Risk factor identification and mortality prediction in cardiac surgery using artificial neural networks

Johan Nilsson, MD, PhD,[a] Mattias Ohlsson, PhD,[b] Lars Thulin, MD, PhD,[a] Peter Höglund, MD, PhD,[c] Samer A.M. Nashef, FRCS,[d] and Johan Brandt, MD, PhD[a]

See related editorial on page 8.

Supplemental material is available online.

Objective: The artificial neural network model is a nonlinear technology useful for complex pattern recognition problems. This study aimed to develop a method to select risk variables and predict mortality after cardiac surgery by using artificial neural networks.

Methods: Prospectively collected data from 18,362 patients undergoing cardiac surgery at 128 European institutions in 1995 (the European System for Cardiac Operative Risk Evaluation database) were used. Models to predict the operative mortality were constructed using artificial neural networks. For calibration a sixfold cross-validation technique was used, and for testing a fourfold cross-testing was performed. Risk variables were ranked and minimized in number by calibrated artificial neural networks. Mortality prediction with 95% confidence limits for each patient was obtained by the bootstrap technique. The area under the receiver operating characteristics curve was used as a quantitative measure of the ability to distinguish between survivors and nonsurvivors. Subgroup analysis of surgical operation categories was performed. The results were compared with those from logistic European System for Cardiac Operative Risk Evaluation analysis.

Results: The operative mortality was 4.9%. Artificial neural networks selected 34 of the total 72 risk variables as relevant for mortality prediction. The receiver operating characteristics area for artificial neural networks (0.81) was larger than the logistic European System for Cardiac Operative Risk Evaluation model (0.79; $P = .0001$). For different surgical operation categories, there were no differences in the discriminatory power for the artificial neural networks ($P = .15$) but significant differences were found for the logistic European System for Cardiac Operative Risk Evaluation ($P = .0072$).

Conclusions: Risk factors in a ranked order contributing to the mortality prediction were identified. A minimal set of risk variables achieving a superior mortality prediction was defined. The artificial neural network model is applicable independent of the cardiac surgical procedure.

From the Departments of Cardiothoracic Surgery[a] and Theoretical Physics,[b] Lund University; Competence Centre for Clinical Research,[c] Lund University Hospital, Lund, Sweden; and Papworth Hospital,[d] Cambridge, UK.

Supported by grants from the Swedish Heart Lung Foundation, Anna Lisa and Sven-Eric Lundgrens foundation for medical research, and by computer resources from Lunarc at Lund University.

Received for publication Sept 1, 2005; revisions received Dec 19, 2005; accepted for publication Dec 29, 2005.

Address for reprints: Dr Johan Nilsson, Department of Cardiothoracic Surgery, Heart and Lung Center, Lund University Hospital, SE 221 85 LUND, Sweden (E-mail: johan.nilsson@thorax.lu.se).

J Thorac Cardiovasc Surg 2006;132:12-9

0022-5223/$32.00

Copyright © 2006 by The American Association for Thoracic Surgery

doi:10.1016/j.jtcvs.2005.12.055

Preoperative evaluation of a patient's surgical risk is an important component in cardiac surgery. Risk stratification can provide patients and their families with insight into the existent risk of complications and mortality and guide the selection of cases for surgery versus alternative, nonsurgical therapies. It can

Abbreviations and Acronyms

ANN = artificial neural network
CABG = coronary artery bypass grafting
CI = confidence interval
ROC = receiver operating characteristics
SD = standard deviation

also predict the need for hospital care resources in cardiac surgery.[1] During the last decades, several scoring systems to calculate the mortality risk before the surgery have been developed.[2-5]

Most risk scoring systems have been created using a biostatistical method based on a generalized linear model with assumptions of linear relationship. Artificial neural networks (ANNs) work in a nonlinear fashion, which may better describe the interaction between health risk factors. ANNs have been used in classification and diagnostic prediction of cancer[6] and electrocardiogram interpretation,[7] among others. Some studies in clinical medicine have demonstrated superiority of the classification or prediction by ANNs compared with other statistical models.[8] In the field of cardiac surgery, only a few studies using ANNs have been published, and the results have been ambiguous.[9-14]

To select risk variables for a model, significance testing (*P* values) is the most common methodology, but this does not assess the importance of the individual variable.[15] On the other hand, ANNs may be used for both variable selection and ranking of individual variables in order of importance.[15] For example, this methodology has been employed to select and minimize a large number of gene expression levels used in cancer classification, with excellent results.[16]

This study aimed to systematically evaluate the accuracy and performance of ANNs to select and rank the most important risk factors for operative mortality in cardiac surgery by using high-performance computer clusters.

Methods

Database

The database used in the present study was that of the multinational European System for Cardiac Operative Risk Evaluation (EuroSCORE) cardiac surgical project. This was a prospective study to assess risk factors for operative mortality, defined as death within 30 days after the operation or within the same hospital admission,[17] and to construct a risk stratification system.[5] The database included 97 risk factors from all patients who underwent cardiac surgery in 128 centers from 8 European countries from September to December 1995. The data collection, quality checks, and validation have been described by Roques and colleagues.[17] A local database including risk factors for adult patients undergoing cardiac surgery at the Lund University Hospital between January 1996 and February 2001 was used to further evaluate the developed ANN risk model by blind testing.

Patients and Study Design

From the 97 original EuroSCORE variables, a subset of 72 variables was selected (Tables 1 and 2). This was done by excluding variables closely linked to other variables and data collected intraoperatively (ie, number of conduits and number of distal coronary anastomoses). Patients with a missing value in any mandatory variable (age, gender, or surgical procedure) or outcome (operative mortality) were excluded from analysis. Imputation was used to substitute missing values in the other variables with the statistical mode for categorical variables and the mean for continuous variables.[11]

Calibration of the ANN Model and Selection of Risk Factors Utilized for Mortality Prediction

An ensemble approach was used where several ANNs were combined into a single prediction model. The individual members of the ensemble were standard multilayer perceptrons with 1 hidden layer and 1 output node that was used to encode the operative mortality.[18] The model selection was performed using a sixfold cross-validation procedure (Figure 1). To select the most important risk variables and to minimize the number of variables included in the final model, a ranking of risk variables was performed.[15]

Performance and Accuracy

The performance and accuracy of the ANN model was compared with the logistic EuroSCORE model[19] and a logistic model. The final prediction models were tested on patients not previously exposed to the models by using a fourfold cross-testing technique (Figure 1).

Statistical Analysis

Mean values (± standard deviation [SD]) were used to describe continuous variables, and frequencies were calculated for categorical variables. Logistic regression analysis was performed to obtain the coefficients for the risk variables included in the logistic model as described by Hosmer and Lemeshow.[20]

To compare the number of correctly classified patients by ANNs versus the logistic EuroSCORE, a proportion test was used. Effective odds ratio for the risk variables were determined as described by Lippmann and Shahian.[11] The 95% confidence intervals (CIs) for both the odds ratio and the output from the ANNs were calculated using the bootstrap technique.[11,21]

Receiver operating characteristics (ROC) curves were used to describe the performance and predictive accuracy for the models.[22] The area (with 95% CI) under the curve was used as a quantitative measure of the ability of the risk predictor models to distinguish between survivors and nonsurvivors. To compare the areas under the resulting ROC curves, the nonparametric approach described by DeLong and coworkers[23] was used.

Computer Cluster and Software

High-performance computing clusters were used to train and evaluate the ANNs. The ANN calibration and analyses were performed with MatLab 7 (2005), Neural Network Toolbox (MathWorks, Natick, Mass). Graphs and statistical analyses were performed using the Intercooled Stata version 9.0 (2005) statistical package (StataCorp LP, College Station, Tex).

Download English Version:

https://daneshyari.com/en/article/2985298

Download Persian Version:

https://daneshyari.com/article/2985298

Daneshyari.com