Perils and Pitfalls in Clinical Trials of Diagnostic Tests for Tuberculosis

Richard O'Brien, MD
Foundation for Innovative New Diagnostics
Geneva

Outline of Presentation

- Statement of the problem
- Common errors in diagnostic clinical trials
- Suggested approaches
  - Trials of new case detection tools
  - Trials of new drug susceptibility testing methods
  - Trials of tests for latent TB infection

TB Diagnostics: Trial Design and Regulatory Perspective

- No internationally accepted standards

Application of Methodological Standards in Evaluation of Diagnostic Tests*

- Specify spectrum of evaluated patients (27%)
- Report subgroup results (8%)
- Avoid workup (verification) bias (46%)
- Avoid review bias (38%)
- Provide numeric precision of test indices (11%)
- Report indeterminate results (22%)
- Specify test reproducibility (23%)

*Reid MC, et al. JAMA 1995;274:645-51.

TB Diagnostics: Trial Design and Regulatory Perspective

- No internationally accepted standards
- Diagnostics not regulated in most countries

TB Diagnostics: Trial Design and Regulatory Perspective

- No internationally accepted standards
- Diagnostics not regulated in most countries
- Standards needed
  - To promote improved clinical trials
  - To assist regulatory authority
  - To inform technical guidelines (e.g., WHO)
  - To guide potential purchasers

Standards for Reporting of Diagnostic Accuracy (STARD) Initiative (1)*

- Introduction
  - Identify study as one of diagnostic accuracy
  - State research questions/study aims
- Methodology (1)
  - Participants
    - Study population
    - Participant recruitment
    - Participant sampling
    - Data Collection

*Bossuyt PM, et al. Clin Chem 2003;49:1-6.

Standards for Reporting of Diagnostic Accuracy (STARD) Initiative (2)

- Methodology (2)
  - Test methods
    - Choice of reference standard and rationale
    - Technical aspects of material and methods
    - Define rationale for units, cut/offs, categories
    - Number, training, expertise of staff
    - Blinded reading of test and reference standard
  - Statistical methods
    - Methods for calculating test accuracy
    - Methods for calculation test reproducibility

Standards for Reporting of Diagnostic Accuracy (STARD) Initiative (3)

- Results
  - Participants
    - When was study done
    - Clinical and demographic characteristics
    - Number eligible that were not enrolled
  - Test results
    - Time from test to reference standard (and Rx)
    - Distribution of severity of disease (and non-disease)
    - Cross tabulation of results of test and reference std.
    - Any adverse events

Standards for Reporting of Diagnostic Accuracy (STARD) Initiative (4)

- Results (2)
  - Estimates
    - Diagnostic accuracy and statistical uncertainty
    - How indeterminate results handled
    - Estimates of variability of accuracy in subgroups
    - Estimates of test reproducibility
- Discussion

Errors in TB Diagnostic Trials (1)

- Failure to assess test in correct population
  - Overestimating test sensitivity
    - Non-hospitalized vs hospitalized patients
    - AFB+ vs paucibacillary TB
    - HIV- vs HIV+ patients

Analysis of SeroDxTic Tests Among 465 Hospitalized Patients in Botswana, 2002

Test	Sens	Spec	PPV	NPV
ICS (PATH)	27	75	39	63
Osborn Sci	37	63	36	64
MycoDot	3	99	56	65
Amer Biologics	0	99	0	64
Anda Bio (Integ)	63	39	38	64
Omega DxTics	43	52	33	63

Talbot EA, et al. Clin Infect Dis 2004;39:e1-e7

Errors in TB Diagnostic Trials (2)

- Failure to assess test in correct population
  - Overestimating test sensitivity
    - Non-hospitalized vs hospitalized patients
    - AFB+ vs paucibacillary TB
    - HIV- vs HIV+ patients
  - Overestimating test specificity
    - Inappropriate control group (healthy controls)

MPB-64 patch test

TB patients vs healthy controls

	MPB64 patch test		
	Pos	**Neg**	**Total**
TB Patients	**52**	**1**	**53**
Healthy controls	**0**	**43**	**43**
Total	**52**	**44**	**96**

Nakamura, RM. Int J Tuberc Lung Dis. 1998;2:541

Analysis of SeroDxTic Tests Among 465 Hospitalized Patients in Botswana, 2002

Test	Sens	Spec	PPV	NPV
ICS (PATH)	27	75	39	63
Osborn Sci	37	63	36	64
MycoDot	3	99	56	65
Amer Biologics	0	99	0	64
Anda Bio (Integ)	63	39	38	64
Omega DxTics	43	52	33	63

Talbot EA, et al. Clin Infect Dis 2004;39:e1-e7

Errors in TB Diagnostic Trials (3)

- Inadequate blinding
  - Patients and controls identified
  - Results of gold-standard test known
- Improper resolution of discrepant results
- Failure to consider indeterminate results
- Too small a sample size

FOUNDATION FOR INNOVATIVE NEW DIAGNOSTICS

FASTPlaque RIF Response Test Direct Test on AFB+ Sputa

- Sensitivity = 100% (11/11)
- Specificity = 100% (134/134)
- Results reported in 2 days
- Conventional method took a mean ± SD of 33.2 ± 7.2 days

Albert H, et al. Int J Tuberc Lung Dis 2004;8:1114-9.

FASTPlaque RIF Response Test

	MGIT + indirect 7H11 proportion method				
FASTPlaque	Resistant	Susceptible	MGIT Culture-negative	Contam**	Total
Resistant	10	1#	0	1	12
Susceptible	0	134	1	14	149
RIF- <100 plaques*	4	16	5	2	27
Contaminated	0	4	0	3	7
Total	14	155	6	20	195

* less than 100 plaques obtained on the RIF- plate
** contaminated on either MGIT culture or 7H11 susceptibility test
# this specimen was found to be rifampicin resistant upon repeat testing by the proportion method

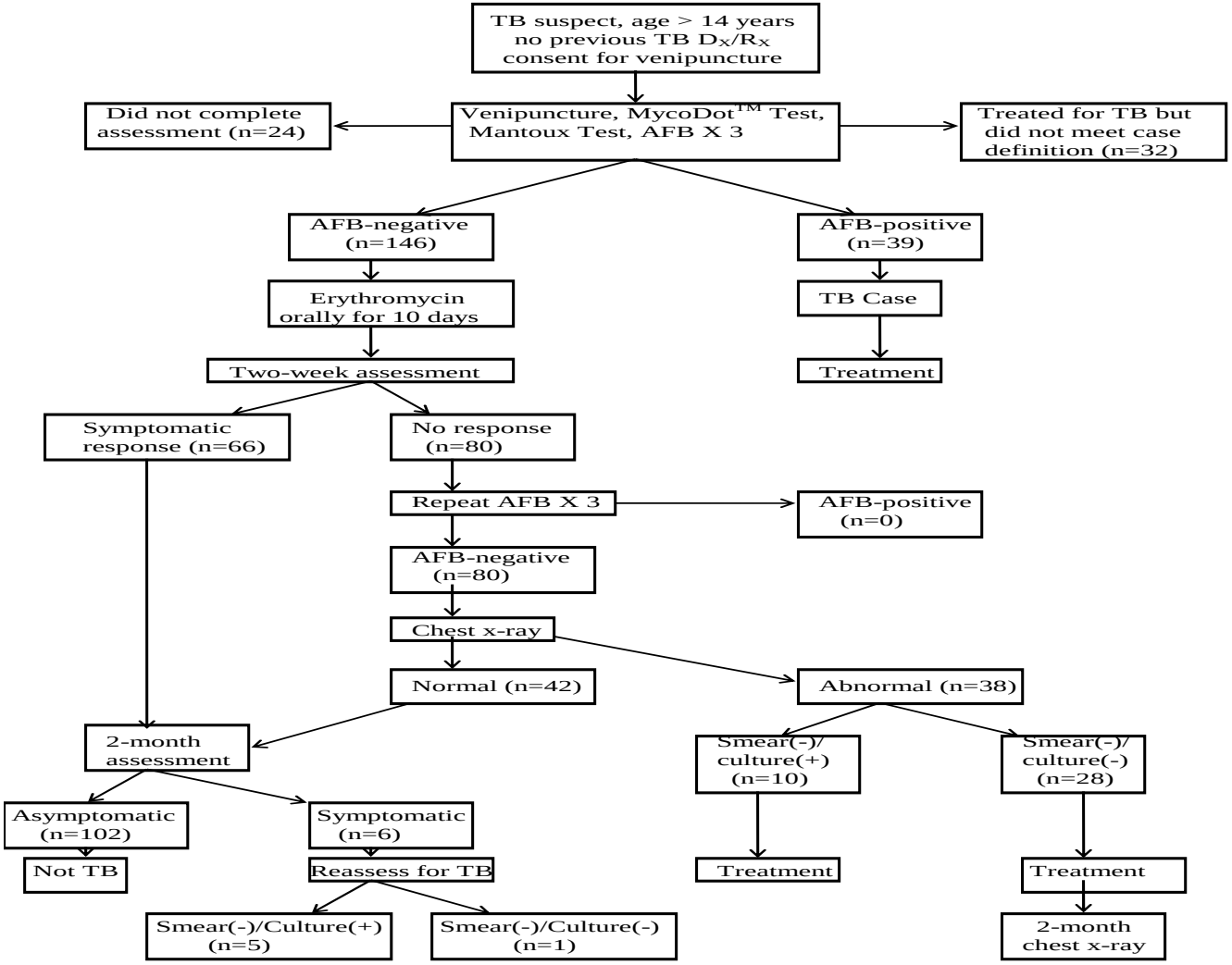
Approach to Trials of New Case Detection Tools

- Evaluation in TB suspects (HIV+/-)
- Multiple trial sites
- Gold standard
  - Culture on solid and liquid media
- Rigorous assessment of all patients for proper assignment
  - Include “clinical” cases
  - Follow-up of patients not diagnosed with TB

TB Case Definitions

- AFB+: ≥ 2 AFB+ sputum smears or 1 AFB+ smear and compatible CXR
- AFB-/culture+: ≥ 1 culture with ≥ 1+ growth
- AFB-/culture-: ≥ 3 negative smears/cultures, compatible signs/symptoms of TB, compatible CXR, +TST, clinical/CXR response to treatment
- Non-TB case: negative bacteriology, stable CXR and/or resolution without TB treatment and/or alternate diagnosis

FOUNDATION FOR INNOVATIVE NEW DIAGNOSTICS



Approach to Trials of Drug Susceptibility Testing Methods

- Study in population at risk of DR, e.g., failures of Category I treatment
- FDA gold standard: proportional method on 7H10 solid medium
- Resolution of discrepant results, e.g., rpo*B* gene sequencing
- Economic analyses important

FIND/TDR Study of Rapid DST Methods in Peru

- To enroll AFB+ patients (1000 new, 250 retreatment): expected 150 MDR TB cases
- Test methods: direct FASTPlaque RIF Response, INNO LiPA, and LJ; indirect MTT
- Standard method: indirect LJ proportional
- All isolates with any RIF-R plus 10% of others: rpo*B* sequencing, DST on 7H10

Approach to Trials of Tests for Latent TB Infection (1)

- Calculation of test sensitivity and specificity problematic
  - New/previously treated TB patients
  - TB contacts with defined exposure risk
  - Persons at low risk for LTBI
- Gold standard (PPD TST) inadequate
  - Sensitivity measured in TB patients, poor specificity

Study of QuantiFERON-2 and Tuberculin Skin Test in TB Contacts*

- High school outbreak in Denmark
- 125 contacts evaluated
  - BCG-unvaccinated (85)
    - Excellent agreement between QFT and TST in both high (93%) and low (95%) groups
  - BCG-vaccinated (40)
    - QFT positivity in high (53%) and low (6%) groups similar to that see among non-vaccinated groups

*Brock I, et al. Am J Respir Crit Care Med 2004;170:65-9.

Evaluation of ELISPOT TB Assay In TB Contacts in School Outbreak*

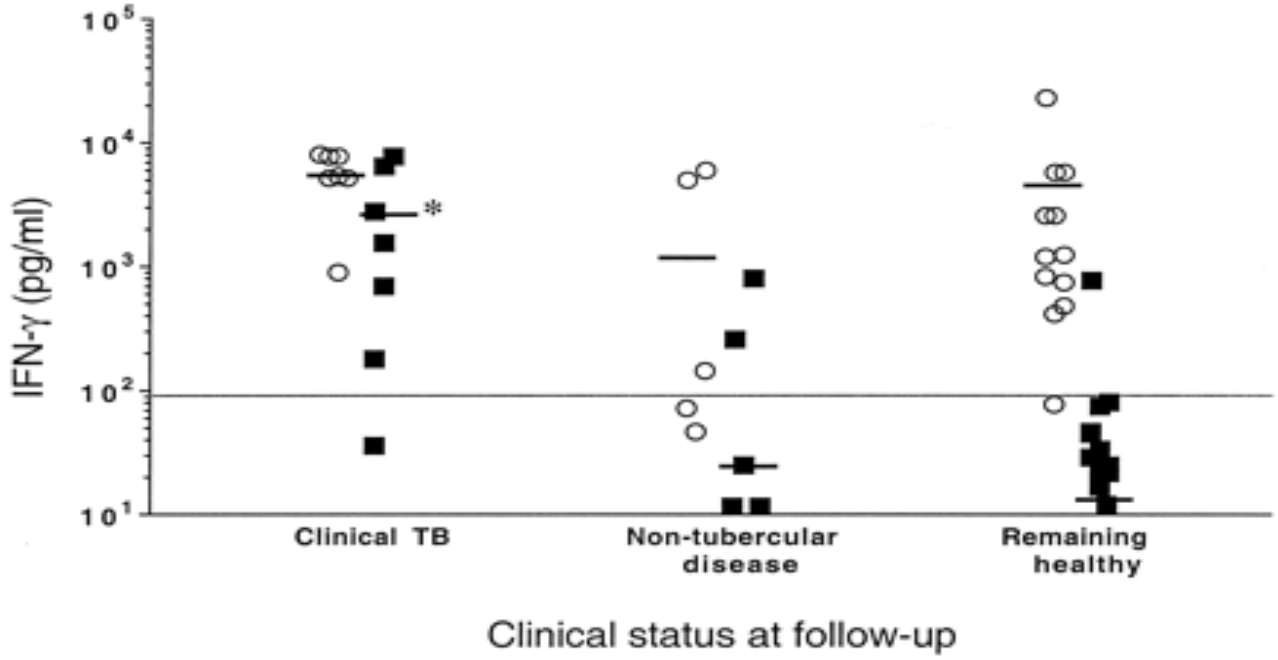
- 535 students in a school outbreak in UK
- TST (Heaf test) and ELISPOT TB assay
- Four exposure groups defined
- Results of ELISPOT more significantly correlated with exposure than TST
  - OR 2.78 (2.22-3.48) vs 2.33 (1.88-2.88), p = 0.03

*Ewer K, et al. Lancet 2003;361:1168-73.

Approach to Trials of Tests for Latent TB Infection (2)

- Calculation of test sensitivity and specificity problematic
  - New/previously treated TB patients
  - TB contacts with defined exposure risk
  - Persons at low risk for LTBI
- Gold standard (PPD TST) inadequate
  - Sensitivity measured in TB patients, poor specificity
- Large cohort studies needed to demonstrate relationship between test result and future TB disease

Immune Response to ESAT-6 in TB Contacts*



Dark squares = ESAT-6, open circles = PPD

*Doherty TM, et al. J Clin Microbiol 2002;40:704-6.