

# Chapter 2

## Estimation of Optimally Combined-Biomarker Accuracy in the Absence of a Gold Standard Reference Test

Leandro Garcia Barrado, Elisabeth Coart, and Tomasz Burzykowski

**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 and takes possible errors in the reference test into account. A simulation study was performed based on 400 data sets. Sample sizes from 100 to 600 observations were considered. Informative as well as non-informative prior information for the diagnostic accuracy of the reference test was considered. In addition, a controlled prior specification is proposed. The obtained average estimates for all parameters were close to the true values; some differences in efficiency were observed. Results indicate an adequate performance of the model-based estimates.

### 2.1 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 [6]. To assess the diagnostic performance of a biomarker-based test,

---

L.G. Barrado (✉)

Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat),  
Hasselt University, Agoralaan building D, 3590 Diepenbeek, Belgium  
e-mail: [leandro.garciabarrado@uhasselt.be](mailto:leandro.garciabarrado@uhasselt.be)

E. Coart

International Drug Development Institute (IDDI), Avenue Provinciale 30,  
1340 Louvain-la-Neuve, Belgium  
e-mail: [elisabeth.coart@iddi.com](mailto:elisabeth.coart@iddi.com)

T. Burzykowski

I-BioStat, IDDI, Louvain-la-Neuve, Belgium  
e-mail: [tomasz.burzykowski@uhasselt.be](mailto:tomasz.burzykowski@uhasselt.be)

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 [4]. Therefore, when developing a biomarker-based diagnostic test, the possibility of an imperfect reference test has to be taken into account.

## 2.2 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 [2].

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.

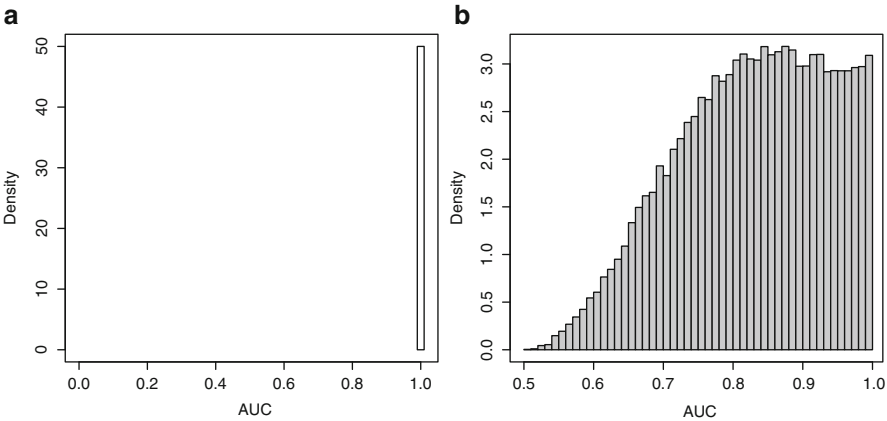
Priors for the remaining parameters were set as standard non-informative priors [1]. This “naive” approach does not enable control of the prior distributions for the variances, correlations, and the AUC. Therefore, an alternative prior specification is proposed. Re-parametrization of the variance-covariance matrices allows a more direct specification of the prior information for variances and correlations [5]. By putting a prior distribution on the difference of the mixture component means, scaled by the sum of the variance-covariance matrices, prior information for the AUC can be precisely specified.

## 2.3 Results

Table 2.1 presents the results of the simulation study for the three considered sample sizes. The table contains the average of the median posterior AUC estimates over the 400 data sets for each of the different simulation settings. The rows marked Naive correspond to the “naive” approach, which leads to the AUC prior as in Fig. 2.1a. The rows marked Controlled correspond to the re-parametrized approach,

**Table 2.1** Mean (standard error) of the median posterior AUC of all 400 fits for all considered settings

Prior formulation	Se/Sp prior	True AUC	Sample size		
			N = 100	N = 400	N = 600
Naive	Non-Inf	0.8786	0.9241 (0.0279)	0.8890 (0.0279)	0.8836 (0.0262)
Naive	Inf	0.8786	0.9068 (0.0344)	0.8827 (0.0286)	0.8785 (0.0263)
Controlled	Non-Inf	0.8786	0.8907 (0.0347)	0.8803 (0.0290)	0.8773 (0.0271)
Controlled	Inf	0.8786	0.8728 (0.0388)	0.8741 (0.0292)	0.8722 (0.0269)



**Fig. 2.1** Simulated implied priors for AUC based on mixture component priors. (a) Implied prior for the naive prior specification. (b) Implied prior for the proposed controlled prior specification

with the AUC prior shown in Fig. 2.1b. Within each parametrization approach, the rows indicated by Non-Inf and Inf represent the results for the non-informative and informative prior for the accuracy of the reference test, respectively.

Considering the naive approach, the results point to overestimation and decreasing efficiency of posterior estimates with decreasing sample size. Increasing the amount of prior information for the accuracy of the reference test resolves, or at least reduces, the bias observed for small data sets. Counterintuitively, the increase of the 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 the AUC essentially becomes a point mass distribution at one (see panel A of Fig. 2.1). This explains the overestimation of the AUC as due to the highly informative AUC prior distribution. Changing the parametrization of the model allows specifying the prior distribution for the AUC as in Fig. 2.1b. As a consequence, introducing this less informative prior reduces the bias. Increasing sample size or the amount of information in the prior for the accuracy of the reference test does not alter the results substantially, as shown in Table 2.1.

## 2.4 Conclusions

The results indicate that the model does provide unbiased estimates of the accuracy of the optimal combination of diagnostic biomarkers, but care has to be taken in the specification of the prior information, especially for the AUC. Under the “naive” approach, an informative prior for the accuracy of the imperfect reference test may overcome the informative prior for the AUC in small data sets.

## References

1. O'Malley AJ, Zou KH (2006) Bayesian multivariate hierarchical transformation models for ROC analysis. *Stat Med* 25:459–479
2. Scott AN, Joseph L, Bélisle MA, Behr KS (2007) Bayesian modelling of tuberculosis clustering from DNA fingerprint data. *Stat Med* 27:140–156
3. Su JQ, Liu JS (1993) Linear combinations of multiple diagnostic markers. *J Am Stat Assoc* 88:1350–1355
4. Valenstein PN (1990) Evaluating diagnostic tests with imperfect standards. *Am J Clin Pathology* 93:252–258
5. Wei Y, Higgins PT (2013) Bayesian multivariate meta-analysis with multiple outcomes. *Stat Med*. doi:10.1002/sim.5745
6. Zhou XH, Obuchowski NA, McClish DK (2002) *Statistical Methods in Diagnostic Medicine*. Wiley, New York

The Contribution of Young Researchers to Bayesian  
Statistics

Proceedings of BAYSM2013

Lanzarone, E.; Ieva, F. (Eds.)

2014, X, 214 p. 31 illus., 21 illus. in color., Hardcover

ISBN: 978-3-319-02083-9