

Editorial Manager(tm) for Medical & Biological Engineering & Computing
Manuscript Draft

Manuscript Number: MBEC2449R2

Title: Prediction of persistence of combined evidence-based cardiovascular medications in patients with acute coronary syndrome after hospital discharge using neural networks

Article Type: Original article

Keywords: Acute coronary syndrome, evidence-based medical therapy, long-term persistence, artificial neural networks.

Corresponding Author: Valérie Bourdès, MD, MPH

Corresponding Author's Institution:

First Author: Valérie Bourdès, MD, MPH

Order of Authors: Valérie Bourdès, MD, MPH; Jean Ferrières, MD; Jacques Amar, MD; Elisabeth Amelineau, MD; Stéphane Bonnevey, PhD; Maryse BERLION, Ph.D; Nicolas Danchin, MD

Abstract: In the PREVENIR-5 study, artificial neural networks (NN) were applied to a large sample of patients with recent first acute coronary syndrome (ACS) to identify determinants of persistence of evidence-based cardiovascular medications (EBCM: antithrombotic + beta-blocker + statin + angiotensin converting enzyme inhibitor-ACEI and/or angiotensin-II receptor blocker-ARB). From October 2006 to April 2007, 1811 general practitioners recruited 4850 patients with a mean time of ACS occurrence of 24 months. Patient profile for EBCM persistence was determined using automatic rule generation from NN. The prediction accuracy of NN was compared with that of logistic regression (LR) using Area Under Receiver-Operating Characteristics-AUROC. At hospital discharge, EBCM was prescribed to 2132 patients (44%). EBCM persistence rate, 24 months after ACS, was 86.7%. EBCM persistence profile combined overweight, hypercholesterolemia, no coronary artery bypass grafting and low educational level (Positive Predictive Value=0.958). AUROC curves showed better predictive accuracy for NN compared to LR models.

Response to Reviewers: From Dr BOURDES Valérie
Société THEMIS
Bioparc - 60 avenue Rockefeller
69008 LYON, FRANCE

To Dr SPAAN Jos
Editor in Chief
Medical & Biological Engineering & Computing

April 21st, 2011



Dear Dr Spaan,

We thank you for accepting our Research Article titled "Prediction of persistence of combined evidence-based cardiovascular medications in patients with acute coronary syndrome after hospital discharge using neural networks" by Valérie Bourdès, Jean Ferrières, Jacques Amar, Elisabeth Amelineau, Stéphane Bonnevey, Maryse Berlion, and Nicolas Danchin.

We carefully read the last comments of the reviewers and we revised the manuscript taking into account all the changes requested by the reviewers.

The replies to the reviewers, comment by comment, are given below and we submit our revised manuscript.

Yours sincerely,

Valérie Bourdès, MD, MPH

REPLY TO REVIEWER:

COMMENTS:

Reviewer #1:

* Methods / Study Design [Line 14]:
"...Good Epidemiological Practice..." requires a reference.

Authors: the reference has been added in Section "Methods / Study design", Line 14 and in the reference list (reference 21).

* Methods / Variable Selection [Line 18]:
The authors used the error of the training set as the objective criterion of the GA. Why not use the test set? Using the training set gives limited information as to the NN's generalisation capability and so might be misleading. A brief comment here would suffice.

Authors: The aim of this process is not to build the NN, but only to select the best set of variables. At each step of this process (each set of variables and each generation: $100 \times 100 = 10000$ step) a NN is build and an error is calculated. Then, with the best set of variables, a NN is build according to a classical learning process with test and training sets.

* Methods / Variable Selection [Line 19]:
The authors use 100 population over 100 generations. Can the authors justify these numbers? Did the population sufficiently converge over 100 generations?

Authors: There's no convergence reason; these choices were only guided by the time-consuming process.

* Methods / Variable Selection:

Can the authors give the cross-over and mutation coefficient values?

Authors: The single-point crossover rate was set to 0.8 and the mutation rate was set to 0.1.

Reviewer #2:

From my point of the view the authors have accomplished the reviewers' suggestions and have incorporated all necessary changes.

Prediction of persistence of combined evidence-based cardiovascular medications in patients with acute coronary syndrome after hospital discharge using neural networks

Valérie Bourdès^a, Jean Ferrières^{b,c}, Jacques Amar^c, Elisabeth Amelineau^d, Stéphane Bonnevey^e, Maryse Berlion^a, Nicolas Danchin^f

^aTHEMIS-ICTA group, Lyon, France; ^bInserm U558, University Hospital, Toulouse, France; ^cRangueil Hospital, Toulouse, France; ^dBristol-Myers Squibb, Rueil Malmaison, France; ^eLaboratoire ERIC, Ecole Polytechnique Universitaire, Villeurbanne, France; ^fGeorges Pompidou European Hospital, Paris, France.

Corresponding author:

Valérie BOURDES, MD, MPH

THEMIS-ICTA Group

Bioparc - 60 avenue Rockfeller

69008 Lyon, FRANCE

Tel.: +33 4 78 01 42 30; Fax: +33 4 78 01 67 29; E-mail: valerie.bourdes@themis-rd.fr

The total number of words of the manuscript, including entire text from title page to figure legends: 5436

The number of words of the abstract: 149

The number of figures: 1

The number of tables: 4

Abstract

In the PREVENIR-5 study, artificial neural networks (NN) were applied to a large sample of patients with recent first acute coronary syndrome (ACS) to identify determinants of persistence of evidence-based cardiovascular medications (EBCM: antithrombotic + beta-blocker + statin + angiotensin converting enzyme inhibitor-ACEI and/or angiotensin-II receptor blocker-ARB). From October 2006 to April 2007, 1811 general practitioners recruited 4850 patients with a mean time of ACS occurrence of 24 months. Patient profile for EBCM persistence was determined using automatic rule generation from NN. The prediction accuracy of NN was compared with that of logistic regression (LR) using Area Under Receiver-Operating Characteristics-AUROC. At hospital discharge, EBCM was prescribed to 2132 patients (44%). EBCM persistence rate, 24 months after ACS, was 86.7%. EBCM persistence profile combined overweight, hypercholesterolemia, no coronary artery bypass grafting and low educational level (Positive Predictive Value=0.958). AUROC curves showed better predictive accuracy for NN compared to LR models.

Keywords: Acute coronary syndrome, evidence-based medical therapy, long-term persistence, artificial neural networks.

Introduction

Acute coronary syndrome (ACS) includes life-threatening clinical conditions ranging from unstable angina to non-ST-elevation myocardial infarction (NSTEMI) and ST-elevation myocardial infarction (STEMI) that are a major cause of emergency medical care and hospitalization in industrialized countries [17, 23]. Patients with ACS after the initial phase carry a high risk of recurrence of ischemic events. Therefore, active secondary prevention is an essential element of long-term management. European clinical practice guidelines (European Society of Cardiology) support the use of four major classes of drugs, i.e., angiotensin-converting enzyme inhibitors (ACEI) and/or angiotensin II receptor blockers (ARB), beta blockers, antithrombotic agents and statins for long-term treatment of patients after an acute coronary event [5, 32, 34]. Unfortunately, there is evidence that these therapies are neither consistently prescribed when appropriate nor adhered to by patients in daily clinical practice [27]. In a recent analysis of real-world use of secondary prevention therapies over 2003-2004 within the 90 days following an hospitalization for ACS, Lee *et al.* [27] found exposure rates of 52% for ACEIs or ARBs, 64% for beta-blockers and 63% for statins in a cohort of 1135 patients. Similar results were obtained in three European observational studies [8, 14, 19]. Long-term persistence of evidence-based cardiovascular medications (EBCM) remains a key concern to achieve optimal outcomes. In a survey of 1326 patients with coronary artery disease assessing the adherence to EBCM by comparing baseline prescription at hospital discharge to self-reported medical regimen at 1 year, Kulkarni *et al.* [26] reported that only 54% of patients were adherent to all of their initial medications. In 1700 patients hospitalized for ACS in 2006 in France, the four-drug combination therapy had been prescribed in 46% of patients at hospital discharge, 80% of them were still taking the combination 14 months after hospital discharge [3]. Defining predictors of long-term EBCM persistence has implications for the disease management approach of individual patient. Traditional statistical approaches such as logistic or Cox regressions, commonly used for prediction analyses in cardiology, tend to select only

variables which have a high level of linear correlation with the outcome variable. Complex non-linear relationships between independent and dependent variables can be accurately detected by artificial neural networks (NNs) [9]. Artificial NNs are biologically inspired by computer programs designed to simulate the way in which the human brain processes information [31]. NN is formed from hundreds of single units, artificial neurons or processing elements, connected with coefficients (weights), which constitute the neural structure and are organized in layers [1]. NNs present the main advantage not to be based on "*a priori*" assumptions and to allow detection of links between factors that conventional statistical techniques may not be able to detect [1, 7, 33]. In studies carried out in the cardiovascular field, NNs provided a better predictive accuracy than did traditional statistical techniques [4, 6, 35].

In the present French, cross-sectional, observational PREVENIR-5 study, NNs (Multi-Layer Perceptron) have been applied to a large sample of patients with recent ACS to identify persistence factors of EBCM. The prediction accuracy of NN was compared with that of logistic regression (LR) using classification performance indices, i.e., Area Under Receiver-Operating Characteristics (AUROC), sensitivity, specificity, accuracy and positive predictive value (PPV).

Methods

Study design

Male or female adult patients hospitalized for a first episode of ACS (NSTEMI, STEMI or unstable angina) during 2004 or 2005, and followed routinely by a general practitioner (GP), were included in this study. No exclusion criteria were applied except for current participation in a clinical trial. The physicians were randomly selected from a comprehensive national database of French GPs. The sample was regionally stratified to ensure the representativeness of the study with regard to French medical practice. The study was proposed by mail to 39665 GPs and 3122 (7.9%) gave their consent to participate and signed the financial agreement. Each participating GP was asked to include the first three consecutive patients, who fulfilled the eligibility criteria over a three-month period. The following data were prospectively collected by the GP at the inclusion visit (day of consultation): patients' and ACS characteristics, medications at hospital discharge and at inclusion visit, cardiovascular risk factors, last biological measurements, concomitant cardiovascular and other diseases, and concomitant treatments. The study met the quality standard described in the declaration of Helsinki and the Good Epidemiological Practices [21] and was approved by local legal authorities ("French Information Technology and Liberties Commission" and "French Medical Committee"). Participation in the study was voluntary. This observational study did not modify the clinical practices of the physicians.

Clinical endpoints

The endpoints were to identify key persistence factors of EBCM in ACS patients using NN and LR and to compare the two models prediction performances. Persistence was defined as current use of EBCM prescribed at hospital discharge and non persistence as the discontinuation of at least one of the drugs the day of consultation. The analysis was performed on the subgroup of patients who received the four-drug combination at hospital discharge (N=2132), i.e., any antithrombotic agent + beta-blocker + statin + ACEI and/or ARB (EBCM).

Variable selection with neural networks

Neural network constructions were carried out using the Statistica Neural Networks software release 7.1. LR was performed with SAS Software 8.2. Two types of NN models were built, one using the LR variable selection (NN-varLR) and the other using the NN variable selection (NN-varNN). Therefore, three models were compared (NN-varLR, NN-varNN and LR). Three different methods were used to select the most significant variables for each NN model: forward and backward stepwise feature selections and a genetic input selection algorithm. Forward selection consists in choosing the most predictable variable then checks for a second variable that, added to the first, most improves the model; this process is repeated until either all variables have been selected or no further improvement is made. Backward stepwise feature selection is the reverse process: it starts with all the variables and then removes a variable at each stage which less degrades the model. Genetic algorithm [16] selection is a heuristic seeking the optimal set of input variables. This heuristic builds a model by a succession of artificial transformations (mutation, crossover, and selection) from an initial population of variables sets. Each of our genetic selections was made from a population of 100 individuals (one individual corresponds to one set of variables) on 100 generations. Each set of variables corresponds to a binary string where a 0 indicates that the variable is not in the set of variables, and a 1 indicates that the variable is in this set. This set is tested with the help of a neural network, and the objective function is the error of this neural network on a training set. With our choice (100 individuals and 100 generations), it performs 10,000 evaluations of sets of variables. For example, the selection of one set of variables is about 60 times longer with the genetic algorithm than the backward selection, but genetic algorithms are well suited for feature selection as there is a large number of possible variables. Because of their differences and complementarities, we decided to combine these three methods to select the inputs of our NN models.

With a view to improving generalization capability of networks and to decreasing the network size and execution size, a penalty can be used to penalize the large sets of variables. In this way,

a penalty parameter is multiplied by the number of selected variables and added to the error level. Following different analyses, we finally used a small penalty equal to 0.0001 with half of selection algorithms and no penalty with the other half selection algorithms. Each method calculates the value of the set of covariates selected at each step while building the NN. We have chosen to perform each method 40 times: 20 without penalty and 20 with a penalty equal to 0.0001. Indeed, 120 (3×40) selections were performed. Covariates were kept as an input for the final model if it was selected in at least 80% of the 120 individual models; this value of 80% leads to the inclusion of a reasonable number of variables regarding the complexity of the NN models. Table 1 displays the selected variables for LR and NN for EBCM subgroup.

Neural networks modelling

The Multi Layer Perceptron (MLP) NN was applied as a classifier, with a single-layer of hidden nodes using hyperbolic activation functions [20]. The number of neurons on the hidden layer was determined according to the number and the nature of variables at the entry. Weights and bias of NN were determined by training with a two-phase procedure. The first phase is a quite short burst of backpropagation, with a moderate training rate. The second phase is a longer run of conjugated gradient descent, a much more powerful algorithm, which is less likely to encounter convergence problems than otherwise due to the use of backpropagation first. During this learning process, the weights in a MLP are adjusted using least squares fitting together with the training two-phase procedure to minimize a root mean square error function. In order to interpret the network outputs as probabilities and to make them comparable to the results of logistic regression, we used a cross entropy error function to adjust weights. This cross entropy function is specially designed for classification problems where it is used in combination with hyperbolic activation function.

A continuous input value is prescaled to a range between 0 and 1; a two-state nominal variable, which corresponds to one entry of the neural network, is represented by transformation into a numeric value (e.g., “Gender” = 0 or 1); a many-state nominal variable is recoded into as many

binary entries as modalities (e.g., “Age” is a 4-state nominal variable which correspond to 4 inputs). The NN-varLR was composed of: (a) one input layer corresponding to the 11 covariates (given by the Logistic Regression). Seven covariates with two modalities correspond to a binary entry. The other covariates are recoded into as many entries as modalities (one covariate with 3 modalities, one with 4 modalities and two covariates with 5 modalities). The network built that way has 24 binary entries; (b) one hidden layer composed of 6 neurons with hyperbolic activation function and (c) one output layer composed of the one neuron with logistic activation function. The NN-varNN was composed of (a) one input layer corresponding to the 8 covariates (given by the neural selection). Three covariates correspond to a binary entry and the other covariates are recoded into as many entries as modalities (2 covariates with 3 modalities, one with 4 modalities and two with 5 modalities). The network built that way has 23 binary entries; (b) one hidden layer composed of 8 neurons with hyperbolic activation function and (c) one output layer composed of the one neuron with logistic activation function.

In order to avoid over-fitting and to obtain networks with a strong capacity of generalization, we divided data randomly in two datasets by cross validation: a learning set to build the models and a testing set for the evaluation (not used for construction). The learning set was composed of 70% of the total random population (testing set with 30% of the remaining population).

Comparison between logistic regression and neural networks

The predictive accuracy of the three models, LR, NN with inputs of LR selection (NN-varLR) and NN with inputs of NN selection (NN-varNN), was assessed using AUROC. The AUROC is a good measure of the overall predictive accuracy of an analytic tool. It represents a plot of sensitivity *versus* (1 - specificity). Sensitivity measures the fraction of positive cases classified as positive and specificity measures the fraction of negative cases classified as negative. The positive predictive value (PPV) corresponds to the number of true positives divided by the sum of true positives and false positives.

Rule extraction methodology

To identify patients profiling, the Orthogonal Search-based Rule Extraction (OSRE) methodology was used to extract rules from NN [12, 13]. OSRE is a rule extraction algorithm that efficiently extracts comprehensive rules from smooth models such as those created by NNs. This method, which is a refinement of RULENEG algorithm, constructed a hierarchical list of rules from the best to the less good: decreasing order of sensitivity and PPV. This construction was performed on the learning set then validated on the remaining data.

Results

From October 2006 to April 2007, 5111 patients were enrolled by 1838 active GPs. A total of 261 patients (5.1%) were excluded from the analyzed population due to non respect of inclusion criteria (n=9), patients hospitalized out of the 2004-2005 period (n=77), patients seen in consultation before the beginning of the survey (n=175). Overall, 4850 patients were enrolled by 1811 GPs with a mean time (\pm SD) of occurrence of first ACS of 24 ± 7 months (range, 10.8 to 39.3 months).

Patients' characteristics

Patients' characteristics (N=4850) are described in Table 2. The majority was male (80.1%) and the mean age was 64 years. ACS included STEMI (43.7%), unstable angina (33.1%), NSTEMI (18.4%) or unknown (4.8%). The majority of patients were regularly followed by their GP with quarterly, monthly or bimonthly visit rates for 40.7%, 36.5% and 19.1% of patients, respectively. Moreover, 96.1% of the patients were also followed by a specialist, mainly a private (65.0%) or hospital (34.6%) cardiologist.

EBCM persistence

At hospital discharge, EBCM was prescribed to 2132 patients (44.0%). At inclusion visit, EBCM was persistent in 1849 patients (86.7%). Persistence rates of antithrombotic agents, ACEI and/or ARB, beta-blockers, statins were 99.4%, 94.3%, 92.9% and 98.1%, respectively (Table 3). Non persistence was mostly due to the discontinuation of beta-blockers (7.1%) and ACEI and/or ARB (5.7%). The main reason for treatment discontinuation reported by the practitioner was intolerance particularly for statins (48.8%), ACEI and/or ARB (36.1%) and beta-blockers (38.4%) (Table 4).

Persistence profile

The rule extracted from the OSRE algorithm was:

- $25 < \text{body mass index-BMI} < 30 \text{ kg/m}^2$ and

- low educational level, and
- hypercholesterolemia, and
- no coronary by-pass grafting (CABG) for the first ACS.

Rules generated indicated that patients with the association of overweight, hypercholesterolemia, no CABG for their first ACS, and low educational level represented a relevant persistence profile with a PPV of 0.958.

Comparison between NN and LR

The AUROC curves of persistence showed better predictive accuracy for NN models compared with LR model for EBCM, whatever the variables selection with AUC of 0.78, 0.80 and 0.69- for NN-varLR, NN-varNN and LR, respectively (Figure 1).

Discussion

This large-scale, epidemiological and cross-sectional study reflects real-life general medical practice regarding the long-term therapeutic management of ACS in France. NN models successfully identified persistence profile of patients with a high PPV, confirming their relevance as predictive tools. The NN model showed a higher prediction performance than LR and should be considered as a helpful tool in medical decision support by identifying specific subgroups of patients.

The MLP neural network is the most widely applied to real world problems in medical diagnosis and prediction [20]. One of the criticisms towards NN is that their process inside is unknown and some authors consider them as "black boxes". To prevent this criticism and to enhance the NN selection process, we decided to use several variables selection techniques perfectly coded and to use one penalty for some selection methods. Using these three different techniques for the variables selection may be criticized as a time-consuming process. However, it was a guarantee for avoiding over-fitting, which is the main limit of the NN [1] and leading to a gain in selection improvement.

The PREVENIR-5 survey confirms the low prescription of combined EBCM at hospital discharge in patients with ACS, despite the well documented efficacy of aspirin, beta-blockers, statins and ACEIs as secondary prevention of coronary artery disease [29]. Only 44% of patients received treatment with an antithrombotic, a beta-blocker, a statin and an ACEI and/or ARB. This prescription rate was consistent with that observed in the recent PREVENIR-4 observational study carried out in 2006, in which 46% of patients received the combination therapy (β -blocker, antithrombotic, statin and ACEI) [3]. However, it seems to be better than results reported by Danchin *et al.* [10] in French patients admitted to intensive care units in 2000 for acute myocardial infarction; only 27% of these patients received the quadritherapy at hospital discharge. Yan *et al.* [36] reported a significant increase in the discharge use of all 4 classes of medications over time with 29% and 52% of patients prescribed optimal medical

therapy in the Canadian ACS I (September 1999-June 2001) and ACS II (October 2002-December 2003) registries, respectively. Some factors associated with lower rate prescription have been identified: elderly patients, female gender, missing of low density lipoprotein-cholesterol measurement, history of peripheral artery disease or stroke, and finally the difficulty of observance [28].

The rates of persistence for EBCM as well as for each compound separately were high in PREVENIR-5 patients as already observed in the most recent studies. In PREVENIR-4 survey carried out in 2006, 80% of coronary patients were still taking the quadritherapy 14 months on average after hospital discharge [3]. When analyzed separately, beta-blockers, antithrombotics, statins and ACEIs were persistent over the 14-month period for 92%, 97%, 96% and 86% of patients, respectively [3]. In 13830 patients from the GRACE registry, therapy was maintained 6 months after hospital discharge for myocardial infarction or unstable angina for 92% of patients taking aspirin on discharge, 88% of those taking beta-blockers, 80% of those taking ACEIs and 87% of those taking statins [11]. The high persistence rates for PREVENIR-5 patients can be explained by the fact that almost all the patients were regularly followed by both GPs and cardiologists. A regular follow-up by the attending physician and/or specialist suggests effective physician-patient interactions and efficient communication is also thought to be an important determinant of adherence to treatment [11]. A patient who clearly understands the usefulness of a drug and the potential consequences of not taking it is more likely to continue the treatment. Moreover, half of the patients were less than 65 years old and few patients presented concomitant disease. Patients were started on medication at the time of hospital discharge and adherence is generally superior in this setting to when therapies are initiated in the outpatient setting [15]. These rates of persistence were higher than those reported in the early to mid 1990s suggesting an improvement in adherence to treatment over the past decade [18, 24, 25, 30]. This gradual improvement may be explained by the diffusion of the findings of major randomized trials into medical practice and a good follow-up of the clinical guidelines.

Using NNs, the associated predictive factors correlated with a persistence profile were absence of CABG, overweight, hypercholesterolemia and low educational level for patients receiving combined treatment at hospital discharge. The cardiovascular risk factors particularly hypercholesterolemia were frequently shown to be associated with treatment persistence, especially for statin therapy [22]. Conversely to our results, adherent patients were significantly more likely to have a higher educational status in an American observational study [26]. Low educational status may be a marker of limited financial resources to afford medications in United States but not in France where patients benefit from a different health care system. In France, cardiovascular risk factors such as overweight or hypercholesterolemia are frequently associated with low educational status. Although prior studies have suggested age-influenced adherence [2], our study did not find an association between patient age and persistence whatever the model used.

Our study had several limitations, the most significant being inherent to the cross-sectional design of the study. As a result it was not possible to examine the time relationship between patient factors and persistence. However, no impact on NN or LR models was expected.

In conclusion, long-term persistence of EBCM prescribed after hospital discharge for a first ACS was high. The neural approach and rule extraction allowed the identification of patient profiling for persistence of EBCM, i.e., absence of CABG, overweight, hypercholesterolemia and low educational level. This large epidemiological French study shows that NN should be recognized as powerful predictive tools, to be considered routinely alongside standard LR.

Acknowledgments

This work was financially supported by Bristol Myers Squibb, France. The authors thank Prof. Paulo Lisboa and his team (Liverpool John Moore University, United Kingdom) for the rule extraction methodology and the data management and statistics department of ICTA-PM (Dijon, France).

References

1. Agatonovic-Kustrin S, Beresford R (2000) Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J Pharm Biomed Anal* 22:717-727
2. Ali RC, Melloni C, Ou FS, Schmader K, Ohman EM, Roe MT, Peterson ED, Alexander KP (2009) Age and persistent use of cardiovascular medication after acute coronary syndrome: results from medication applied and sustained over time. *JAGS* 57:1990-1996
3. Amar J, Ferrières J, Cambou JP, Amelineau E, Danchin N (2008) Persistence of combination of evidence-based medical therapy in patients with acute coronary syndromes. *Arch Cardiovasc Dis* 101:301-306
4. Baldassarre D, Grossi E, Buscema M, Intraligi M, Amato M, Tremoli E, Pustina L, Castelnuovo S, Sanvito S, Gerosa L, Sirtori CR (2004) Recognition of patients with cardiovascular disease by artificial neural networks. *Ann Med* 36:630-640
5. Bassand JP, Hamm CW, Ardissino D, Boersma E, Budaj A, Fernández-Avilés F, Fox KA, Hasdai D, Ohman EM, Wallentin L, Wijns W (2007) Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes. *ESC Guidelines. Eur Heart J* 28:1598-1660
6. Bigi R, Gregori D, Cortigiani L, Desideri A, Chiarotto FA, Toffolo GM (2005) Artificial neural networks and robust Bayesian classifiers for risk stratification following uncomplicated myocardial infarction. *Int J Cardiol* 101:481-487
7. Bourdès V, Bonnevey S, Lisboa PJ, Defrance R, Pérol D, Chabaud S, Bachelot T, Gargi T, Négrier S (2010) Comparison of artificial neural network with logistic regression as classification models for variable selection for prediction of breast cancer patient outcomes. *Advances in Artificial Neural Systems 2010*, Article ID 309841.
8. Carruthers KF, Dabbous OH, Flather MD, Starkey I, Jacob A, Macleod D, Fox KA; GRACE Investigators (2005) Contemporary management of acute coronary syndromes:

does the practice match the evidence? The global registry of acute coronary events (GRACE). *Heart* 91:290-298

9. Cross SS, Harrison FR, Kennedy RL (1995) Introduction to neural networks. *Lancet* 346:1075-1079
10. Danchin N, Cambou JP, Hanania G Kadri Z, Genès N, Lablanche JM, Blanchard D, Vaur L, Clerson P, Guéret P, USIC 2000 investigators (2005) Impact of combined secondary prevention therapy after myocardial infarction: data from a nationwide French registry. *Am Heart J* 150:1147-1153
11. Eagle KA, Kline-Rogers E, Goodman SG Gurfinkel EP, Avezum A, Flather MD, Granger CB, Erickson S, White K, Steg PG (2004) Adherence to evidence-based therapies after discharge for acute coronary syndromes: an ongoing prospective, observational study. *Am J Med* 117:73-81
12. Etchells TA, Lisboa PJ (2006) Orthogonal search-based rule extraction (OSRE) for trained neural networks: a practical and efficient approach. *IEEE Transactions on Neural Networks* 17:374-384
13. Etchells TA, Harrison MJ (2006) Orthogonal search-based rule extraction for modelling the decision to transfuse. *Anaesthesia* 61:335-338
14. EUROASPIRE II Study Group (2001) Lifestyle and risk factor management and use of drug therapies in coronary patients from 15 countries; principal results from EUROASPIRE II Euro Heart Survey Programme. *Eur Heart J* 22:554-572
15. Feely J, Chan R McManus J, O'Shea B (1999) The influence of hospital-based prescribers on prescribing in general practice. *Pharmacoeconomics* 16:175-181
16. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Kluwer Academic Publishers, Boston, MA
17. Grech ED, Ramsdale DR (2003) Acute coronary syndrome: unstable angina and non-ST segment elevation myocardial infarction. *Brit Med J* 326:1259-1261

18. Harder S, Thürmann P, Thierolf C, Klepzig H (1998) Prescription of cardiovascular drugs in outpatient care: a survey of outpatients in a German university hospital. *Int J Clin Pharmacol Ther* 36:195-201
19. Hasdai D, Behar S, Wallentin L, Danchin N, Gitt AK, Boersma E, Fioretti PM, Simoons ML, Battler A (2002) A prospective survey of the characteristics, treatments and outcomes of patients with acute coronary syndromes in Europe and the Mediterranean basin; the Euro Heart Survey of Acute Coronary Syndromes (Euro Heart Survey ACS). *Eur Heart J* 23:1190-1201
20. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward neural networks are universal approximators. *Neural Networks* 2:259-366
21. IEA guidelines for proper conduct in epidemiologic research (2007) Good Epidemiological Practice (GEP). November, 2007. <http://www.dundee.ac.uk/iea>
22. Jackevicius CA, Mamdani M, Tu JV (2002) Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *JAMA* 288:462-467
23. Joint European Society of Cardiology/American College of Cardiology Committee (2000) Myocardial infarction redefined--a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *Eur Heart J* 21:1502-1513
24. Jones JK, Gorkin L, Lian JF, Staffa JA, Fletcher AP (1995) Discontinuation of and changes in treatment after start of new courses of antihypertensive drugs: a study of a United Kingdom population. *Brit Med J* 311:293-295
25. Krumholz HM, Vaccarino V, Ellerbeck EF, Kiefe C, Hennen J, Kresowik TF, Gold JA, Jencks SF, Radford MJ (1997) Determinants of appropriate use of angiotensin-converting enzyme inhibitors after acute myocardial infarction in persons \geq 65 years of age. *Am J Cardiol* 79:581-586
26. Kulkarni SP, Alexander KP, Lytle B, Heiss G, Peterson ED (2006) Long-term adherence with cardiovascular drug regimens. *Am Heart J* 151:185-191

27. Lee HY, Cooke CE, Robertson TA (2008) Use of secondary prevention drug therapy in patients with acute coronary syndrome after hospital discharge. *J Manag Care Pharm* 14:271-280
28. Philippe F, Cambou JP, Danchin N, Thomas D (2005) Changes in the prescription of cardiovascular prevention drugs in France between 1995 and 2003: factors influencing the gap between evidence base medicine and clinical practice. *Ann Cardiol Angeiol* 54 (Suppl 1):S30-S36
29. Ramanath VS, Eagle KA (2007) Evidence-based medical therapy of patients with acute coronary syndromes. *Am J Cardiovasc Drugs* 7:95-116
30. Roe CM, Motheral BR, Teitelbaum F, Rich MW (2000) Compliance with and dosing of angiotensin-converting-enzyme inhibitors before and after hospitalization. *Am J Health Syst Pharm* 57:139-14
31. Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 65:386-408
32. Silber S, Albertsson P, Fernandez-Avilès F, Camici PG, Colombo A, Hamm C, Jørgensen E, Marco J, Nordrehaug JE, Ruzyllo W, Urban P, Stone GW, Wijns W; Task Force for Percutaneous Coronary Interventions of the European Society of Cardiology (2005) Guidelines for percutaneous coronary interventions. ESC Guidelines. *Eur Heart J* 26:804-807
33. Tu JV (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol* 49:1225-1231
34. Van de Werf F, Ardissino D, Betriu A, Cokkinos DV, Falk E, Fox KA, Julian D, Lengyel M, Neumann FJ, Ruzyllo W, Thygesen C, Underwood SR, Vahanian A, Verheugt FW, Wijns W; Task Force on the Management of Acute Myocardial Infarction of the European Society of Cardiology (2003) Management of acute myocardial infarction in patients presenting with ST-segment elevation. ESC Guidelines. *Eur Heart J* 24:28-66

35. Voss R, Cullen P, Schulte H, Assmann G (2002) Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Munster Study (PROCAM) using neural networks. *Int J Epidemiol* 31:1253-1562
36. Yan AT, Yan RT, Tan M Huynh T, Soghрати K, Brunner LJ, DeYoung P, Fitchett DH, Langer A, Goodman SG; Canadian ACS Registries Investigators (2007) Optimal medical therapy at discharge in patients with acute coronary syndromes: temporal changes, characteristics, and 1-year outcome. *Am Heart J* 154:1108-1115

Tables

Table 1 Selection of variables for neural networks and logistic regression for EBCM subgroup (N=2132)

Variable	Neural networks	Logistic regression ^{&}
Interventional procedures	93%	X
Number of medications per day at inclusion visit	93%	X
ACS oldness	88%	X
Education level	85%	X
Fasting glycemia	85%	
Body mass index	83%	X
Smoking status	83%	X
Treated diabetes mellitus	80%	
Treated hypercholesterolemia	75%	X
Coronarography results	70%	
Treated hypertension	70%	
Triglycerides	65%	
Oral antidiabetics at inclusion visit	63%	
Gender	53%	
Peripheral vasodilators	45%	X
Accommodation	43%	X
Severe renal failure	43%	
Respiratory failure, chronic obstructive pulmonary disease, asthma	38%	X
Insulin	18%	
Hypnotics at inclusion visit	10%	
Stroke with sequeleae	5%	
Hormone replacement treatment at inclusion visit	3%	X

Abbreviation: ACS, acute coronary syndrome.

Grey sections correspond to selected variables for NN and the X correspond to selected variables for LR.

[&] Univariate analysis with $p < 0.20$ then multivariate analysis using backward selection with $p < 0.10$.

Table 2 Patients and disease characteristics

		Overall patients (N=4850)		
		Mean ± SD	Number of patients	%
Gender	Male		3885	80.1
Age, years		64 ± 12		
	<55 years		1054	21.7
	[55-65[years		1428	29.5
	[65-75[years		1398	28.8
	≥75 years		969	20.0
BMI	Healthy weight		1566	32.3
	Overweight		2361	48.7
	Obesity		919	19.0
Education level	Secondary school		2097	43.3
Current employment			1220	25.2
Acute coronary syndrome	Unstable angina		1604	33.1
	STEMI		2122	43.7
	NSTEMI		892	18.4
	Unknown		232	4.8
Time of occurrence of ACS, months		24 ± 7		
Cardiovascular risk factor	At least one risk factor		4644	95.8
	Treated hypertension		3058	65.9
	Treated hypercholesterolemia		3999	86.1
	Treated diabetes		929	20.0
	Current or former smokers		3047	65.6
Left ventricular ejection fraction	<40%		424	13.8
	≥40%		2646	86.2
Coronarography	Normal or lesion <50%		354	8.4
	One vessel disease		1665	39.6
	Two vessel disease		1324	31.5
	Three vessel disease		865	20.6
Interventional procedures	None		1011	20.9
	PCI		3276	67.6
	CABG		514	10.6
	PCI and CABG		45	0.93
	Other		1	0.02

Abbreviations: ACS, acute coronary syndrome; BMI, body mass index; CABG, coronary artery bypass grafting; NSTEMI, non-ST-elevation myocardial infarction; PCI, percutaneous coronary intervention; SD, standard deviation; STEMI, ST-elevation myocardial infarction.

Table 3 Persistence of treatments at consultation visit

Treatment at consultation visit, n (%) (at least one treatment or combination)	EBCM subgroup at hospital discharge (n=2132)
Antithrombotic agent	2119 (99.4%)
Beta-blocker	1981 (92.9%)
ACEI and/or ARB	2010 (94.3%)
Statin	2091 (98.1%)
Antithrombotic agent + statin	2080 (97.6%)
Antithrombotic agent + beta-blocker + statin	1939 (90.9%)
Antithrombotic agent + beta-blocker + ACEI and/or ARB	1876 (88.0%)
Antithrombotic agent + ACEI and/or ARB + statin	1970 (92.4%)
Antithrombotic agent + beta-blocker + ACEI and/or ARB + statin (EBCM)	1849 (86.7%)

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin II receptor blocker; EBCM, evidence-based cardiovascular medications.

Table 4 Reasons for treatment discontinuation in EBCM subgroup

Reason of non persistence - EBCM	ACEI and/or ARB (n=122)	Beta-blocker (n=151)	Statin (n=41)	Aspirin (n=13)
Bad tolerance	44 (36.1%)	58 (38.4%)	20 (48.8%)	3 (23.1%)
Replacement by another class	9 (7.4%)	22 (14.6%)	5 (12.2%)	3 (23.1%)
No indication	33 (27.0%)	17 (11.3%)	1 (2.4%)	1 (7.7%)
Benefit/risk ratio judged as insufficient	13 (10.7%)	4 (2.6%)	1 (2.4%)	1 (7.7%)
Patient's refusal	2 (1.6%)	6 (4.0%)	4 (9.8%)	0
Contra-indication	5 (4.1%)	17 (11.2%)	2 (4.9%)	2 (15.4%)
Other	15 (12.3%)	25 (16.6%)	6 (14.6%)	2 (15.4%)
Missing data	1 (0.8%)	2 (1.3%)	2 (4.9%)	1 (7.7%)

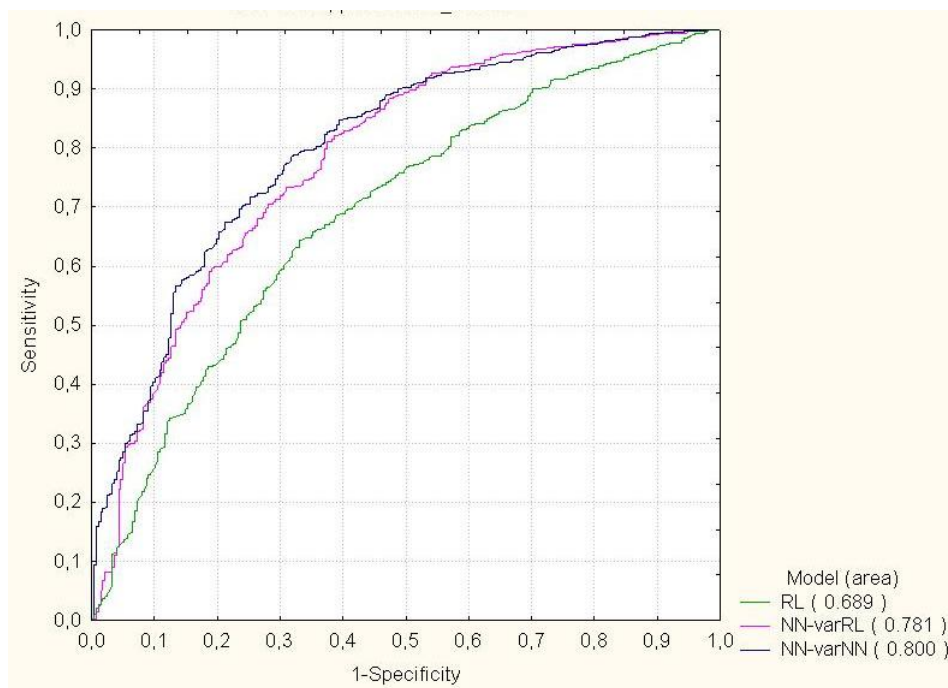
Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin II receptor blocker; EBCM, evidence-based cardiovascular medications.

Figure legend

Fig. 1 ROC curves for EBCM subgroup

	NN – varLR	NN – varNN
Sensitivity	79.2	80.3
Specificity	62.8	63.9

FIGURE 1



From Dr BOURDES Valérie
Société THEMIS
Bioparc - 60 avenue Rockefeller
69008 LYON, FRANCE

To Dr SPAAN Jos
Editor in Chief
Medical & Biological Engineering & Computing

April 21st, 2011

Dear Dr Spaan,

We thank you for accepting our Research Article titled “Prediction of persistence of combined evidence-based cardiovascular medications in patients with acute coronary syndrome after hospital discharge using neural networks” by Valérie Bourdès, Jean Ferrières, Jacques Amar, Elisabeth Amelineau, Stéphane Bonnevey, Maryse Berlion, and Nicolas Danchin.

We carefully read the last comments of the reviewers and we revised the manuscript taking into account all the changes requested by the reviewers.

The replies to the reviewers, comment by comment, are given below and we submit our revised manuscript.

Yours sincerely,

Valérie Bourdès, MD, MPH

REPLY TO REVIEWER:

COMMENTS:

Reviewer #1:

- *Methods / Study Design [Line 14]:*
"...Good Epidemiological Practice..." requires a reference.

Authors: the reference has been added in Section "Methods / Study design", Line 14 and in the reference list (reference 21).

- *Methods / Variable Selection [Line 18]:*
The authors used the error of the training set as the objective criterion of the GA. Why not use the test set? Using the training set gives limited information as to the NN's generalisation capability and so might be misleading. A brief comment here would suffice.

Authors: The aim of this process is not to build the NN, but only to select the best set of variables. At each step of this process (each set of variables and each generation: $100 \times 100 = 10000$ step) a NN is build and an error is calculated. Then, with the best set of variables, a NN is build according to a classical learning process with test and training sets.

- *Methods / Variable Selection [Line 19]:*
The authors use 100 population over 100 generations. Can the authors justify these numbers? Did the population sufficiently converge over 100 generations?

Authors: There's no convergence reason; these choices were only guided by the time-consuming process.

- *Methods / Variable Selection:*
Can the authors give the cross-over and mutation coefficient values?

Authors: The single-point crossover rate was set to 0.8 and the mutation rate was set to 0.1.

Reviewer #2:

From my point of the view the authors have accomplished the reviewers' suggestions and have incorporated all necessary changes.