Evidence Synthesis for Diagnostic Tests with Partially Ordered Performance and No Reference Standard

Shaun P. Forbes, BSBA[1], Mengyang Di, MD, PhD[1], Joshua A. Salomon, PhD[2], Benjamin P. Linas, MD, MPH[3] and Thomas A. Trikalinos, MD[1]
(1)Brown University, Providence, RI, (2)Harvard School of Public Health, Boston, MA, (3)Boston Medical Center, Boston, MA

www.cebm.brown.edu

Purpose

1. In many health problems, the true disease status is unknown; thus the sensitivity and specificity of tests cannot be identified (estimated) directly from individual studies.
2. Yet, mathematical models of such problems often require sensitivity and specificity estimates.
3. We develop evidence synthesis models for estimating sensitivity and specificity of tests when there is no gold standard, by borrowing information across multiple studies and leveraging expert input on how tests' performance is ordered.
4. To explicate, we estimate the sensitivity and specificity of tuberculin skin test (TST) and interferon-gamma release assays (IGRA) to diagnose Latent Tuberculosis Infection (LTBI). Experts suggest that the specificity of IGRA is better than that of TST.
5. The example was used in a mathematical model for policy analysis.

Methods

Model

We develop a Bayesian hierarchical random effects model that
1. models the unobserved (latent) disease status,
2. respects the relationship between sensitivity and specificity across studies,
3. models between-study heterogeneity,
4. respects stratification by study,
5. allows for dependencies in tests' performance conditional on disease status,
6. models the effect of covariates, and
7. can use external information on the ordering of test performance to improve identifiability.

The model consists of observational and structural parts.

Observational part (connects data to study level parameters)

Counts of cross-classified test results in study *k*

Test results	D(+)	D(-)	Marginal
T1(+), T2(+)	?	?	$r_{k11} = 133$
T1(+), T2(-)	?	?	$r_{k10} = 44$
T1(-), T2(+)	?	?	$r_{k01} = 26$
T1(-), T2(-)	?	?	$r_{k00} = 62$

$\mathbf{r}_k \sim \text{Multinomial}(\boldsymbol{\lambda}_k, N_k)$

Probabilities of cross-classified test results in study *k*

Test results	D(+)	D(-)	Marginal
T1(+), T2(+)	$s_{k1}s_{k2} + s_k^*$	$(1-c_{k1})(1-c_{k2}) + c_k^*$	λ_{k11}
T1(+), T2(-)	$s_{k1}(1-s_{k2}) - s_k^*$	$(1-c_{k1})c_{k2} - c_k^*$	λ_{k10}
T1(-), T2(+)	$(1-s_{k1})s_{k2} - s_k^*$	$c_{k1}(1-c_{k2}) - c_k^*$	λ_{k01}
T1(-), T2(-)	$(1-s_{k1})(1-s_{k2}) + s_k^*$	$c_{k1}c_{k2} + c_k^*$	λ_{k00}

Methods (Continued)

Structural part (models relationship of parameters across studies)

Prevalence

$$\text{logit}(\pi_k) \sim N(\mu_\pi, \tau_\pi^2) \quad [1]$$

Sensitivities and specificities

$$\begin{bmatrix}\text{logit}(s_{k1})\\ \text{logit}(s_{k2})\\ \text{logit}(c_{k1})\\ \text{logit}(c_{k2})\end{bmatrix} = \begin{bmatrix}\eta_{k1}\\ \eta_{k2}\\ \xi_{k1}\\ \xi_{k2}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\eta}_k\\ \boldsymbol{\xi}_k\end{bmatrix} \sim N\left(\begin{bmatrix}\mathbf{H}\\ \boldsymbol{\Xi}\end{bmatrix}, \mathbf{T}\right) \quad [2]$$

Subject to

$$\begin{aligned}\xi_{k2} - \xi_{k1} &\geq 0 \text{ for all } k = 1, \dots K,\\ \Xi_2 - \Xi_1 &\geq 0.\end{aligned} \quad [3]$$

Obtained from experts

Covariates

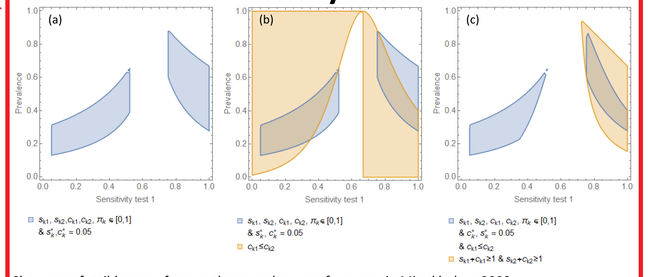
To incorporate between-study effects of covariates ($\mathbf{x}$), such as
- originating from TB endemic area (x_1),
- history of active TB case contact (x_2),
- [centered] mean age (x_3),
- chronic disease (x_4),
- having HIV (x_5),
- history of BCG vaccination (x_6), and
- having other risks (x_7),

set $\mu_\pi := \mu_\pi(\mathbf{x})$ in [1] and $\mathbf{H} := \mathbf{H}(\mathbf{x})$, $\boldsymbol{\Xi} := \boldsymbol{\Xi}(\mathbf{x})$ in [2].

We fit the model with a Bayesian Algorithm

- MCMC with 3 chains
- Burn-in of 50k iterations
- Update until convergence (Gelman-Rubin diagnostic <1.1 for means and variances in [1] and [2])

Using partially-ordered performance we can better identify latent variables



Shown are feasible areas for prevalence and test performance in Mirtskhulava 2008.
(a) feasible areas with no external information; (b) using constraint [3]; (c) if the tests are non-adversarial

Selected Results

To inform the mathematical model we can obtain posterior distributions of LTBI prevalence and test performance for 12 cost effectiveness analysis scenarios. We show 2 of them:

Scenario 1
- **Not from endemic TB area**
- No history of TB case contact
- Age 40
- **No chronic disease**
- No HIV
- **No BCG vaccination**
- Has other risks

Scenario 10
- **Born in endemic TB area**
- No history of TB case contact
- Age 40
- **Has chronic disease**
- No HIV
- **Has BCG vaccination**
- Has other risks

Conclusions

1. We developed a meta-analysis model that estimates test performance measures in the absence of an error-free reference standard and that can use information on partially-ordered performance.
2. We address a very common need in cost-effectiveness analyses of test-and-treat strategies.

Limitations

1. Because we are dealing with a fundamentally under-identified problem, results can be sensitive to specification of priors.
2. In some cases a simpler solution/approximation could be adequate, especially if test performance is not an important parameter in the mathematical model.

Future Directions

1. Extensions
   - Model within-study (person-level) effects of covariates
2. Explorations
   - Sensitivity of results to choice of priors & model structure
   - Exploration of (non)identifiability with prior-posterior predictive checks

Acknowledgements

This project was funded by the U.S. Centers for Disease Control and Prevention, National Center for HIV, Viral Hepatitis, STD, and TB Prevention Epidemiologic and Economic Modeling Agreement 5U38PS004642
SPF is supported by the Brown University Presidential Fellowship