Estimation of optimally combined-biomarker accuracy in the absence of a gold standard reference test

Leandro Garcia Barrado[1], Elisabeth Coart[2], Tomasz Burzykowski[1,2]

Bayesian Young Statisticians Meeting (BAYSM), Milan June, 5-6, 2013
Paper no. 2

[1] Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat)
Hasselt University, Agoralaan building D, 3590 Diepenbeek, Belgium.
leandro.garciabarrado@uhasselt.be
tomasz.burzykowski@uhasselt.be
[2] International Drug Development Institute (IDDI), Avenue Provinciale 30, 1340 Louvain-la-Neuve, Belgium.
elisabeth.coart@iddi.com

Abstract

The reference diagnostic test used to establish the discriminative properties of a combination of biomarkers could be imperfect. This may lead to a biased estimate of the accuracy of the combination. A Bayesian latent-class mixture model is proposed to estimate the Area Under the ROC Curve (AUC) of a combination of biomarkers. The model allows selecting the combination that maximizes the AUC. Due to the latent structure of the model, possible errors in the reference test are taken into account. A simulation study was performed based on 400 data sets. Sample sizes from 100 to 600 observations were considered. The prior information about the diagnostic accuracy of the reference test was varied from non-informative to informative. The obtained average estimates for all parameters were close to the true values; some bias and differences in efficiency were observed. Results indicate an adequate performance of the model-based estimates. The observed bias could be overcome by increasing the amount of information in the prior for the diagnostic-accuracy parameters of the reference test.

Keywords: Bayesian estimation; latent class mixture models; AUC

1 Introduction

Biomarkers can be used for developing a diagnostic test for a disease. Often, to increase the diagnostic accuracy of the test, a combination of several biomarkers

is considered [1]. To assess the diagnostic performance of a biomarker-based test, a so-called "reference test", establishing disease status of an individual, is needed. Depending on the disease of interest, the reference test may be imperfect, i.e., it may misclassify the control and diseased individuals. In such a case, the estimate of the diagnostic accuracy of a biomarker could be biased [2]. Therefore, when developing a biomarker-based diagnostic test, the possibility of an imperfect reference test has to be taken into account.

2 Methods

We use the Area Under the ROC Curve (AUC) as a measure of the diagnostic accuracy and a model to derive the linear combination of biomarkers maximizing the AUC [3]. In particular, a Bayesian latent-class mixture model is fitted to obtain estimates of the distributional parameters of the multivariate distributions for the biomarkers that form the mixture components for the diseased and control populations. By estimating the latent true disease status and component parameters through a mixture model with both reference test and biomarker values contributing to the likelihood, the misclassification probabilities of the reference test are taken into account [4].

A simulation study was performed to investigate the performance of the model under several settings for 400 simulated data sets. Sample sizes of 100, 400, and 600 observations were considered, split equally between the diseased and control groups. The prior distributions for the sensitivity and specificity of the reference test were varied from non-informative to informative.

3 Results

The results point to bias and decreasing efficiency of posterior estimates when sample size decreases. Increasing prior information resolves, or at least reduces, the bias observed for small data sets (see Table 1). Counterintuitively, the increase in prior information leads to a decrease in efficiency of the AUC estimates. It appears that the consequence of assuming non-informative priors for the parameters of the biomarker related distributions is that the prior for AUC essentially becomes a point mass distribution at one. This explains the result as an increased effort to overcome the highly informative AUC prior information in favour of the data.

Table 1: Mean (standard error) of the median posterior AUC of all 400 fits for all considered settings.

		Sample size		
Se/Sp Prior info	**True AUC**	**N=100**	**N=400**	**N=600**
Non-informative	0.8786	0.9241 (0.0279)	0.8890 (0.0279)	0.8836 (0.0262)
Informative	0.8786	0.9068 (0.0344)	0.8827 (0.0286)	0.8785 (0.0263)

4 Conclusions

The results indicate that the model does provide unbiased estimates of the accuracy of the optimal combination of diagnostic biomarkers, but an informative prior about the imperfect reference test may be needed to overcome the lack of sufficient information for the reference test accuracy in small data sets. For larger sample sizes, the efficiency of the model-based estimates increases. The proper specification of the AUC prior requires further investigation.

References

[1] X.H. Zhou, N.A. Obuchowski, D.K. McClish. *Statistical Methods in Diagnostic Medicine*. New York: Wiley inc.; 2002.

[2] P.N. Valenstein. Evaluating diagnostic tests with imperfect standards. *American Journal of Clinical Pathology*; 1990; 93(2); pp. 252-258.

[3] J.Q. Su, J.S. Liu. Linear combinations of multiple diagnostic markers. *Journal of the American Statistical Association*; 1993; 88(424); pp. 1350-1355.

[4] A.N. Scott, L. Joseph, L. Bélisle, M.A. Behr, K. Schwartzman. Bayesian modelling of tuberculosis clustering from DNA fingerprint data. *Statistical Medicine*; 2007; 27(1); pp. 140-156.