International Journal of Cardiology xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

Letter to the Editor

A new memetic pattern based algorithm to diagnose/exclude coronary artery disease ☆

Michael J. Zellweger [a,*], Miriam Brinkert [a], Urs Bucher [a], Andrew Tsirkin [b], Peter Ruff [b], Matthias E. Pfisterer [a]

[a] Cardiology Department, University Hospital Basel, Switzerland
[b] Exploris AG, Zürich, Switzerland

ARTICLE INFO

Article history:
Received 14 February 2014
Accepted 29 March 2014
Available online xxxx

Keywords:
Coronary artery disease
Diagnosis
Risk scores
Decision making
Artificial intelligence

Currently used coronary artery disease (CAD) risk scores estimate CAD-related event risk rather than the presence or the absence of CAD. Newer statistical methods such as neural network systems are rarely used in this context. They could help to diagnose/exclude CAD based on easily available patient data thereby reducing the need of unnecessary tests and related costs. We applied such a memetic pattern based algorithm (MPA) to data of 2 separate patient cohorts using simple clinical variables to determine the diagnostic accuracy and compare it to the Framingham risk score.

The MPA was developed by combining binary classification methods particularly ensemble methods with an evolutionary learning optimization engine aiming to create models with low type-II errors [1–10]. For statistical details see http://exploris.info/Cardioexplorer/ where the modeling process is demonstrated in detail illustrated by an open access movie. Patients evaluated for CAD by routine testing were enrolled into the study. Results were compared with invasive coronary angiography: 245 patients were used ("training population") to build the MPA and 128 separate patients for its validation. MPA results were compared with Framingham risk scores (FRSs) using receiver operating curves (ROCs) for overall accuracy (areas-under-the-curve (AUCs)), predictive values and likelihood ratios for individual patient results.

Of the 373 patients, 66% were male, aged 65 ± 10 years and 66% had CAD. Variables most predictive by MPA in the training population were then used to evaluate the validation population. The AUC of MPA and FRS was 0.79 and 0.66, respectively, in the training population and 0.82 and 0.74, respectively, in the validation cohort. The positive predictive values of MPA and FRS to correctly diagnose CAD were 90% and 84%, respectively. The positive predictive values of MPA and FRS for exclusion of CAD were 100% and 50%, respectively. The likelihood ratio was $\geq 1.75\times$ higher by MPA than by FRS for diagnosing CAD and $\geq 5\times$ higher for excluding CAD, with most striking differences at the ends of the ROCs (Fig. 1). Since training and validation cohorts were restricted in sample size, additional sensitivity analyses were performed to estimate how the results would turn out in larger populations and with potential "noise" (Table 1). They demonstrate that the models are stable and appropriate for diagnostic decision making. In the validation population, anticipating the worst case, only 3.8% were false positive when it came to the exclusion of CAD. The quality of binary classification stayed on the same level, despite the slight decline of the overall model quality.

This study presents a new statistical algorithm to diagnose/exclude CAD based on readily available clinical parameters only. The MPA developed in a training population based on a heterogeneous ensemble classifier optimized with an evolutionary learning optimizing engine proved to provide a better discrimination between patients with versus without angiographically documented CAD than the FRS. The validation population confirmed the diagnostic power of the MPA and its stability was verified with Monte Carlo simulation. In a particular patient the MPA provided one of the following answers: the patient "suffers from CAD", "does not suffer from CAD" (with a probability of more than 96%) or "does not sufficiently fit into the algorithm and therefore cannot be categorized into these two categories despite having a certain probability of CAD and therefore needs "conventional" work up". This is an important step towards personalized medicine and decision making.

In daily practice, the gold-standard to diagnose or exclude CAD still is invasive coronary angiography. However, a large number of non-invasive tests including echocardiography, scintigraphy as well as computed tomography and magnetic resonance imaging based techniques are available to evaluate individual patients with respect to the presence or the absence of CAD. This leads to an "overuse" of these tests with inherent risks [11]. Understanding the multifactorial nature of cardiovascular disease and the interrelations among risk factors as reflected by recent CAD scores and new statistical methods as used in the MPA should result in avoiding unnecessary and expensive invasive and

☆ This study was supported by a grant from the Swiss Cardiology Foundation (Grant 2007).
* Corresponding author at: Cardiology Department, University Hospital, Petersgraben 4, CH-4031 Basel, Switzerland. Tel.: +41 61 265 54 73; fax: +41 61 265 45 98.
E-mail address: michael.zellweger@usb.ch (M.J. Zellweger).

http://dx.doi.org/10.1016/j.ijcard.2014.03.184

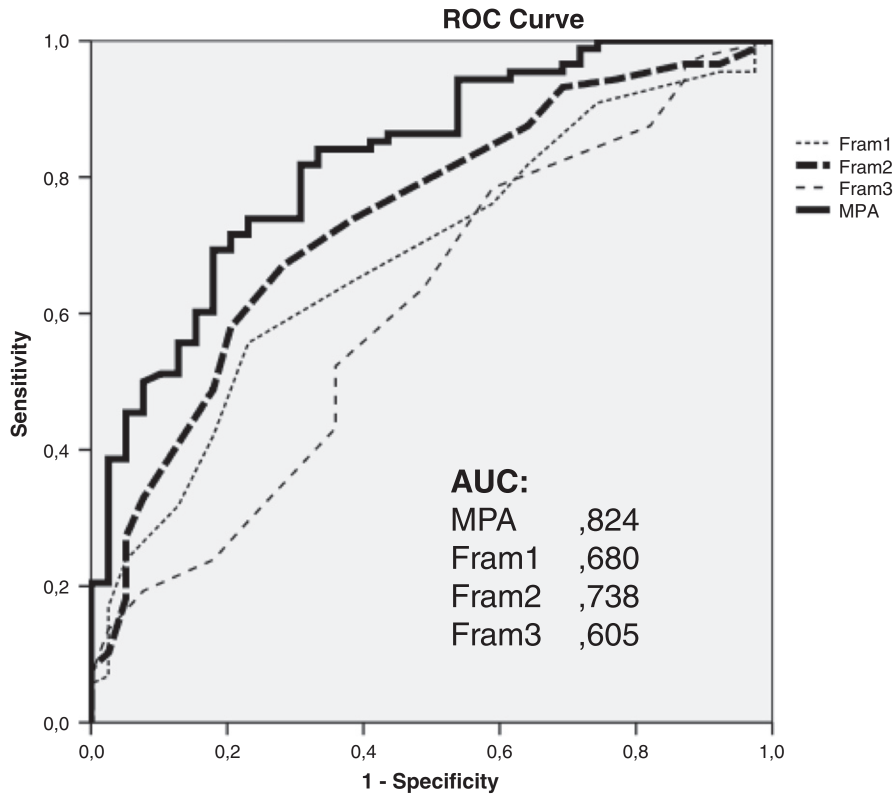

Fig. 1. Comparison of the overall predictive value of the most often cited Framingham 1, 2, and 3 and MPA by receiver operating curves (ROCs) and areas under the curve (AUCs) with respect to the diagnosis of CAD.

non-invasive procedures particularly in patients at low disease probability. This would be an important achievement.

Comparing results of MPA in the present study to performances of logistic regression, classification and regression tree, and neural networks for predicting CAD reported by Kurt et al. showed that AUC values of MPA described here are just above the highest of all statistical modeling methods considered in that article [12]. The main advantage of the new statistical models described by Kurt and tested in the present study is that these methods seem better suited for diagnostic classifications with lower levels of a beta error in patients with a high or a low probability of CAD (the ends of the ROC).

The study population of the current study was restricted to patients referred for angiographic evaluation of suspected CAD, i.e. with a relatively high pre-test probability of CAD, however without prior evidence of the disease. The strength is that the MPA was developed and validated in two separate independent patient populations and that risk score evaluations were confirmed by invasive coronary angiography in all patients. In addition, sensitivity analyses showed that the stability of MPA is given, particularly also for patients in whom CAD could be excluded indicating that similarly reliable results may be anticipated also for larger populations of the same risk category.

Since MPA provided a higher diagnostic accuracy for CAD than the FRS, particularly for identifying patients with very high/low CAD probabilities, these findings form a first step to introduce MPA into the clinical arena aiming to prevent unnecessary diagnostic procedures. This novel approach needs to be corroborated in a large prospective study with patients at low and high pre-test probabilities of CAD.

Disclosures

Peter Ruff is part-owner, board member and head of Exploris AG which is a privately owned Swiss research company focusing in development of novel diagnostic solutions.

Andrew Tsirkin is head modeling and development of Exploris AG.

Table 1
Diagnostic results in the validation population and results of the simulations with and without noise.

	Validation population		Simulation results (no noise) 90%-CI	Simulation results (5% noise used) 90-CI	Binominal 90%-CI
	Actual value	Median			
CAD+ true positive ratio	52.3%	53.4%	42.2–60.4%	31.2–61.0%	49.5–62.4%
CAD+ false positive ratio	12.5%	9.5%	4.5–15.8%	4.5–23.3%	5.2–14.4%
LHR CAD+ (TP/FP)	4.1%	6.2%	3.5–10.6%	2.4–8.7%	3.5–10.9%
H+ true positive ratio	12.5%	11.5%	6.7–16.7%	6.5–17.4%	6.1–16.8%
H+ false positive ratio	0%	1.0%	0–3.8%	0–3.8%	0.1–2.8%
LHR H+ (TP/FP)	>11[a]	12.9%	3.6–26.1%	3.8–26.3%	2.9–89.7%

90%-CI: 90% confidence interval.
LHR: likelihood ratio.
TP: true positive.
FP: false positive.
CAD+: presence of CAD.
H+: absence of CAD.
[a] Since H+ specificity is 100% and the model has no type II errors, technically LHR tends to infinity. Therefore one type II error was assumed for calculation purposes.

Please cite this article as: Zellweger MJ, et al, A new memetic pattern based algorithm to diagnose/exclude coronary artery disease, Int J Cardiol (2014), http://dx.doi.org/10.1016/j.ijcard.2014.03.184

References

[1] Carpenter GA, Grossberg S, Markuzon N, Reynolds JH, Fuzzy Rosen DB. ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE Trans Neural Netw 1992;3:698–713.
[2] Catto JW, Linkens DA, Abbod MF, et al. Artificial intelligence in predicting bladder cancer outcome: a comparison of neuro-fuzzy modeling and artificial neural networks. Clin Cancer Res 2003;9:4172–7.
[3] Knowles JD, Corne DW. Approximating the nondominated front using the Pareto Archived Evolution Strategy. Evol Comput 2000;8:149–72.
[4] Hammer B, Micheli A, Sperduti A, Strickert M. Recursive self-organizing network models. Neural Netw 2004;17:1061–85.
[5] Johnson DS, Hoeting JA. Autoregressive models for capture-recapture data: a Bayesian approach. Biometrics 2003;59:341–50.
[6] Kohonen T. Self-organizing neural projections. Neural Netw 2006;19:723–33.
[7] Kuncheva LI. Evaluation of computerized medical diagnostic decisions via fuzzy sets. Int J Biomed Comput 1991;28:91–100.
[8] Moon H, Ahn H, Kodell RL, Baek S, Lin CJ, Chen JJ. Ensemble methods for classification of patients for personalized medicine with high-dimensional data. Artif Intell Med 2007;41:197–207.
[9] Ontrup J, Ritter H. Large-scale data exploration with the hierarchically growing hyperbolic SOM. Neural Netw 2006;19:751–61.
[10] Hart WE. Locally-adaptive and memetic evolutionary pattern search algorithms. Evol Comput 2003;11:29–51.
[11] Shaw LJ, Marwick TH, Zoghbi WA, et al. Why all the focus on cardiac imaging? JACC Cardiovasc Imaging 2010;3:789–94.
[12] Kurt I, Ture M, Kurum AT. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Syst Appl 2008;34:366–74.

Please cite this article as: Zellweger MJ, et al, A new memetic pattern based algorithm to diagnose/exclude coronary artery disease, Int J Cardiol (2014), http://dx.doi.org/10.1016/j.ijcard.2014.03.184