Bayesian Meta-analysis of Diagnostic Tests Allowing for Imperfect Reference Standards

Joris Menten*[12], Marleen Boelaert[1], Emmanuel Lesaffre[23]

[1] Institute of Tropical Medicine; [2] L-Biostat KULeuven; [3] Dep. of Biostatistics, Erasmus University Rotterdam

1 Introduction

This work is done in the context of a systematic review and meta-analysis of the accuracy of rapid diagnostic tests (index test) for diagnosing visceral leishmaniasis (VL) disease in first-line health services in VL-endemic areas. The diagnostic accuracy of a test is usually summarized using the proportion of diseased subjects who test positive (Sensitivity S) and the proportion of non-diseased subjects who test negative (Specificity C). Diagnostic studies of VL are however hampered by the lack of a perfect reference test which can classify all subjects as diseased or not. This leads to widely varying approaches with regard to the reference standard: (1) in some studies Latent Class Analysis (LCA) is used to calculate S and C of the index test; (2) in other studies a less than perfect reference standard may have been used. The statistical model used in the meta-analysis must consequently combine results from studies analyzed with a reference standard with those from LCA while allowing for imperfect reference tests

2 Model and Model Estimation

We extend the hierachical bivariate logistic normal model of Reitsma (2005) as follows:

(1) At the higher level of the model, the study-specific diagnostic accuracy measures $\text{logit}(S_i) = \theta_{Si}$ and $\text{logit}(C_i) = \theta_{Ci}$ are modelled using a bivariate normal distribution:

$$\begin{pmatrix} \theta_{Si} \\ \theta_{Ci} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_S \\ \mu_C \end{pmatrix}, \Sigma \right) \text{ with } \Sigma = \begin{pmatrix} \sigma_S^2 & \sigma_{SC} \\ \sigma_{SC} & \sigma_C^2 \end{pmatrix}$$

(2) At the lower level, the individual study data is modeled differently for primary studies that use LCA and those that use a, possibly imperfect, reference standard.

Primary studies that use LCA to estimate the diagnostic accuracy will typically report $\hat{S}_i$ and $\hat{C}_i$ together with 95% credible intervals. From these, we obtain $\text{logit}(\hat{S}_i) = \hat{\theta}_{Si}$, $\text{logit}(\hat{C}_i) = \hat{\theta}_{Ci}$, $\sigma_{\hat{\theta}_{Si}}$, and $\sigma_{\hat{\theta}_{Ci}}$. We assume the observed $\hat{\theta}_{Si}$ and $\hat{\theta}_{Si}$ are drawn from two independent normal distributions:

$$\hat{\theta}_{Si} \sim N(\theta_{Si}, \sigma^2_{\hat{S}i})$$
$$\hat{\theta}_{Ci} \sim N(\theta_{Ci}, \sigma^2_{\hat{C}i})$$

using $\sigma^2_{\hat{S}i}$ and $\sigma^2_{\hat{S}i}$ as plug-in estimators of σ^2_{Si} and σ^2_{Ci}.

For studies using a reference standard, the 2x2 table of index vs. reference test results is modelled using a multinomial distribution for the cell counts. Specifically, if y_{ijk} is the number of subjects in study i with result j (0=negative, 1=positive) for the index test and result k for the reference test, then: $y_{ijk} \sim \text{Mu}(n_i, p_{ijk})$ with

$$p_{ijk} = \pi_i[S_i^j\,(1-S_i)^{1-j}\,S_{Ri}^k\,(1-S_{Ri})^{1-k} + (-1)^{j-k}\,cov_{i|D=1}] +$$
$$(1-\pi_i)\,[C_i^{1-j}\,(1-C_i)^j\,C_{Ri}^{1-k}\,(1-S_{Ri})^k + (-1)^{j-k}\,cov_{i|D=0}]$$

with n_i the sample size and π_i the prevalence in study i, S_{Ri} and C_{Ri} the S and C of the reference test, and $cov_{i|D=1}$ and $cov_{i|D=0}$ the correlations between index and reference test results in diseased and non-diseased subjects, respectively.

For model identifiability, we need deterministic or probabilistic constraints on this model. In our meta-analysis, we assume $cov_{i|D=1} \equiv cov_{i|D=0} \equiv 0$. In addition, we define L types of reference standard and classify for each study i the reference standard R_i in one of these categories and obtain expert opinion concerning the diagnostic accuracy of the L different reference test categories.

As for the index test, we fit a bivariate normal model to the logits of S and C of the reference test:

$$\begin{pmatrix}\theta_{SRi}\\ \theta_{CRi}\end{pmatrix} \sim N\left(\begin{pmatrix}\mu_{S_{l(i)}}\\ \mu_{C_{l(i)}}\end{pmatrix}, \Sigma_{l(i)}\right) \text{ with } \Sigma_l = \begin{pmatrix}\sigma^2_{S_{Rl(i)}} & \sigma_{S_{Rl(i)}C_{Rl(i)}}\\ \sigma_{S_{Rl(i)}C_{Rl(i)}} & \sigma^2_{C_{Rl(i)}}\end{pmatrix}$$

where $\theta_{SRi} = \text{logit}(S_{Ri})$, $\theta_{CRi} = \text{logit}(C_{Ri})$ and $l(i)$ indicates the type of reference test used in study i.

We estimate the model using Markov Chain Monte Carlo methods using WinBUGS, called from within R. We use non-informative priors for the disease prevalences and measures of diagnostic accuracy of the index test and use expert opinion to construct informative priors for $(\mu_{S_{l(i)}}, \mu_{C_{l(i)}})$ and $\Sigma_{l(i)}$.

Region	Model 1 Estimate (95% CI)	Model 2 Estimate (95% CI)
Eastern Africa		
S	82.3 (71.3, 89.5)	86.6 (70.9, 93.6)
C	89.1 (77.3, 94.8)	95.8 (84.9, 98.7)
Indian Subcontinent		
S	96.2 (91.8, 98.3)	95.7 (86.3, 98.4)
C	89.6 (76.0, 95.6)	91.6 (63.9, 97.6)
Latin-America and the Mediterranean		
S	93.4 (84.0, 97.3)	90.0 (67.7, 96.9)
C	96.6 (88.9, 99.1)	98.8 (84.9, 99.9)

Table 1: Combined estimates of the sensitivity S and specificity C of a visceral leishmaniasis rapid diagnostic test by geographic region, as calculated using Bayesian meta-analysis assuming perfect reference standards (Model 1) and allowing for imperfect reference standards (Model 2)

3 Preliminary Results

To assess the feasibility of this modelling approach, we fitted models assuming perfect reference standard and allowing for imperfect reference standard to the data. In this analysis, we used preliminary informative priors for the diagnostic accuracy of the reference tests obtained from the literature. The diagnostic accuracy of the index test varied by geographic region. Table 1 shows the average (95% CI) S and C of the index test assuming perfect reference standards (Model 1) and adjusting for imperfect reference standards (Model 2). The analysis indicates that the assumption of a perfect reference standard may have resulted in a downward bias of C estimates in the primary publications.

References

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH (2005). Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. Journal of Clinical Epidemiology 2005; 58(10):98290.