selection (8), it is important to determine the effect of these techniques on the accuracy of the predictive model. Therefore, this study aimed to provide answers to these questions: 1) Do data-mining techniques, such as ANNs and decision trees, provide models with higher predictive accuracy than logistic regression, and 2) How does the technique of variable selection affect the predictive accuracy of these models?

3. PATIENT POPULATIONS, MATERIALS, AND METHODS

3.1. Patient populations

As described in previous publications (14, 15), 491 patients who had undergone surgical resection for '*first primary*' colorectal carcinoma between 1981 to 1993 at the University of Alabama at Birmingham (UAB) hospital were identified for this study. These patients were identified from the UAB Tumor Registry, following the selection criteria described below. During the initial selection process, patients who died within a week of surgery, whose archival tissues were not available, who had surgical margin-involvement, who had an unspecified tumor location, who had multiple primaries within the colorectum, who had multiple malignancies (except non-melanotic lesions of the skin), or who had a family or personal history of CRC were excluded from the study population. To control for treatment bias, only patients who underwent surgery as a therapeutic intervention were included, and patients who received any pre- or post-surgical therapy were excluded. Since adjuvant chemotherapy was not in widespread use during this study's time frame (1981-1993), a large number of Stage III and IV patients (n=212) who had not received adjuvant therapy were included.

3.2. Pathological features

Slides, stained with hematoxylin and eosin, were reviewed to determine the degree of histologic differentiation and categorized as well, moderate, poor, or undifferentiated. Tumors that were either well or moderately differentiated were designated as low-grade and those classified as either poor or undifferentiated as high-grade (16). The pathologic staging was determined according to the criteria of the American Joint Commission on Cancer (17). The codes of the International Classification of Diseases for Oncology (ICD-O) were used to specify anatomic location (colon versus rectum) of the tumor (18).

3.3. Follow-up

Patients were followed by the UAB Tumor Registries until their death or the date of the last documented contact within the study time frame. The tumor registries ascertain outcome information directly from patients (or living relatives) and from the physicians of the patients through telephone and mail contacts. This information was further validated against State Death Lists. The tumor registries update information every six months. Follow-up of our cohort ended in August 2008. The median follow-up period of the complete study population was 8.91 years (range <1->20 years).

3.4. Immunohistochemistry

Archival tissues, formalin-fixed and paraffin-embedded, were collected from the Surgical Pathology Division of the UAB Hospital. Earlier publications describe evaluation of immunohistochemical staining and immunostaining of nuclear accumulation of p53 and Bcl-2 (14, 19, 20). The staining was evaluated semi-quantitatively; the investigators involved were blinded to the patients' clinicopathologic data and their treatment status. The expression of Bcl-2 was in the cell cytoplasm; p53 accumulated in the nucleus (p53^{nac}). Both the percentage of positive cells and the staining intensity were taken into consideration in determining the final immunostaining score (ISS), as described in prior reports (14, 19-21). Molecular marker expression was dichotomized into high expressers and low expressers, based on the cut-off values described below. Consistent with findings from our prior studies (19, 20), 0.5 ISS was chosen as the cut-off value for Bcl-2 expression. Only tumor cells with distinct nuclear immunostaining for p53 were considered as positive; the tumor was considered positive only if $\geq 10\%$ of all malignant cells in a tissue section were positive, as described in earlier publications (14, 19-21).

3.5. Statistical analyses

The outcome variable for the predictive models was a binary variable indicating survival (or death) five years post-surgery for patients with CRC (disease-specific five-year survival). The training data set consisted of 234 Caucasians (80% of initial 292), and 159 (80% of 199) African-Americans; the remaining individuals formed the test set for validation (98 patients). The variables were age (≥ 65 and < 65 years), race (Caucasians *vs.* African-Americans), tumor stage (I & II *vs.* III & IV), tumor

differentiation (low grade *vs.* high grade), location of tumor (proximal colon, distal colon and rectum), and p53^{nac} and Bcl-2 expression. All variables were dichotomized (22-24).

3.6. Variable selection

No variable selection was used for the first type of analyses performed on the data set, as all available variables were used to build the predictive models. In subsequent analyses, the two procedures for variable selection were decision-tree (25) and stepwise regression. These techniques were performed with SAS® Enterprise Miner, version 5.3. The software default settings were used for each selection technique.

The two techniques were used on the full data set, and then on data sets stratified by race. The race-based analysis was conducted because our prior studies indicated that p53^{nac} was a strong predictor of poor survival, but only among non-Hispanic Caucasians who had tumors located in the proximal colon. In contrast, p53^{nac} was not a useful prognostic marker for Caucasians with distal tumors, or for African-Americans with tumors of any anatomic location in the colorectum (14). For each data set, the techniques for variable selection were conducted by including the tumor stage in the list of variables, and again after omitting the tumor stage from the list of variables available for selection. Since there was interest in judging the predictive value of the biomarker, Bcl-2, in the absence of the tumor stage variables, the analysis was conducted twice. For the first analysis, the list of all possible variables that could be included in the model had the tumor-stage variable; for the second, the stage variable was excluded from the variable list.

3.7. Predictive models

By use of all variables and the subset of variables identified by the decision-tree and the stepwise-regression methods, three predictive models were built. These included logistic regression, decision trees, and ANNs. These models were chosen because of their value, as described in recent reports (1-4). The models were built by use of SAS® Enterprise Miner software. Following is a brief description of each model type.

3.8. Logistic regression

A logistic regression was built to model the probability of survival (or lack of survival) five years post-surgery for CRC. Logistic regression is used to model data when the outcome variable is binary (e.g., survive-yes/no; recurrence-yes/no). The probability of an outcome is related to a set of predictor variables by an equation $\log[(p/(1-p)] = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k$, where p is the probability of survival five years post-surgery for CRC, beta$_0$ is an intercept term, beta$_1$,...., beta$_k$ are the coefficients associated with each variable, X_1,...., X_k are the values of the predictor variables, and k denotes a unique subscript for each variable. The standard assumption is that the predictor variables are related in a linear fashion to the log odds $_{\{\log[p/(1-p)]\}}$ of the outcome of interest.

3.9. Artificial neural networks (ANN)

Development of the ANN was inspired by the mechanism through which the brain recognizes patterns (26). The goal of an ANN is the same as in logistic regression, predicting an outcome based on the values of predictor variables. The approach used in developing the ANN model is different from logistic regression.

ANNs have the capacity to "learn" mathematical relationships between a series of input (predictor) variables and the corresponding output (outcome) variables. This is achieved by "training" the network with a data set that consists of the predictor variables and a known outcome variable. Once the ANN has been "trained," the model can be used for classification of a validation data set.

Figure 1 is a diagram illustrating an ANN that has been trained to predict the probability of a patient dying of CRC five years post-surgery based on only two predictor variables, age and race. ANNs are often represented in diagrams such as this. The circles are known as *nodes*. A typical ANN consists of three layers of nodes: input, hidden, and output. The values of the predictor variables reside in the input node. The output node contains the predicted output of the network. The hidden nodes in the diagram contain an *activation function* that allows the network to model complex nonlinear associations between the predictor variables and the outcome. Of several activation functions examined, the tangent hyperbolic gave the best results; it was chosen for use in this study.

Each input node is connected to each hidden node, and each hidden node is connected to the output node. In this example, there are two input nodes where the values of age (X_1) and race (X_2) are input into the network, along with a bias weight, which is equivalent to an intercept term in a regression model.

The input nodes are connected to the hidden nodes by a *connection weight*. (These are the lines in Figure 1 connecting the input and hidden nodes.) The connection weights can be thought of as the ANN equivalent of the beta coefficients in a logistic regression model. At each hidden node, the connection weights are passed to an activation function, most commonly the sigmoid function. The activation function uses the connection weights to model any non-linear relationships among the predictor variables and the outcome variable. Another set of connection weights are then passed from the hidden node to the output node to obtain the output of the network, which corresponds to the predicted probability of the outcome variable.

In the ANN analysis, there were as many input nodes as predictor variables. (The number of predictor nodes varies depending on the method used for variable selection.) There were three hidden nodes (the default setting in SAS® Enterprise Miner) and one output node (probability of survival five years after surgery for patients with CRC).
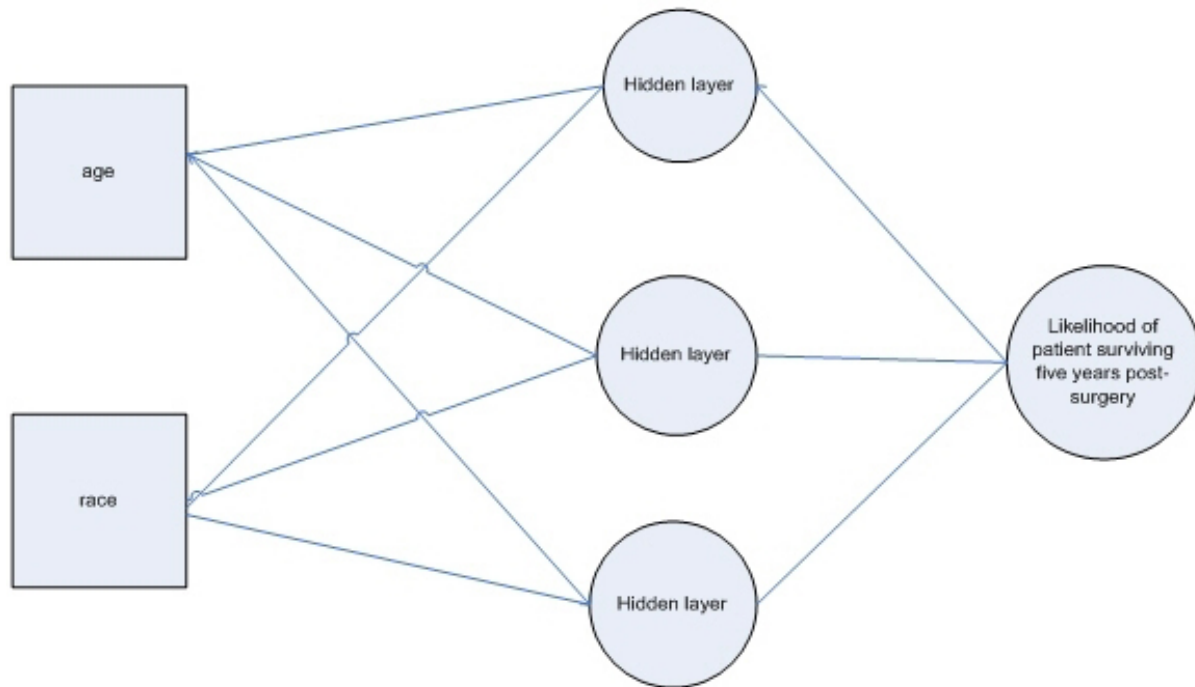
Figure 1. Diagram of the neural network trained to predict the probability of CRC patient survival of five years post-surgery on the basis of age (X_1) and race (X_2). Each box represents an input node in which the predictor variables are input into the network, and each line represents a connection weight. Each circle in the middle of the diagram represents hidden layers, where the relationship between the predictor variables and the outcome are modeled. The circle at the end of the diagram is where the probability of survival is output from the network.

3.10. Decision trees

The type of decision tree was CART (Classification and Regression Tree) (27). The settings in SAS® Enterprise Miner were set to create such a tree. CART is an algorithm used to split the data into smaller segments called 'nodes' that are homogeneous with respect to the outcome variable. At each node, the algorithm examines all predictor variables and all values of these predictors with respect to determining the best predictor variable and a value of that variable that will "best" separate the data into more homogenous subgroups with respect to the outcome variable. In other words, each node is a classification question, and the branches of the tree are partitions of the data set into different classes (those patients who will survive/not survive five years after surgery). This process repeats itself in a recursive, iterative manner until no further separation of the data is feasible. Therefore, the nodes at the end of the branches of the decision tree represent the different classes.

The second part of the algorithm is *pruning*. Pruning is applied to the decision tree to ensure that the algorithm does not over-fit the training data. At each subsequent node, smaller numbers of observations are available. Towards the end of the splitting algorithm, idiosyncrasies of the training observations at a particular node can display a pattern that is specific only to those observations that become meaningless and detrimental for prediction when applied to larger populations. Pruning removes smaller branches that fail to generalize use of the validation data set.

3.11. Measures of performance evaluation

The Kolmogorov-Smirnov (KS) statistic, which measures the difference between two different distributions, was used to evaluate model performance. The actual KS statistic is the maximum difference between two different distributions. In this case, the two distributions of interest are the estimated probabilities of belonging to the survival or non-survival groups produced by the models. If the two distributions are the same, the model does not effectively separate between survivors and non-survivors (implying a small KS statistic). On the other hand, significantly different distributions suggest good separation between the two groups (implying a larger KS statistic). Since the KS statistic has a probability distribution, a p-value is computed to determine if the two distributions are significantly different.

The predictive models were built on the set of training data, and the set of validation data was used to obtain the KS statistics and the corresponding p-value.

4. RESULTS

Table 1 displays the KS statistics for each of the predictive models when all variables were used. For all groups of interest (model categories), the logistic-

Predictive qualities of prognostic models

Table 1. Predictive accuracy of logistic regression, ANN and decision-tree models

Model category	KS statistics[1] (*P* value)		
	Logistic[3] regression	ANN[3]	Decision[3] tree
All variables[2] (including tumor stage)	1.97 (0.0009)	1.97 (0.0009)	1.46 (0.0277)
All variables (excluding tumor stage)	1.22 (0.1026)	0.99 (0.2848)	0.52 (0.9479)
Caucasians (including tumor stage)	1.44 (0.0308)	1.44 (0.0308)	1.18 (0.1224)
Caucasians (excluding tumor stage)	0.66 (0.7818)	0.92 (0.3668)	0.26 (1.0000)
African-Americans (including tumor stage)	1.47 (0.0266)	1.59 (0.0126)	0.98 (0.2923)
African-Americans (excluding tumor stage)	0.98 (0.2923)	0.73 (0.6527)	0.78 (0.5842)

KS: Kolmogorov-Smirnoff; ANN: artificial neural networks, [1] KS statistics was used to measure the predictive quality of a model, as described in the Methods section, [2] All variables were dichotomized as described in the Methods section, [3] The variables age, tumor differentiation, tumor location, p53nac and Bcl-2 expression, race, and tumor stage were included according to the model category

Table 2. Predictive accuracy of logistic regression, ANN and decision-tree models based on the variables selected by a decision-tree approach

Model category	KS statistics[1] (*P* value)			Variables[2] in final model
	Logistic regression	ANN	Decision tree	
All variables (including tumor stage)	1.46 (0.0277)	1.46 (0.0277)	1.46 (0.0277)	Tumor stage
All variables (excluding tumor stage)	0.80 (0.5375)	0.82 (0.5109)	0.52 (0.9479)	Age, tumor differentiation, and Bcl-2
Caucasians (including tumor stage)	1.18 (0.1224)	1.18 (0.1224)	1.18 (0.1224)	Tumor stage
Caucasians (excluding tumor stage)	0.39 (0.9978)	0.39 (0.9978)	0.26 (1.0000)	Age, tumor differentiation And bcl-2
African-Americans (including tumor stage)	0.98 (0.2923)	0.98 (0.2923)	0.98 (0.2923)	Tumor stage
African-Americans (excluding tumor stage)	0.78 (0.5842)	0.78 (0.5842)	0.78 (0.5842)	Tumor differentiation

KS: Kolmogorov-Smirnoff; ANN: artificial neural networks, [1] KS statistics was used to measure the predictive quality of a model, as described in the Methods section, [2] All variables were dichotomized as described in the Methods section.

Table 3. Predictive accuracy of logistic regression, ANN and decision-tree models based on the variables selected by a stepwise regression approach

Model category	KS statistics[1] (*P* value)			Variables[2] in final model
	Logistic regression	ANN	Decision tree	
All variables (including tumor stage)	1.81 (0.0029)	1.46 (0.0277)	1.81 (0.0029)	Age, tumor stage, and tumor differentiation
All variables (excluding tumor stage)	0.80 (0.5375)	0.82 (0.5109)	0.52 (0.9479)	Age, tumor differentiation, and Bcl-2
Caucasians (including tumor stage)	1.18 (0.1224)	1.18 (0.1224)	1.18 (0.1224)	Age and tumor stage
Caucasians (excluding tumor stage)	0.26 (1.0000)	0.26 (1.0000)	0.26 (1.0000)	Tumor differentiation and Bcl-2
African-Americans (including tumor stage)	0.98 (0.2923)	0.98 (0.2923)	0.98 (0.2923)	Tumor stage
African-Americans (excluding tumor stage)	0.78 (0.5842)	0.78 (0.5842)	0.78 (0.5842)	Tumor differentiation

KS: Kolmogorov-Smirnoff; ANN: artificial neural networks, [1] KS statistics was used to measure the predictive quality of a model as described in the Methods section, [2] All variables were dichotomized as described in the Methods section.

regression and the ANN methods outperformed the decision-tree method, except for the group of African-Americans without tumor stage. The ANN outperformed the logistic regression method in the groups 'Caucasians without tumor stage' and 'African-Americans with tumor stage.'

Table 2 displays the KS statistics for the predictive models when a decision tree was used for variable selection. For the three groups that included the tumor stage variable, the decision tree selected only tumor stage as a predictive variable. For the four groups that were stratified by race, all three predictive models had the same KS statistic for each predictive model, regardless of the variable(s) used in building the predictive model. The group 'all variables with stage' also had the same KS statistic for all three predictive models. The only group in which there was a difference among the KS statistics was 'all variables without stage,' which had a slightly higher KS statistic for the ANN than for the logistic regression method, with both of these models outperforming the decision tree.

Table 3 presents the KS statistics for the models when stepwise regression was used as a variable selection

technique. Once again, the groups that were stratified on race had the same KS statistic for all three predictive models. The logistic regression and the decision tree had the same KS statistic for the group 'all variables with stage,' in which these models outperform the ANN. The ANN model was best for the group 'all variables without stage.'

5. DISCUSSION

Although the focus of this manuscript is a comparison of the predictive quality of three statistical models, it is noteworthy that increased phenotypic expression of Bcl-2 in colorectal cancer tissues emerged as a strong predictor of five-year post-surgery survival, especially for non-Hispanic, Caucasian patients when the tumor stage variables were not included (Tables 2 & 3). Another clinically relevant finding is that the pathologic feature, tumor differentiation, is an important predictor of survival for both African-American and Caucasian patients when the information on tumor stage was not available. These findings are relevant to the treatment of colorectal cancer, particularly in predicting the outcome of patients who undergo excisional biopsies. Since information on all

components of TNM staging cannot be obtained from histologic assessment of biopsy specimens, these findings may be useful in assessing the aggressiveness of the tumor at the time of biopsy. In the analyses, regardless of the variable(s) and the methods for variable selection used in the models, all three prediction approaches including logistic regression, neural networks and decision tree-based models have identified survivors and non-survivors with the same level of statistical accuracy.

In a study in which the predictor variables have been dichotomized, there is the potential for residual confounding in the prediction analysis. The value of categorization of continuous predictor variables in medical studies has been considered. Most of the discussion has focused on issues such as power, specifying the correct outcome and predictor variable association, or the selection of cut-points for the continuous predictors (28, 29). Placing a continuous predictor variable into two categories reduces the bias in the model up to 64% (30, 31). Most of the confounding from a predictor variable is removed by placing the variable into two categories; it is rarely necessary to have more than five categories (32).

Overall, the logistic regression and ANN techniques outperformed the decision-tree method. In the instances where the ANN outperformed logistic regression (such as the group 'all variables without stage' by use of a decision tree as a variable selection technique), the difference between the accuracy of the KS statistic was minimal (0.80 versus 0.82). In cases such as these, it is difficult to justify use of an ANN, which is a much more complex model and is not as well understood as logistic regression. However, logistic regression analysis does not take censoring into account. The recommendation from this study is that, unless the ANN outperforms a simpler model (such as a logistic regression or a decision tree), the simpler model should be used. Relative to interpretation and understandability, more will be gained by use of the less complex model, although it may not be as predictive).

Thus, the techniques for variable selection had a significant effect on the predictive accuracy of the models. For example, the highest KS statistics were obtained when all variables were used. The models involving all the predictor variables outperformed the models built by use of the two techniques for variable selection across all model categories. Yet, when the models were compared by use of different forms of techniques for variable selection, there was less consistency.

For example, Table 2 shows that the logistic and ANN developed by use of a decision-tree approach for variable selection outperforms the logistic and ANN developed by use of an approach involving stepwise regression for the model category "Caucasians excluding tumor stage." Table 3 shows that the logistic and decision tree outperforms the logistic and decision tree developed in Table 2 for the model category that included all features, including tumor stage. Tables 2 and 3 demonstrate that when variable selection procedures are used to develop predictive models, the results are not consistent (i.e., there is not one model consistently outperforming the others for the different model categories).

In conclusion, various models can be used to predict survival of CRC patients five years post-surgery. This study used stepwise regression and decision trees to select the variables to be entered into the models. The KS statistic was then used on a validation data set to determine how well the models separated between the survival and non-survival groups by use of predictor variables chosen by the techniques for selection of variables. This *proof of principle* study demonstrates the potential of predictive models to assess the survival probabilities of patients with CRC. These findings might be useful in cancer biomarker data analyses specifically to address a binary outcome with all binary predictive variables. We are now examining the capacity of these models to discover additional associations and patterns in the prognostic markers of CRC.

6. ACKNOWLEDGEMENTS

7. REFERENCES

1. Anand, S. S., A. E. Smith, P. W. Hamilton, J. S. Anand, J. G. Hughes & P. H. Bartels: An evaluation of intelligent prognostic systems for colorectal cancer. *Artif Intell Med*, 15, 193-214 (1999)

2. Grumett, S., P. Snow & D. Kerr: Neural networks in the prediction of survival in patients with colorectal cancer. *Clin Colorectal Cancer*, 2, 239-244 (2003)

3. Lundin, M., J. Lundin, H. B. Burke, S. Toikkanen, L. Pylkkanen & H. Joensuu: Artificial neural networks applied to survival prediction in breast cancer. *Oncology*, 57, 281-286 (1999)

4. Snow, P. B., D. J. Kerr, J. M. Brandt & D. M. Rodvold: Neural network and regression predictions of 5-year survival after colon carcinoma treatment. *Cancer*, 91, 1673-1678 (2001)

5. Burke, H. B., P. H. Goodman, D. B. Rosen, D. E. Henson, J. N. Weinstein, F. E. Harrell, Jr., J. R. Marks, D. P. Winchester & D. G. Bostwick: Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer*, 79, 857-862 (1997)

6. Burke, H. B., A. Hoang, J. D. Iglehart & J. R. Marks: Predicting response to adjuvant and radiation therapy in patients with early stage breast carcinoma. *Cancer*, 82, 874-877 (1998)

7. Grizzle, W. E. & S. Meleth: Clarification in the point/counterpoint discussion related to surface-enhanced laser desorption/ionization time-of-flight mass spectrometric identification of patients with adenocarcinomas of the prostate. *Clin Chem*, 50, 1475-6; author reply 1476-1477 (2004)

8. Meleth, S., C. Chatla, V. Katkoori, B. Anderson, J. Hardin, N. Jhala, A. Bartolucci, W. Grizzle & U. Manne: Comparison of predicted probabilities of proportional hazards regression and linear discriminant analysis methods using a colorectal cancer molecular biomarker database. *Cancer Informatics*, 2, 1-8 (2007)

9. Finne, P., R. Finne, A. Auvinen, H. Juusela, J. Aro, L. Maattanen, M. Hakama, S. Rannikko, T. L. Tammela & U. Stenman: Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. *Urology*, 56, 418-422 (2000)

10. Matsui, Y., S. Egawa, C. Tsukayama, A. Terai, S. Kuwao, S. Baba & Y. Arai: Artificial neural network analysis for predicting pathological stage of clinically localized prostate cancer in the Japanese population. *Jpn J Clin Oncol*, 32, 530-535 (2002)

11. Remzi, M., T. Anagnostou, V. Ravery, A. Zlotta, C. Stephan, M. Marberger & B. Djavan: An artificial neural network to predict the outcome of repeat prostate biopsies. *Urology*, 62, 456-460 (2003)

12. Song, J. H., S. S. Venkatesh, E. A. Conant, P. H. Arger & C. M. Sehgal: Comparative analysis of logistic regression and artificial neural network for computer-aided diagnosis of breast masses. *Acad Radiol*, 12, 487-495 (2005)

13. Tu, J. V.: Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol*, 49, 1225-1231 (1996)

14. Manne, U., H. L. Weiss, R. B. Myers, O. K. Danner, C. Moron, S. Srivastava & W. E. Grizzle: Nuclear accumulation of p53 in colorectal adenocarcinoma: prognostic importance differs with race and location of the tumor. *Cancer*, 83, 2456-2467 (1998)

15. Meleth, S., C. Chatla, V. R. Katkoori, B. Anderson, M. J. Hardin, N. C. Jhala, A. Bartolucci, W. E. Grizzle & U. Manne: Comparison of Predicted Probabilities of Proportional Hazards Regression and Linear Discriminant Analysis Methods Using a Colorectal Cancer Molecular Biomarker Database. *Cancer Informatics*, 2, 1-8 (2007)

16. Compton, C., C. M. Fenoglio-Preiser, N. Pettigrew & L. P. Fielding: American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. *Cancer*, 88, 1739-1757 (2000)

17. Greene, F. L.: The American Joint Committee on Cancer: updating the strategies in cancer staging. *Bull Am Coll Surg*, 87, 13-15 (2002)

18. WHO: International classification of diseases for oncology. 2nd ed., World Health Organization. (1990).

19. Manne, U., R. B. Myers, C. Moron, R. B. Poczatek, S. Dillard, H. Weiss, D. Brown, S. Srivastava & W. E. Grizzle: Prognostic significance of Bcl-2 expression and p53 nuclear accumulation in colorectal adenocarcinoma. *International Journal of Cancer*, 74, 346-358 (1997)

20. Grizzle, W. E., R. B. Myers, U. Manne & S. Srivastava: Immunohistochemical evaluation of biomarkers in prostatic and colorectal neoplasia. In: John Walker's Methods in Molecular Medicine-Tumor Marker Protocols. Eds: M. Hanausek & Z. Walaszek. Humana Press, Totowa, NJ (1998)

21. Manne, U., H. L. Weiss & W. E. Grizzle: Bcl-2 expression is associated with improved prognosis in patients with distal colorectal adenocarcinomas. *Int J Cancer*, 89, 423-430 (2000)

22. Delen, D., G. Walker & A. Kadam: Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods. *Artificial Intelligence in Medicine*, 34, 113-127 (2005)

23. Lavrac, N.: Selected Techniques for Data Mining in Medicine. *Artificial Intelligence in Medicine*, 16, 3-23 (1999)

24. Hobbs, G. R.: Data mining and health informatics. *American Journal of Health Behavior*, 25, 285-289 (2001)

5. Povalej, P., M. Lenic, M. Zorman, P. Kokol & D. Dinevski: Accuracy of Intelligent Medical Systems. *Computer Methods and Programs in Biomedicine*, 80, S95-S105 (2005)

26. White, H.: Learning in artificial neural networks. *Neural Computing*, 1, 425-426 (1989)

27. Brieman, L., R. A. Olshen & C. J. Stone: Classification and Regression Trees. Wadsworth and Books Cole, Belmont, CA (1984)

28. Greenland, S.: Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. *Epidemiology*, 6, 450-454 (1995)

29. Lagakos, S. W.: Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. *Stat Med*, 7, 257-274 (1988)

30. Becher, H.: The concept of residual confounding in regression models and some applications. *Stat Med*, 11, 1747-1758 (1992)

31. Cochran, W. G.: The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, 24, 295-313 (1968)

32. Rothman, K. J.: Modern Epidemiology. Little and Brown (1986)

Send correspondence to: Upender Manne, Department of Pathology, University of Alabama at Birmingham, 515B1-Kracke Building 619, 19th Street South, Birmingham, AL, 35294-7331, Tel: 205-934-4276, Fax: 205-934-4418, E-mail: manne@uab.edu