Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: http://www.elsevier.com/locate/trprot

PanelomiX: A threshold-based algorithm to create panels of biomarkers

Xavier Robin[a], *Natacha Turck*[a], *Alexandre Hainard*[a], *Natalia Tiberti*[a], *Frédérique Lisacek*[b], *Jean-Charles Sanchez*[a,*], *Markus Müller*[b]

[a] *Translational Biomarker Group, Department of Human Protein Sciences, University Medical Centre, Geneva, Switzerland*
[b] *Swiss Institute of Bioinformatics, University Medical Centre, Geneva, Switzerland*

ARTICLE INFO

Article history:
Received 13 February 2013
Received in revised form
24 April 2013
Accepted 26 April 2013

Keywords:
Biomarkers
Panel
Combination of biomarkers
Machine learning
Clinical study
Aneurysmal subarachnoid haemorrhage

ABSTRACT

In order to increase their predictive power, medical biomarkers can be combined into panels. However, the lack of ready-to-use tools generating interpretable results and implementing rigorous validation standards hampers the more widespread application of panels and their translation into clinical practice.

The computational toolbox we present here – PanelomiX – uses the iterative combination of biomarkers and thresholds (ICBT) method. This method combines biomarkers and clinical scores by selecting thresholds that provide optimal classification performance. To speed up the calculation for a large number of biomarkers, PanelomiX selects a subset of thresholds and parameters based on the random forest method. The panels' robustness and performance are analysed by cross-validation (CV) and receiver operating characteristic (ROC) analysis.

Using 8 biomarkers, we compared this method against classic combination procedures in the determination of outcome for 113 patients with an aneurysmal subarachnoid haemorrhage. The panel classified the patients better than the best single biomarker ($p < 0.005$) and compared favourably with other off-the-shelf classification methods.

In conclusion, the PanelomiX toolbox combines biomarkers and evaluates the performance of panels to classify patients better than single markers or other classifiers. The ICBT algorithm proved to be an efficient classifier, the results of which can easily be interpreted.

1. Introduction

The translation of panels of biomarkers into clinical practice is principally obstructed by two critical factors [1]. Firstly, methods and results can often be difficult to understand for non-experts; secondly, there is a general lack of robust validation steps, which are critical for the reproducibility of results given high biological variation.

Abbreviations: ROC, receiver operating characteristic; AUC, area under the ROC curve; pAUC, partial AUC; CV, cross-validation; SE, sensitivity; SP, specificity; aSAH, aneurysmal subarachnoid haemorrhage; SVM, support vector machines.

* *Corresponding author*. Tel.: +41 223795486.
E-mail addresses: Xavier.Robin@unige.ch (X. Robin), Natacha.Turck@unige.ch (N. Turck), Alexandre.Hainard@unige.ch (A. Hainard), Natalia.Tiberti@unige.ch (N. Tiberti), Frederique.Lisacek@isb-sib.ch (F. Lisacek), Jean-Charles.Sanchez@unige.ch (J.-C. Sanchez), Markus.Mueller@isb-sib.ch (M. Müller).

2212-9634
http://dx.doi.org/10.1016/j.trprot.2013.04.003

To overcome the first issue, a combination method must produce clear and easily interpretable results, where patient classification can be understood in terms of the contribution of each individual biomarker. Medical practitioners have long been used to clinical scores, such as the Hoffer–Osmond test to diagnose schizophrenia [2,3], or the Ranson score [4] for the prognosis and operative management of acute pancreatitis. These methods were recently applied to assess the probability of pulmonary embolism [5] and acute pancreatitis [6]. These types of scores have become popular because they are clear and easy to interpret, granting access to the intermediate results of individual sub-tests. This is in contrast to black box classifiers, such as neural networks or support vector machines (SVM), which may display high accuracy, but which do not reveal the contribution of each individual marker directly. While black boxes are acceptable in specific applications, they may not always be suitable in expert systems for medical decision-making [7–9]. In contrast, many methods present results in a user-friendly format referred to as "white boxes".

Combining biomarkers is an application of statistical learning. Over the years, this field has developed countless methods to tackle the task. Linear or logistic regression methods determine a factor, generally multiplicative, for each biomarker included in the panel. A straightforward interpretation of these factors is to see them as the "weights" of influence of the biomarkers. Methods based on decision trees also provide an easy interpretation, where one follows a sequence of binary splits. As long as a tree contains only a fairly limited number of such decisions (or branches), these are easy to track and to justify how a decision was reached. Decision trees are graphically expressive (see [1]) for easier understanding. Finally, in threshold-based methods, all biomarker tests are analysed at the same time (instead of sequentially), and the number of positive tests defines a score used for classification.

The second issue is the lack of a robust validation step. Panel validation requires an independent test set – preferably measured in a different laboratory – in order to compute the panel's true performance and avoid performance overestimation due to over-fitting the data during the learning process [1]. If no independent set is available, computational methods such as cross-validation (CV) or bootstrapping allow the simulation of such sets [10,11].

Two useful and quite common performance measures are sensitivity (the proportion of positive patients correctly detected by the test) and specificity (the proportion of negative patients correctly rejected by the test), as they give clear estimates of how patients are classified [1]. When no biomarker level cut-off is preferred or pre-defined, receiver operating characteristic (ROC) analysis can be performed to weight the trade-off between sensitivity and specificity [10]. The area under the ROC curve (AUC) is also a very common performance metric in medical decision-making [12], bioinformatics [13] and statistical learning [14]. An important and often neglected step is the panel's performance comparison against that of single biomarkers. A fair evaluation would process the panel and single biomarkers with the same tools (sensitivity and specificity or AUC) on the same independent test set or with the same CV procedure [1]. Then performance could be compared either with McNemar's test (for sensitivity or specificity) or using ROC curves.

The methods we propose here, which use single biomarker thresholds as the base of their decisions, are part of the PanelomiX software. In threshold-based combinations, thresholds are often chosen in a univariate manner. For example, Ranson et al. [4] selected convenient prognostic sign cut-off values outside the range of the mean plus or minus one standard deviation; Morrow and Braunwald [15] chose the 99th percentile of the control distribution; Sabatine et al. [16] used the cut-offs described in the literature. In contrast, Reynolds et al. [17] adopted a multivariate approach and tested many thresholds by 10% increments. This approach takes into account the interaction that may arise when biomarkers are combined. PanelomiX can combine biomarkers (molecule levels, clinical scores, etc.) in a multivariate manner. Therefore we developed an exhaustive search algorithm to select the optimal thresholds, and called it iterative combination of biomarkers and thresholds (ICBT). To minimize execution times, we developed several approaches to reduce complexity and hence increase search speed. As it has been shown to be an efficient feature selection method [11], we used random forest [18,19] as a filtering method to reduce both the number of biomarkers and thresholds that account for the search space size. Random forest builds a large number of decision trees that are made slightly different by bootstrapping. In the end, the classification is the average prediction of all trees.

PanelomiX has already been applied to predict the outcome of an aneurysmal subarachnoid haemorrhage (aSAH) [20] and to assess the progression of human African trypanosomiasis [21]. Below, we demonstrate the PanelomiX methodology and performance, using 8 parameters for the determination of outcome for patients with an aSAH.

2. Methods

2.1. *Iterative combination of biomarkers and thresholds (ICBT)*

2.1.1. *Combining biomarkers*

The approach adopted here is based on the ICBT method. A threshold is defined for each biomarker by an optimization procedure defined in the following sections. A patient's score is the number of biomarkers exceeding their threshold values.

We can write this as:

$$S_p = \sum_{i=1}^{n} I(X_{ip} \geq T_i) \quad (1)$$

where S_p is the score for patient p, n is the number of biomarkers, X_{ip} is the concentration of the ith biomarker in patient p, T_i is the threshold for the ith biomarker, and $I(x)$ is an indicator function which takes the value of 1 for x = true and 0 otherwise.

If biomarker concentrations are higher in the control than in the disease group, then they are multiplied by -1 before applying the previous formula.

To classify a patient, a threshold on the S_p score is required and defined as T_s. Patients with a score $S_p \geq T_s$ are positive; negative otherwise.