

Predictive Modeling Techniques in Prostate Cancer

ASHUTOSH TEWARI, M.D.,¹ CHRISTOPHER PORTER, M.D.,² JAMES PEABODY, M.D.,¹
E. DAVID CRAWFORD, M.D.,³ RAYMOND DEMERS, M.D., M.P.H.,¹ CHRISTINE C. JOHNSON, Ph.D.,¹
JOHN T. WEI, M.D.,⁴ GEORGE W. DIVINE, Ph.D.,¹ COLIN O'DONNELL, B.S.,³
EDUARD J. GAMITO, B.S.,³ and MANI MENON, M.D.¹

ABSTRACT

A number of new predictive modeling techniques have emerged in the past several years. These methods can be used independently or in combination with traditional modeling techniques to produce useful tools for the management of prostate cancer. Investigators should be aware of these techniques and avail themselves of their potentially useful properties. This review outlines selected predictive methods that can be used to develop models that may be useful to patients and clinicians for prostate cancer management.

INTRODUCTION

PREDICTIVE MODELING IN MEDICINE involves the use of a given data set, which is collected either retrospectively or prospectively, to derive a mathematical model to predict outcomes for future patients. A number of predictive models are available to health care professionals engaged in the care of patients with cancer of the prostate (CaP). Perhaps the most widely used of these are the tables developed by Partin et al.¹ These tables are nomograms for predicting pathologic stage from clinical variables in men with clinically localized CaP. A number of models also are available for predicting treatment outcomes and survival in patients with various stages of CaP. These models can provide physicians and patients with a scaffold on which to base their therapeutic decision-making.

The majority of predictive models currently used in CaP management have been developed utilizing traditional statistical techniques such as multivariate logistic regression. New techniques have emerged, such as artificial neural networks (ANNs), that can be used to complement traditional techniques. Investigators should be aware of these new techniques that can be used independently or in combination with traditional statistical methods to develop accurate predictive models.²⁻¹⁴ This review examines techniques that are available for developing predictive models for use in CaP management.

PREDICTIVE MODELING TECHNIQUES

The goal of predictive modeling is a model that is both reproducible and accurate in predicting the outcome of interest. Traditionally, prediction of an event, y , from a circumstance, x , has been accomplished by simple linear regression. However, biological systems often require multiple variables ($x_1, x_2, x_3, \dots, x_n$) to predict an outcome, y . For these tasks, more sophisticated methods may be required. Some of these methods are described below.

Mechanistic Modeling

Mechanistic models are mathematical tools that describe the cause and effect relationships of a proposed action in terms of a mathematical function, such as the pull of gravity on an object or the resistance to a current in a wire.¹⁵⁻¹⁸ Mathematical models can come in all degrees of complexity, from simple variations of mass balance equations to highly complex multivariate systems. However, purely mechanistic mathematical functions may fail to describe fully systems that have strong stochastic elements (random processes) contributing to the model output. For these systems, including many biological systems, statistical routines have been developed that include an estimation of the error associated with model output. Many such

¹Vattikuti Urology Institute and Josephine Ford Cancer Center, Henry Ford Hospital, Detroit, Michigan.

²Department of Urology, State University of New York, Stony Brook, New York.

³ANNs in CaP Project, Institute for Clinical Research, Denver, Colorado.

⁴Division of Urology, University of Michigan, Ann Arbor, Michigan.

modeling methods have been adapted for computers. Some statistical models may not always be useful in medical predictions, as biological systems are often quite complex, necessitating approximations and simplifying assumptions that may introduce further uncertainty and error. However, even when there is substantial uncertainty about the true nature of the mechanisms and the values of critical parameters, astute use of alternative models can provide valuable guidance in making medically informed management decisions.

Stochastic Modeling

In statistical models, the *dependent* or *effect* variable, y , is *predicted* by one or more independent or causal variables, x .^{19,20} Statistical models incorporate methods to determine how much of the observed variability can be explained by the independent variables, and how much is due to “chance”; i.e., is unexplained. Inferential statistics, by definition, measure the reliability of conclusions drawn about a population based on a population sample.

There are five advantages of statistical methods. First, they use actual observations, either random or imposed. Second, more observations should lead to more reliable predictions in a predictable and quantifiable manner. Third, the predictions are on a continuous scale and allow for fine distinctions if there are enough data. Fourth, there is a rich set of analytical methods available. Finally, each prediction comes with an estimate of its reliability. The main disadvantages are as follows. First, the form of the statistical relation is not obvious *a priori* and several models may give similar results in terms of goodness-of-fit. Second, assumptions about the distribution of the predictor and response variables are required and may be impossible to verify. Third, unusual observations such as outliers may cause problems with the model. Fourth, extrapolation to values of the independent variable outside the domain of calibration is not justified. Fifth, it is not obvious how to define the sample space, and errors here invalidate the inferences. Sixth, it is not obvious which independent variables should be included in the relation. Finally, the precision of the statistical relation may not be useful enough for meaningful predictions, especially with observational data.

Simple Linear Regression

Simple linear regression assumes a linear relation, within a certain range, between a single independent predictor and the dependent variable.^{3,21–24} Rarely, however, is a single predictor enough to explain most of the observed response in a biological system. The process of fitting a regression equation to observed data is called “calibration,” which yields a goodness-of-fit measure, such as R^2 . If the sample is truly representative of the desired sample space, we would expect to obtain the same parameters, within experimental and observational error, in similar repeated studies. These models are the most commonly used predictive models.

Cox Proportional Hazards Model

If we were to plot survival against time, the result would be a survival plot; i.e., the function $S(t)$. This plot would describe those patients who survive beyond a given time, t . An alternative way of looking at such data would be to look at the pop-

ulation who do not survive beyond time, t . This can be described as a logarithmic function where $h(t)$ is known as the hazard function and can be represented as follows, $h(t) = -d/dt (\ln S(t))$. This equation essentially describes the negative slope of the survival plot. When examining survival, it must be looked at in reference to time and may be looked at in reference to more than one variable. When this is the case, Cox’s proportional hazards model can be used.^{3,21,25–29} Cox’s model assumes that the independent variables are related to survival by the hazard function (3). The hazard function—those not surviving beyond time, t —is influenced by the independent variable(s) and regression coefficients ($b_1, b_2 \dots b_n$). The hazard function and the relation to independent variables and regression coefficients can be expressed as $h_0(t) e^{(x_1 b_1 + x_2 b_2 + \dots x_n b_n)}$, where x_1 is the first independent variable and b_1 is the first regression coefficient. The regression coefficients can be estimated together with the standard errors. This will then allow an estimate of confidence intervals to be made for the relative risks described by Cox’s model. Cox’s model allows us to compare the survival of two groups of patients in reference to multiple independent variables. Logistic regression may also be used to analyze survival data; however, while logistic regression models estimate the effects of the independent variables over a constant time period, the Cox method models the risk as a continuous function over time.

Nonlinear Regression

The basic purpose of regression is to predict an outcome from a given set of observations. The regression model allows for a best-fit line to be used to describe the data.^{3,21,25–29} If simple (linear) regression is used and the degree of correlation between the dependent and independent variables is high, the residual variance (error of the residuals) will be low, resulting in a good predictor of outcome. Unfortunately, in biological systems, the response or outcome may have a nonlinear relation to the variable. In fact, quite often, there are experimental results or theoretical considerations suggesting that the response may be nonlinear. Many nonlinear effects can be linearized; that is, new variables can be created by transforming the original variables, so that the resulting regression on the new variables is linear. An investigator may try any number of transformations in an attempt to linearize the relations within the data at hand. Ideally, these transformations should have some theoretical basis. If the transformation succeeds, the predictor is said to be “intrinsically linear” because we can now fit a linear equation (in the parameters) to the transformed data. The need for a transformation may be indicated if: (1) for significance testing in regression, it is required that experimental (observational) errors be independently and normally distributed with a common variance. Often, we can see that this is not the case; certain transformations will restore these conditions; or (2) a plot of the residuals (fitted v observed values of the dependent variable) show a pattern when plotted against the predictor. Some common transformations include:

Logarithmic: This transformation equalizes the variances when standard deviations are proportional to the mean;

Square root: When enumeration (small integer) data follow a Poisson distribution where the mean and variance are equal, taking the square root of the response variable restores normality. This is equivalent to squaring the predictor variable;

Angular (inverse sine): This manipulation transforms percentile data to a normal distribution.

Multivariate Statistical Methods

In medical applications, several factors may influence outcomes. Multiple regression integrates various input variables and calculates a single equation by one of two methods: (1) trying every possible combination of variables (all possible multiple regressions) or (2) picking the one with the best fit.^{3,21,25-29} Both approaches have inherent advantages and disadvantages.

A practical approach is to run a linear regression on each variable and use the best single predictor. Predictors are then added one at a time, always adding the next predictor that most improves the fit, until the fit (R^2) does not significantly improve. This method is called "forward multiple regression." An advantage of this approach is that it uses the minimum number of predictors, computation is relatively easy, and it uses the least number of variables necessary to explain the result. In backward multiple regression, the least-important variables are eliminated one at a time until the fit becomes significantly worse. Stepwise multiple regression involves the reexamination of all variables at each step.

There are several practical and theoretical problems with multiple linear regression including: (1) a poor understanding of physical significance of predictors, i.e., does the equation have any explanatory power?; (2) a conceptual conflict between using the most meaningful equation versus statistically significant relation; and (3) uncertain accounting for interactions because input variables will often interact either positively (synergistically) or negatively (compensatory). These interactions must be determined for each case. Linear regression cannot account for these, and the resulting equations are intrinsically nonlinear.

Principal Component Analysis

Principle component analysis,³⁰ also called eigenvector decomposition, is a method for transforming a multidimensional space to another one of the same dimension (i.e., same number of axes or variables) with two very interesting and important properties. First, the first component (or synthetic variable) explains the highest proportion of the total variance, the second variable the second-highest proportion, etc. Therefore, the less significant variables (dimensions) can usually be discarded as insignificant noise, thereby effectively reducing the dimensionality of the problem. Second, the axes are orthogonal (mutually perpendicular) in multidimensional Euclidean space, so that the principal components are completely uncorrelated. Mathematically, this is accomplished by finding the eigenvectors (synthetic variables) and eigenvalues (their variances; i.e. importance) of the variance-covariance matrix of the predictor variables.

This method is a theoretically satisfactory way to handle the problem of correlated predictors. The first few components should be sufficient for a stepwise regression. Also, in a stepwise regression, additional synthetic predictor variables do not change the coefficients of variables already in the equation, because the predictors are uncorrelated.

Dynamic Simulation Modeling

Statistical modeling attempts to describe a static relation. In many situations, this may not give satisfactory results because

of the dynamic (time-dependent) nature of problem. One approach for solving this problem is simulation, which involves building mathematical or computer models of a system and using these models to study the properties of the system.^{2,9,31-34} There are different types of models such as dynamic, explanatory, descriptive or realistic.

Dynamic models include time as an explicit element. In dynamic models, the state of the system at one time, plus the driving forces, follow definite transformation relations to reach the next state, and so on till the end of the simulation. This is sometimes called the "state-variable approach." Explanatory models attempt to explain how a system works from some first principles. Descriptive models simply attempt to characterize a system for predictive purposes, without pretending to explain it. Statistical models are a subclass of these models. All realistic models contain large doses of subjectivity, judgment, and empirical parameters. Although it would seem that a dynamic simulation model, being more mechanistic and explanatory than a statistical model, would be better able to extrapolate, this is not always so and must be established by validation over the expected range of inputs, just as in a statistical model. There is no assurance, except accumulated evidence, that the physical basis of the model is correct. In fact, medical problems are very complex and require immense computing power to model basic predictions. This field, however, is rapidly growing and should provide promising results in the near future.

Bayesian Modeling

Classical inferential models do not take into account "prior knowledge" in the calculations. For the purist, this is an appropriate response to prevent the introduction of extraneous data that might skew the experimental results. However, there are times when the prior knowledge would be a useful contribution to the evaluation process. Based on this philosophy, Bayes' theorem was developed with its powerful tools, both for graphically representing the relations among a set of variables, and for dealing with uncertainties in expert system.^{24,35-39} The major benefit of Bayesian inference over classical statistical inference (which deals with confidence levels rather than statements of probability) is that it explicitly describes the fact that observation alone cannot predict the probability of unobserved events without some prior information about the latter. In the Bayesian interpretation, a probability describes the strength of the belief an observer can justifiably hold that a certain statement is true (*subjective* probability). The subject, after observing the outcome of an "experiment" (i.e., collecting new data), updates the belief held before the experiment (the "prior probability"), producing a "posterior probability." The need to assume prior beliefs is a key part of Bayesian inference.

In the field of CAP, these models have been tested by Tewari and Ghosh (in press) and have shown promising results in the pilot studies involving prediction of the pathologic stage of the cancer.

Artificial Neural Networks

Artificial neural networks (ANNs) are another means of modeling data to provide an outcome prediction based on a set of given independent variables.^{6,39-50} They are nonlinear mathematical models that are characterized by a complex structure of interconnected computational elements. These computational

elements (nodes) aggregate a series of inputs (prognostic factors) using a summation operation and produce an output, such as the probability of 10-year survival. Inputs to each neuron are multiplied by a weight factor that reflects the excitatory or inhibitory strength of the connection from the input source to the neuron. The sum of the weighted inputs plus a bias term then goes through an activation function that behaves like a “switch” to determine whether or not the neuron will “fire” and thus send an output signal. The bias term may be viewed as the threshold that the weighted sum of inputs must exceed before the neuron sends an output signal. The information-processing capacity of an ANN is a function of the type and quantity of nodes in a given network and by the arrangement of interconnections between nodes. The “knowledge” or the “processing capability” of an ANN is made possible by the actual values of the interconnection. The knowledge is acquired through a learning phase, during which examples of data are repeatedly fed through the ANN, and the connection weights in the ANN are adjusted adaptively for the ANN as a whole to satisfy some predetermined performance goals.

Existing ANNs are used primarily to evaluate outcome and have been introduced as an alternative to classical statistical models. As a tool for predicting outcomes they can be judged against classical statistical methods by comparing their performances on a receiver operator characteristic (ROC) curve. In such comparisons, ANNs have performed comparably, and in some cases favorably, when predicting outcomes for CaP patients.

Genetic Adaptive Modeling

Genetic adaptive modeling is based on genetic mechanisms of evolution.^{47,48,51–55} These mechanisms are known to result in survival and genesis of intelligent, self-organizing, self-repairing, self-motivating organisms—namely, those that are the strongest among the pool of individuals and genetic patterns. The Darwinian theory of evolution depicts biological systems as products of the ongoing process of natural selection. Likewise, genetic algorithms allow scientists to use a computer to evolve solutions over time, instead of designing them at the outset, without knowing the trend and final outcome. These algorithms emulate the process of natural selection and survival of the fittest by searching high-dimensional spaces for superior solutions. The algorithms are simple, robust, and general; no knowledge of the search space is usually assumed.

In genetic algorithms, selection operates on strings of binary digits stored in the computer’s memory, and over time, the functionality of these strings evolves in much the same way that natural populations of individuals evolve. These algorithms evolve individuals using principles of variation, selection, inheritance, crossover, and mutation. Once all individuals in the population have been evaluated, their individual fits are used as the basis for selection. Eliminating individuals with poor goodness of fit in the population achieves selection, and inheritance is implemented by making multiple copies of individuals with high goodness of fit. Using mutation and crossover, the model transforms the previous set of good individuals into a new one, thus generating individuals of which some may be better than those of previous generations. Conventional optimization techniques are based on adjusting the parameters of a model to produce a desired result. For example, training an

ANN involves modifying its weights so that it produces the desired relation between inputs and outputs. Genetic algorithms, on the other hand, optimize the performance on the basis of biological genetics and natural selection. This involves maintenance and modification of the characteristics of a population of solutions (individuals) over a large number of generations.

Genetic algorithms operate on coded parameters rather than the raw parameters. This is quite like the strand of DNA, which encodes all of the characteristics of a human in chains of amino acids, so the parameters of the problem must be encoded in finite-length strings. Optimization is performed on a set of strings, where each string is composed of a sequence of characters. Given an initial population of strings, a genetic algorithm produces a new population of strings according to a set of genetic rules. This constitutes one generation. The rules are devised so that the new generations tend to have strings that are superior to those of previous generations, as measured by some objective function. Thus, successive generations are better than previous ones. These genetic adaptive models have been used recently in the management of CaP.

Self-Organizing Models

A regression-based method for model self-organization is the group method of data handling (GMDH).^{32,33,44,56} The GMDH combines the principles of statistics and ANNs under the framework of the principle of induction. The GMDH creates models adaptively guided by this cybernetic principle from data in the form of networks of optimized transfer functions (active neurons). This is done in an evolutionary fashion of repetitive generation of populations, such that alternative models of increasing complexity are developed. Corresponding model validation and survival-of-the-fittest selection is performed until an optimal model has been created. Neither the number of neurons and number of layers in the network nor the actual behavior of each created neuron (transfer function of active neuron) are predefined. All this is adjusted during the process of self-organization by the process itself. As a result, an explicit analytical model representing relevant relations between input and output variables is available immediately after modeling. This model contains the extracted knowledge applicable for interpretation, prediction, classification, or diagnosis problems.

Sometimes, we can also use principles of rule-based models in the form of binary or fuzzy logic. This approach involves rule induction from data using genetic algorithms where the representation of models is in the familiar disjunctive normal form.

Another approach is to perform symbolic modeling using self-organizing structured modeling that generates symbols of an appropriate model structure (algebraic formula or complex process models) and optimizes by means of genetic algorithms. This approach assumes that the elementary components are predefined (model base) and suitably coded genetically.

Nonparametric Models

Nonparametric models include analog complexing and objective cluster analysis. Analog complexing selects nonparametric prediction models from a given data set representing one or more patterns of a trajectory of past behavior, which are analogous to a chosen reference pattern. The purpose of objective

cluster analysis algorithms is to automatically subdivide a given data set optimally into groups of data with similar characteristics (classification). The algorithms automatically select the optimal number of clusters, their width, and their composition.

Predictive Model Validation

Model validation is the testing of a model with data from the population under study that were not used to develop the model. The validation data set should be drawn at random from the study data and not used in any way during the development of the model. A measure of a predictive model's performance can be derived from a ROC curve. The area under the ROC curve generated on the validation set can be reported as an objective measure of the model's overall performance. In addition, sensitivity, specificity, negative predictive value, and positive predictive value at a specific model output cut-off may be reported to further describe the model's performance.

CONCLUSIONS

Investigators should not limit themselves to traditional statistical techniques when developing predictive models for medical applications. The use of a combination of techniques can produce reproducible and accurate models. These models may benefit patients and clinicians by providing additional information to aid in the management of CAP.

REFERENCES

- Partin AW, Kattan MW, Subong EN, et al. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer: A multi-institutional update. *JAMA* 1997;277:1445-1451.
- Albertsen PC. Computer modeling: What should we look for? *J Urol* 1998;159:934.
- Bauer JJ, Connelly RR, Sesterhenn IA, et al. Biostatistical modeling using traditional variables and genetic biomarkers for predicting the risk of prostate carcinoma recurrence after radical prostatectomy. *Cancer* 1997;79:952-962.
- Clewell HJ 3rd, Andersen ME. Risk assessment extrapolations and physiological modeling. *Toxicol Ind Health* 1985;1:111-131.
- De Laurentiis M, Ravdin PM. Survival analysis of censored data: Neural network analysis detection of complex interactions between variables. *Breast Cancer Res Treat* 1994; 32:113-118.
- Douglas TH, Moul JW. Applications of neural networks in urologic oncology. *Semin Urol Oncol* 1998;16:35-39.
- Hobbs GR. Data mining and healthcare informatics. *Am J Health Behav* 2001;25:285-289.
- Keller BE. Mathematical modeling. *Int J Radiat Oncol Biol Phys* 1977;2:823.
- Lumb JR. Computer simulation of biological systems: Current trends. *Mol Cell Biochem* 1987;73:91-98.
- Najjar YM, Basheer IA, Hajmeer MN. Computational neural networks for predictive microbiology I: Methodology. *Int J Food Microbiol* 1997;34:27-49.
- Sakallaris BR, Jastremski CA, Von Rueden KT. Clinical decision support systems for outcome measurement and management. *AACN Clin Issues* 2000;11:351-362.
- Steen PM. Approaches to predictive modeling. *Ann Thorac Surg* 1994;58:1836-1840.
- Wei J, Tewari A. Artificial neural networks: Pro. *Urology* 1999;54:945-948.
- Tewari A. Artificial intelligence and neural networks: Concept, applications and future in urology. *Br J Urol* 1997;80(suppl3):53-58.
- Bartoszy, nacute, Ski R, et al. Modeling cancer detection: Tumor size as a source of information on unobservable stages of carcinogenesis. *Math Biosci* 2001;171:113-142.
- Dourson ML, Andersen ME, Erdreich LS, MacGregor JA. Using human data to protect the public's health. *Regul Toxicol Pharmacol* 2001;33:234-256.
- Kavlock RJ. Recent advances in mathematical modeling of developmental abnormalities using mechanistic information. *Reprod Toxicol* 1997;11:423-434.
- Rhomberg L. Risk assessment and the use of information on underlying biologic mechanisms: A perspective. *Mutat Res* 1996;365: 175-189.
- Philippe P. Chaos, population biology, and epidemiology: Some research implications. *Hum Biol* 1993; 65:525-546.
- Vessey JT, Howard KI, Lueger RJ, Kachele H, Mergenthaler E. The clinician's illusion and the psychotherapy practice: An application of stochastic modeling. *J Consult Clin Psychol* 1994;62:679-685.
- Bauer JJ, Connelly RR, Seterhenn IA, Srivastava S, McLeod DG, Moul JW. Statistical modeling using preoperative prognostic variables in predicting extracapsular extension and progression after radical prostatectomy for prostate cancer. *Mil Med* 1998;163:615-619.
- Djulgovic B, Woodcock TM. Mathematical modeling or waiting decades for an empirical answer? *Med Decis Making* 1992; 12:77.
- Harrell FE Jr, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. *J Natl Cancer Inst* 1988;80:1198-202.
- Slate EH, Turnbull BW. Statistical models for longitudinal biomarkers of disease onset. *Stat Med* 2000;19:617-637.
- Bostwick DG. Practical clinical application of predictive factors in prostate cancer: A review with an emphasis on quantitative methods in tissue specimens. *Anal Quant Cytol Histol* 1998; 20:323-342.
- Ohno Machado L. A comparison of Cox proportional hazards and artificial neural network models for medical prognosis. *Comput Biol Med* 1997;27:55-65.
- Trikalinos TA, Ioannidis JP. Predictive modeling and heterogeneity of baseline risk in meta-analysis of individual patient data. *J Clin Epidemiol* 2001;54:245-252.
- Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. *Stat Med* 1992;11:2093-2109.
- Sargent DJ. A general framework for random effects survival analysis in the Cox proportional hazards setting. *Biometrics* 1998; 54:1486-1497.
- Sun DQ, Feng CH, Bocking A, Biesterfeld S. Principal component analysis: A factor analytical technique for the determination of interobserver variation in histomorphologic tumor grading. *Anticancer Res* 1994;14:1525-1528.
- Cao C, Leong TY, Leong AP, Seow FC. Dynamic decision analysis in medicine: A data-driven approach. *Int J Med Inf* 1998;51:13-28.
- Frohn H, Geiger H, Singer W. A self-organizing neural network sharing features of the mammalian visual system. *Biol Cybern* 1987;55:333-343.
- Kuijpers K, Smith J. A computer simulation of a neuron net model as a self-organizing system. *Kybernetik* 1973;12:216-222.
- Zupan B, Demsar J, Kattan MW, Beck JR, Bratko I. Machine learning for survival analysis: A case study on recurrence of prostate cancer. *Artif Intell Med* 2000;20:59-75.

35. Gangnon RE, Clayton MK. Bayesian detection and modeling of spatial disease clustering. *Biometrics* 2000;56:922-935.
36. Gustafson P. Flexible Bayesian modelling for survival data. *Life-time Data Anal* 1998;4:281-299.
37. Tweedie RL, Scott DJ, Biggerstaff BJ, Mengersen KL. Bayesian meta-analysis, with application to studies of ETS and lung cancer. *Lung Cancer* 1996;14(suppl 1):S171-S194.
38. Muller P, Parmigiani G, Schildkraut J, Tardella L. A Bayesian hierarchical approach for combining case-control and prospective studies. *Biometrics* 1999;55:858-866.
39. Wu YC, Gustafson DH. Removing the assumption of conditional independence from Bayesian decision models by using artificial neural networks: Some practical techniques and a case study. *Artif Intell Med* 1994;6:437-454.
40. Dayhoff JE, DeLeo JM. Artificial neural networks. *Cancer* 2001; 91:1615-1635.
41. Faraggi D, Simon R. A neural network model for survival data. *Stat Med* 1995;14:73-82.
42. Liestol K, Andersen PK, Andersen U. Survival analysis and neural nets. *Stat Med* 1994;13:1189-1200.
43. Lundin J. Artificial neural networks in outcome prediction. *Ann Chir Gynaecol* 1998;87:128-130.
44. Marshall JA, Srikanth V. Curved trajectory prediction using a self-organizing neural network. *Int J Neural Syst* 2000;10:59-70.
45. Montie JE, Wei JT. Artificial neural networks for prostate carcinoma risk assessment: An overview. *Cancer* 2000;88:2655-2660.
46. Naguib RN, Sherbet GV. Artificial neural networks in cancer research. *Pathobiology* 1997;65:129-139.
47. Tewari A, Narayan P. Novel staging tool for localized prostate cancer: A pilot study using genetic adaptive neural networks. *J Urol* 1998;160:430-436.
48. Tewari A, Carlson G, Calvanese C, et al. Genetic adaptive neural network models to predict outcome of needle biopsy in 3773 patients with elevated PSA [abstract]. *J Urol* 1999;161:1256A.
49. Tewari A, Sonke G, Verbeek A, et al. Genetic adaptive neural network models to predict urodynamic obstruction in 1900 men with BPH: A multi institutional study. Presented at the Society of Female Urology and Urodynamics. Dallas, May 1999.
50. Wei JT, Zhang Z, Barnhill SD, Madyastha KR, Zhang H, Oesterling JE. Understanding artificial neural networks and exploring their potential applications for the practicing urologist. *Urology* 1998;52:161-172.
51. Wessels LF, van Someren EP, Reinders MJ. A comparison of genetic network models. *Pac Symp Biocomput* 2001:508-519.
52. Tewari A, Calvanese C, Carlson G, et al. An Artificial intelligence based genetic adaptive neural network model in management of prostate cancer. Presented at the American Association for Cancer Research. Palm Springs, December, 1998.
53. Tewari A, Mager J, Kamerer A, Shukla A, Narayan P. Genetic adaptive probabilistic neural network models in prediction of pathological stage in management of localized prostate cancer: A pilot study [abstract]. *J Urol* 1997;157:A293.
54. Potter SR, Miller MC, Mangold LA, et al. Genetically engineered neural networks for predicting prostate cancer progression after radical prostatectomy. *Urology* 1999;54:791-795.
55. Goodman ED. An Introduction to GALOPPS—The Genetic ALgorithm Optimized for Portability and Parallelism System, Release 3.2. Lansing: Intelligent Systems Laboratory and Case Center for Computer-Aided Engineering and Manufacturing, Michigan State University, 1996, p 84.
56. Polanski J, Walczak B, Gieleciak R, Lukaszuk S. Self-organizing neural networks for modeling 3D QSAR: A comparative study. *Acta Pol Pharm* 2000;57(suppl):76-79.

Address reprint requests to:
Eduard J. Gamito
ANNs in CAP Project
300 S. Jackson St., Suite 540
Denver, CO 80209

E-mail: ed.gamito@uchsc.edu