The Importance of Prevalence on Test Thresholds & Outcomes: A Tuberculosis Testing Model

Tanya G.K. Bentley, PhD[1], Antonio Catanzaro, MD[2], Ted Ganiats, MD[2]
[1]Partnership for Health Analytic Research, LLC, Beverly Hills, California, [2]University of California, San Diego, San Diego, California

MOTIVATION

Overview

- Diagnostic & screening tests often evaluated by # of:
  - True positives (TPs)
  - False positives (FPs)
  - True negatives (TNs)
  - False negatives (FNs)
- These outcomes are largely determined by:

Sensitivity & Specificity

- For tests measured on *continuous* scale, sensitivity & specificity are largely determined by:

Positive Test Threshold

Relative Effects of Sensitivity & Specificity in Infectious Disease

Test characteristics	High Sensitivity	High Specificity
Impact on results	↑ True positives	↑ True negatives
Testing goal	• Identify people with disease	• Identify people without disease
Treatment goal	• Treat infection • Prevent future illness & disease spread	• Avoid unnecessary treatment
Potential harms of opposite test characteristic	*Low sensitivity:* ↑ False Negatives ↑ Potential future illness and suffering ↑ Potential future spread of disease	*Low specificity:* ↑ False Positives ↑ Bodily harms, toxicity, and financial costs of unnecessary treatment ↑ Social stigmatization ↓ Confidence in screening program

Sensitivity, Specificity, & TP & FP Balance

- Sensitivity & specificity are negatively associated within a given test:
  - ↑ Sensitivity (e.g., by changing positive test threshold) leads to:
    - ↓ specificity and ↑ TPs & FPs
  - ↑ Specificity leads to:
    - ↓ sensitivity and ↓ TPs & FPs
- Sensitivity-specificity **balance** sought between # FPs tolerated per additional TP gained
- Many factors affect:
  - the **relative effects** of sensitivity & specificity
  - the **balance** of true & false results
  - **decisions** regarding optimal thresholds
- Interaction of two factors – **disease prevalence** and **positive threshold** – cause results to differ in **high-** *vs.* **low-prevalence** settings
- Across settings:
  - disease prevalence varies
  - positive thresholds are set uniformly

Prevalence affects the performance of a test, and this study evaluates the magnitude of that impact to see if it has potential policy significance.

OBJECTIVE

To estimate the impact of disease prevalence in decisions regarding positive thresholds & test strategies, by:
Applying two simple models:

a) A hypothetical generic model
b) A worked example of screening for latent tuberculosis infection (LTBI)

in settings of varying prevalence

METHODS

Generic Model

- We modeled # TPs & FPs in scenarios defined by:
  - Test sensitivity: 50%, 60%, 70%, 80%
  - Test specificity: 90%, 95%, 98%, 99%
  - Disease prevalence: 20%, 40%, 60%, 80%
- We calculated results as:
  - TP = sensitivity * prevalence * *N*
  - FP = (1-specificity) * (1-prevalence) * *N*

where *N* = 1,000 hypothetical individuals

Tuberculosis Model

- We estimated TPs & FPs when switching between two tests for latent tuberculosis infection (LTBI):
  - In-tube QuantiFERON-TB Gold (QFT-IT)*

  to:
  - T-SPOT.TB*
- In 5 countries of varying LTBI prevalence**

* QFT-IT: *Cellestis*, Carnegie, Australia; T-SPOT.TB: *Oxford Immunotec*, Oxford, U.K.
** World Health Organization Global TB Database

RESULTS & IMPLICATIONS

Generic Model

Disease Prevalence	# True Positives when Varying Test Sensitivity: 50%	60%	70%	80%	# False Positives when Varying Test Specificity: 90%	95%	98%	99%
20%	100	120	140	160	80	40	16	8
40%	200	240	280	320	60	30	12	6
60%	300	360	420	480	40	20	8	4
80%	400	480	560	640	20	10	4	2

- Increasing sensitivity increased true positives
- Increasing specificity decreased false positives

So the absolute impact of:

- Sensitivity was greater in high-prevalence settings
- Specificity was greater in low-prevalence settings

Tuberculosis Model

Country	LTBI prevalence	Change in test outcomes with: *7% ↑ in sensitivity, 11% ↓ in specificity*: ↑ in TPs	↑ in FPs	FP / TP
U.S.	**5%**	329	10,483	31.9
Mexico	**29%**	2,018	7,829	3.9
Brazil	**39%**	2,712	6,739	2.5
Thailand	**47%**	3,272	5,859	1.8
Ivory Coast	**55%**	3,823	4,992	1.3

With greater prevalence:
•7% increase in sensitivity increased true positives.
•11% decrease in specificity decreased in false positives.
So:
•Settings with lower prevalence would have to pay a "price" of accepting more false positives for each true positive gained than would settings of higher prevalence.

Consider implications for two different settings:

	Two Countries: *Developed*	*Not Developed*
Health Care Access	Better	Limited
TB Prevalence	Low	Higher
Resistant TB	Rare	More common

→ ***BOTH face TRADEOFF introducing* T-SPOT.TB:** 11% ↓ specificity, 7% ↑ sensitivity

- For the developed country, the 7% increase in early detection may benefit too few people to justify the high burden of false positives.
- For the developing country, with higher disease prevalence, the greater increase in early detection may be worth the increased treatment of false positives
- *However*, this is not to say that the trade-off is not worthwhile in the developed country, or that it is worthwhile in the developing country
  - Resources and local priorities and values should determine that.
- Rather, the tradeoff may differ by orders of magnitude between settings, as prevalence varies.

CONCLUSIONS

In summary:

Positive test thresholds tend to be set globally
⇒ This has unintended consequences
Within a given test, sensitivity & specificity vary with positive test thresholds
⇒ This results in different outcomes between settings

Therefore we conclude that:

Decisions regarding positive test thresholds within tests should be made locally not globally
and
Strategic decisions between tests should be made locally not globally
…by incorporating disease prevalence (along with other factors)