

Challenges in measuring impact of new diagnostics

Outcomes, context & designs

Samuel G. Schumacher - McGill University
samuel.schumacher@mail.mcgill.ca

Guidance for Accuracy Studies

STARD, DEEP, QUADAS, Cochrane DTA

ACADEMIA AND CLINIC

Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative

Patrick M. Bossuyt, Johannes B. Reitsma, David E. Bruns, Constantine A. Gatsonis, Paul P. Glasziou, Les M. Irwig, Jeroen G. Lijmer, David Moher, Drummond Rennie, and Henrica C.W. de Vet, for the STARD Group*

STARD
Reporting



EVALUATING DIAGNOSTICS

Evaluation of diagnostic tests for infectious diseases: general principles

DEEP
Design & conduct

Annals of Internal Medicine | RESEARCH AND REPORTING METHODS

QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies

Penny F. Whiting, PhD; Anne W.S. Rutjes, PhD; Marie E. Westwood, PhD; Susan Mallett, PhD; Jonathan J. Deeks, PhD; Johannes B. Reitsma, MD, PhD; Mariska M.G. Leeflang, PhD; Jonathan A.C. Sterne, PhD; Patrick M.M. Bossuyt, PhD; and the QUADAS-2 Group*

QUADAS-2
Quality assessment



Published on The Cochrane Collaboration website (<http://www.cochrane.org>)

Cochrane Handbook for Diagnostic Test Accuracy Reviews

Cochrane DTA Handbook
Systematic Reviews

What are the challenges?

Outcomes, context, designs

1. "A major obstacle to conducting impact assessments is **a lack of consensus on what 'impact' really means**, and what patient-, population-, health systems-, and epidemiology-important **outcomes** should be measured to decide on impact." INT J TUBERC LUNG DIS 14(12):1506-1507 © 2010 The Union EDITORIAL
2. "There is also **a lack of guidance on the methods (i.e. study designs) to be used** to measure them, which methods will be most rapid and cost-effective" **Assessing the impact of new diagnostics on tuberculosis control**
3. "[...] the impact of a new, more accurate diagnostic tool on tuberculosis epidemiology [is] **substantially affected by contextual factors** unrelated to tool performance."

Challenges & Lenses

1. What is impact / which **outcomes**?
2. How to consider **context**?
3. How to study it / which **designs**?

1. **test developer**
2. **researcher**
 - primary research
 - systematic reviews
3. **policy maker / guideline developer**
 - global
 - country level
4. **Funders**

Challenges & Lenses

1. What is impact / which **outcomes**?
2. How to consider **context**?
3. How to study it / which **designs**?

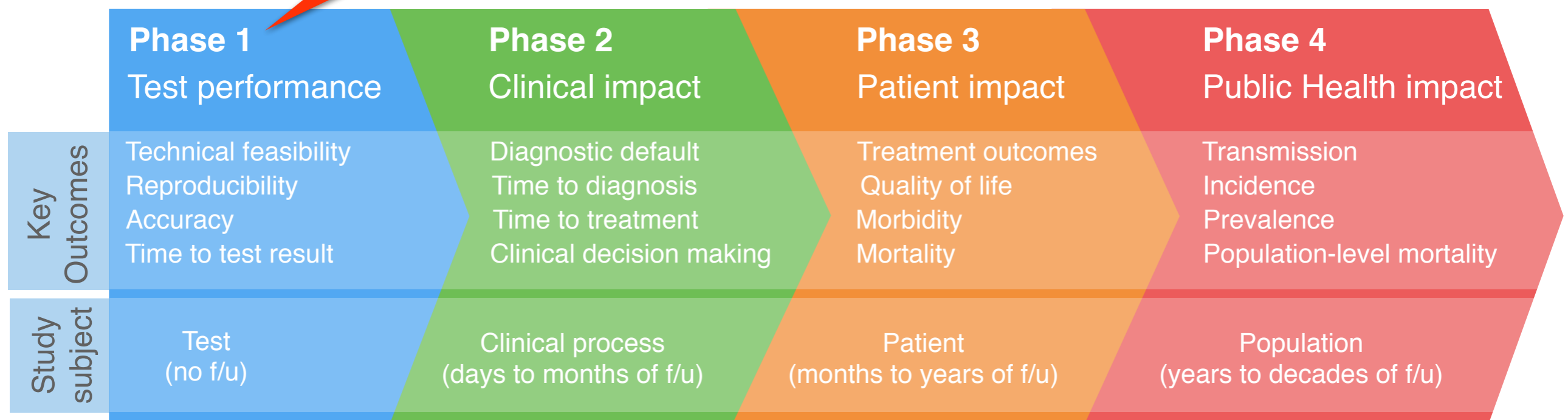
- 1. test developer**
- 2. researcher**
 - primary research
 - systematic reviews
- 3. policy maker / guideline developer**
 - global
 - country level
- 4. Funders**

1. What is impact /
which **outcomes**?

Phases of Diagnostic Research

Going beyond diagnostic accuracy

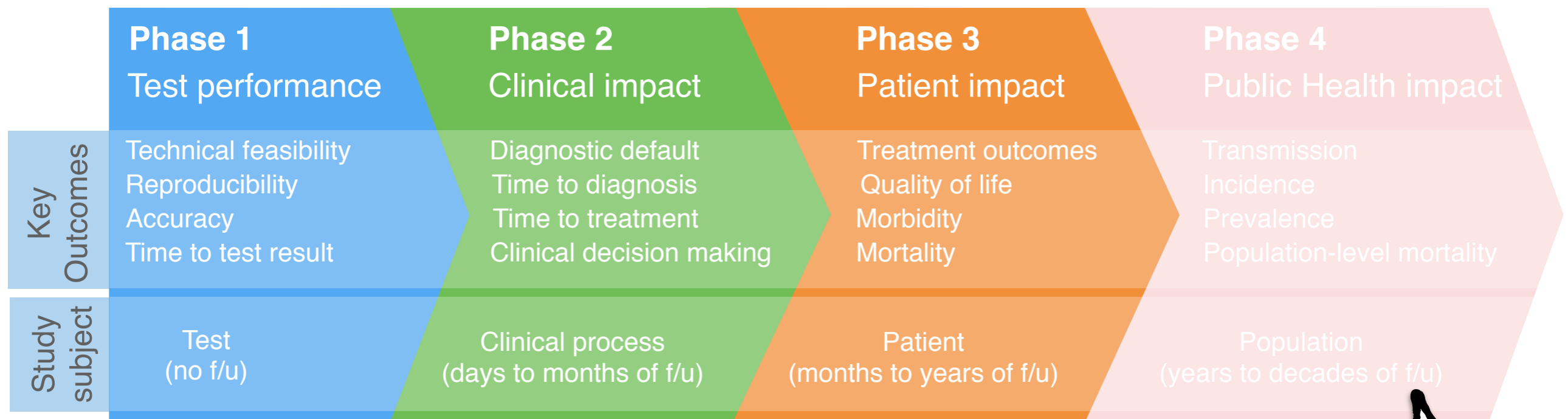
~95% of TB diagnostic research¹



¹Pai M, Brunet L, Minion J, Steingart K, Ramsay ARC, Lienhardt C. Mapping the landscape and quality of TB diagnostic research. Report to Stop TB Partnership. 2009 Nov 12;:1–24.

Tough choices...

which leap of faith are you willing to take?

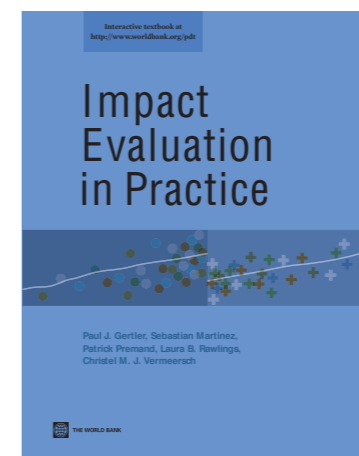
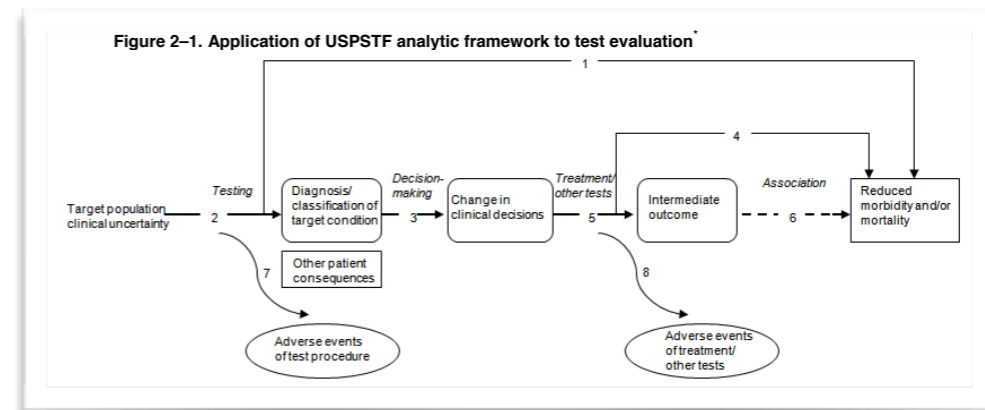
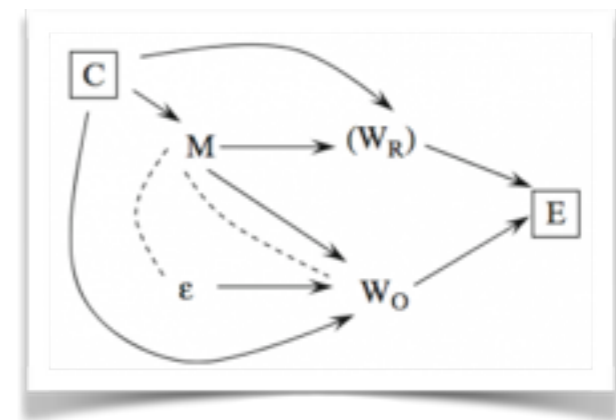
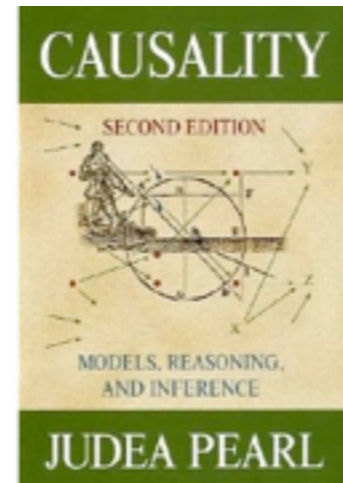


directness ↑
feasibility ↓
risk of bias ↑
generalizability ↓

What do methodologists say?

A quick look beyond the TB world...

- Epidemiologists
- Health Technology Assessment units
- Economists



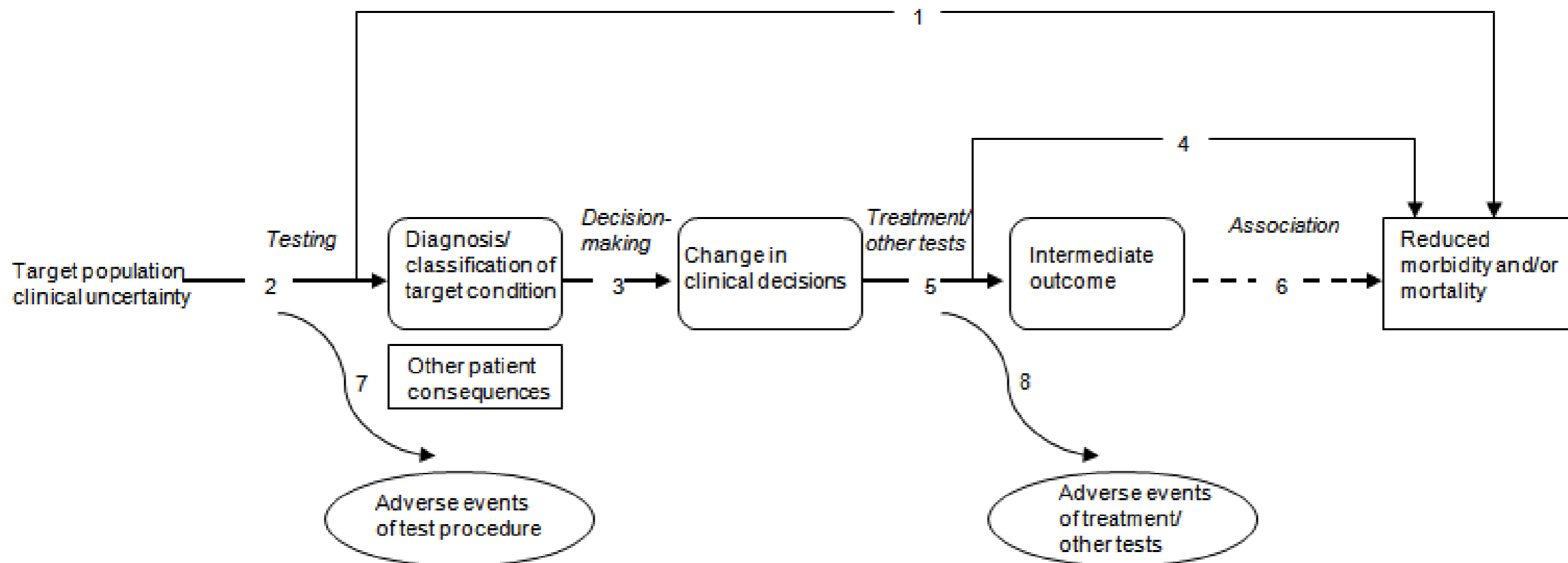
What it is and how to assess it

3 important points

1. impact = **causal effect** on **final outcomes**

- ▶ comparison group or extrapolation

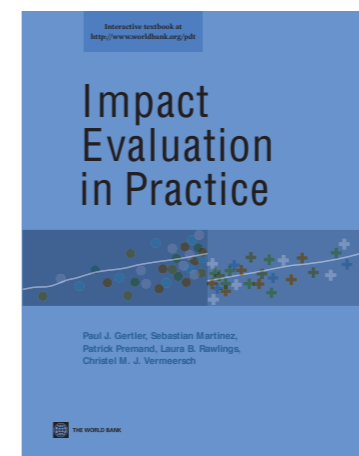
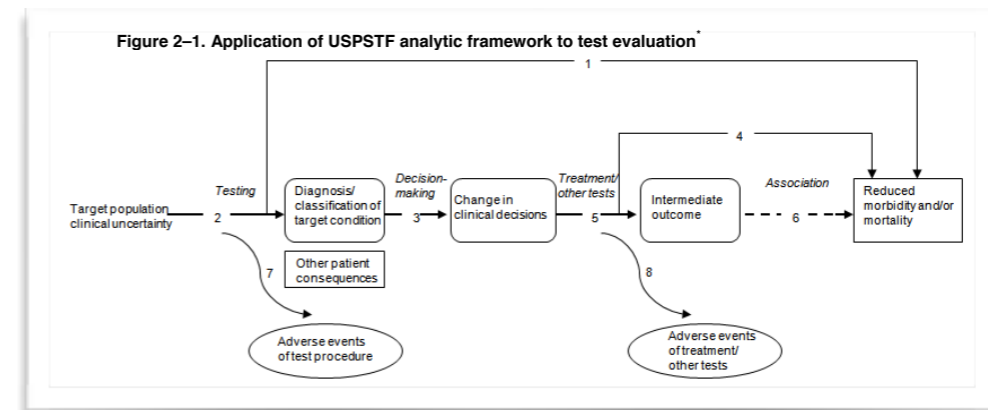
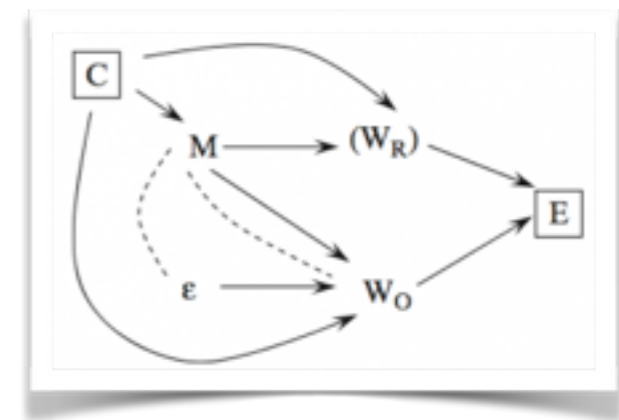
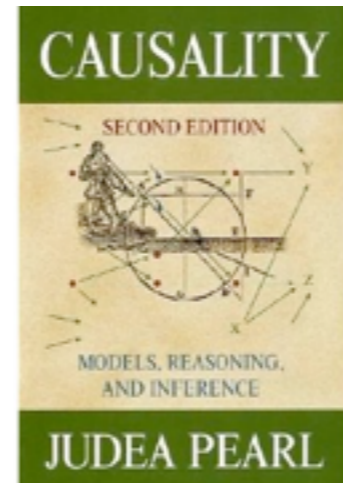
Figure 2–1. Application of USPSTF analytic framework to test evaluation*



What do methodologists say?

A quick look beyond the TB world...

- Epidemiologists
- Health Technology Assessment units
- Economists



Epidemiologists

Causal Inference

Causal diagrams

- ▶ tool to express and communicate beliefs about causal pathways

Causal inference

- ▶ counterfactuals as formal language & conceptual framework to understand bias

Methodology

- ▶ statistical methods for causal inference from observational data

Health Technology Assessment

Frameworks for evaluation & policy decision-making

Surrogate outcomes

- ▶ how to link evidence from various sources on intermediate outcomes

Big picture view

- ▶ cost-effectiveness, budget impact, societal consequences, organizational implications etc.

Policy decision-making

- ▶ pathway for decision about adopting new tests

Economists

Impact Evaluation

Evaluation of country-wide programs

- ▶ guidance specifically for this perspective

Methodology

- ▶ focus on RCTs and quasi-experimental designs

Mixed methods

- ▶ recognition of importance of qualitative methods

2. How to consider **context?**

Contextual factors affect impact

...beyond their effects on accuracy

★ epidemiological context

- HIV
- paucibacillary TB
- MDR-TB

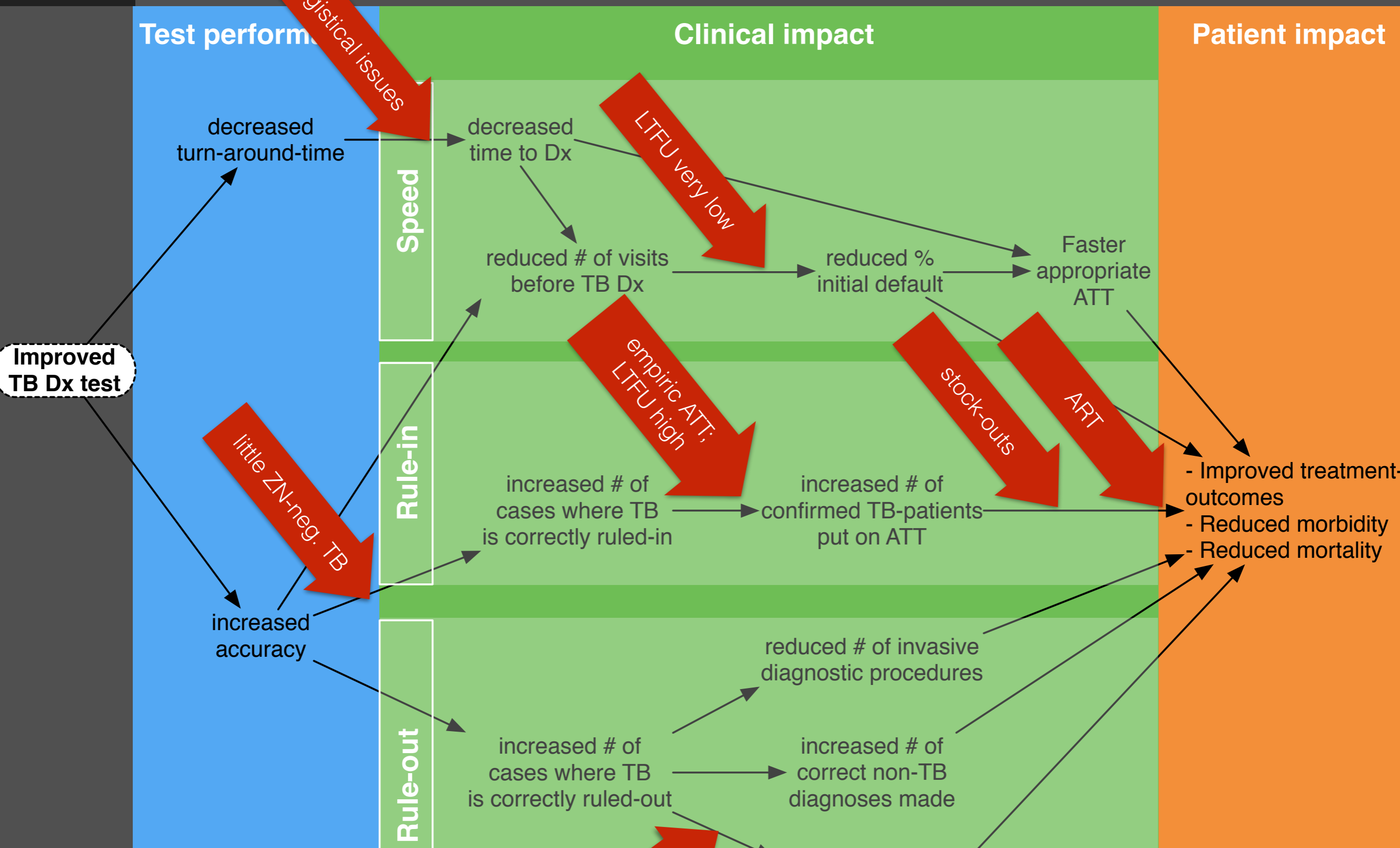
★ quality of baseline

- quality of smear microscopy
- empiric Rx

★ operational context

- “send the patient” vs “send the sample” model
- follow-up procedures, pre-Rx LTFU

Diagnostic Causal Chain & Spectrum of diagnostic research



Influence of contextual factors increases as we look further downstream!

3. How to study it /
which **designs**?

Study Designs

1. RCTs
2. Quasi-experimental studies
3. Diagnostic before-after studies
4. Hypothetical trials (uncontrolled / model-based)

RCTs

Types & Features

- allocation to ***intervention*** or ***no intervention*** assigned at random
- randomization leads to equal distribution of measured **and unmeasured** covariates
- **units of randomization** can be individuals or clusters (e.g. geographical, time)
- **explanatory** (efficacy) or **pragmatic** (effectiveness)

RCTs

Study Design Examples

OPEN ACCESS Freely available online

PLOS MEDICINE

A Multi-Country Non-Inferiority Cluster Randomized Trial of Frontloaded Smear Microscopy for the Diagnosis of Pulmonary Tuberculosis

Luis Eduardo Cuevas

OPEN ACCESS Freely available online

PLOS MEDICINE

LED Fluorescence Microscopy for the Diagnosis of Pulmonary Tuberculosis: A Multi-Country Cross-Sectional Evaluation

Luis Eduardo Cuevas^{1,2*}, Najla Al-Sonboli³, Lovett Lawson⁴, Mohammed Ahmed Yassin¹, Isabel Arbide⁵, Nasher Al-Aghbari³, Jeevan Bahadur Sherchand⁶, Amin Al-Absi⁷, Emmanuel Nnamdi Emenyonu⁴, Yared Merid⁸, Mosis Ifenyi Okobi⁹, Juliana Olubunmi Onuoha⁴, Melkamsew Aschalew⁸, Abraham Aseffa¹⁰, Greg Harper¹, Rachel Mary Anderson de Cuevas¹, Sally Jane Theobald¹, Carl-Michael Nathanson², Jean Joly², Brian Faragher¹, Stephen Bertel Squire¹, Andrew Ramsay²

Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial

Comparison of two methods for acquisition of sputum samples for diagnosis of suspected tuberculosis in smear-negative or sputum-scarce people: a randomised controlled trial



Jonathan G Peter, Grant Theron, Anil Pooran, Johnson Thomas, Melissa Pascoe, Keertan Dheda

Feasibility, accuracy, and clinical effect of point-of-care Xpert MTB/RIF testing for tuberculosis in primary-care settings in Africa: a multicentre, randomised, controlled trial



Grant Theron, Lynn Zijmah, Duncan Chanda, Petra Clowes, Andrea Rachow, Mala Lesosky, Wilbert Bara, Stanley Mungofa, Madhukar Pal, Michael Hoelscher, David Dowdy, Alex Pym, Peter Mwaba, Peter Mason, Jonny Peter, Keertan Dheda, for the TB-NEAT team*

Impact of Xpert MTB/RIF on Antiretroviral Therapy-Associated Tuberculosis and Mortality: A Pragmatic Randomized Controlled Trial

L Mupfumi^{1,2}, B. Makamure^{1,2}, M. Chirehwa², T. Sagonda², S. Zinyowera³, P. Mason^{1,2}, J. Z. Metcalfe^{4*} and R. Mutetwa^{2,4}

¹University of Zimbabwe College of Health Sciences, Harare, Zimbabwe; ²Biomedical Research and Training Institute, Zimbabwe; ³National Microbiology Reference Laboratory, Harare, Zimbabwe; and ⁴Division of Pulmonary and Critical Care Medicine, San Francisco General Hospital, and Francis J. Curry International Tuberculosis Center, University of California, San Francisco, California

Down

OPEN ACCESS Freely available online

PLOS MEDICINE

Impact of Xpert MTB/RIF for TB Diagnosis in a Primary Care Clinic with High TB and HIV Prevalence in South Africa: A Pragmatic Randomised Trial



Helen S. Cox^{1,2*}, Slindile Mbhele³, Neisha Mohess³, Andrew Whitelaw^{3,4}, Odelia Muller², Widaad Zemanay³, Francesca Little⁵, Virginia Azevedo⁶, John Simpson⁴, Catharina C. Boehme⁷, Mark P Nicol¹

Xpert MTB/RIF vs microscopy as the first line TB test in South Africa: mortality, yield, initial loss to follow up and proportion treated. The XTEND study

GJ Churchyard
On behalf of the XTEND team
(Xpert for TB - Evaluating a New Diagnostic)

RCTs

Strengths & Weaknesses

Strengths

- avoids confounding, therefore analysis and causal inference straightforward

Weaknesses

- ethics
- cost, complexity, speed
- generalizability

Quasi-experimental studies

Types & Features

1. pre-post implementation studies

- without control group
- with control group (difference in differences)
- interrupted time-series

2. instrumental-variable estimation

- exploit natural experiment

3. regression-discontinuity

- exploit discontinuity

Quasi-experimental studies

Types & Features

1. pre-post implementation studies

- without control group
- with control group (difference in differences)
- interrupted time-series

2. instrumental-variable estimation

- exploit natural experiment

3. regression-discontinuity

- exploit discontinuity

Quasi-experimental studies

Study Design Examples

OPEN ACCESS Freely available online

PLOS ONE

A Multi-Site Evaluation of Innovative Approaches to Increase Tuberculosis Case Notification: Summary Results

Jacob Creswell^{1*}, Suvanand Sahu¹, Lucie Blok², Mirjam I. Bakker³, Robert Stevens⁴, Lucica Ditiu¹

OPEN ACCESS Freely available online

PLOS ONE

Impact of Xpert MTB/RIF Testing on Tuberculosis Management and Outcomes in Hospitalized Patients in Uganda

Christina Yoon^{1,2}, Adithya Cattamanchi^{1,2,3}, J. Lucian Davis^{1,2}, William Worodria^{4,5}, Saskia den Boon⁶, Nelson Kalema⁶, Wincelous Katagira⁶, Sylvia Kaswabuli⁶, Cecily Miller¹, Alfred Andama⁴, Heidi Albert⁷, Pamela Nabeta⁸, Christen Gray⁸, Irene Ayakaka⁶, Laurence Huang^{1,3}

Engaging the private sector to increase tuberculosis case detection: an impact evaluation study

Aamir J Khan, Saira Khawaja, Faisal S Khan, Fahad Qazi, Ismat Lotia, Ali Habib, Shama Mohammed, Uzma Khan, Farhana Amanullah, Hamidah Hussain, Mercedes C Becerra, Jacob Creswell, Salmaan Keshavjee

ORIGINAL ARTICLE

Accuracy and impact of Xpert MTB/RIF for the diagnosis of smear-negative or sputum-scarce tuberculosis using bronchoalveolar lavage fluid

Grant Theron¹, Jonny Peter¹, Richard Meldau¹, Hoosain Khalfey¹, Phindile Gina¹, Brian Matinyena¹, Laura Lenders¹, Gregory Calligaro¹, Brian Allwood¹, Gregory Symons¹, Ureshnie Govender¹, Mashiko Setshedi¹, Keertan Dheda^{1,2}

Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study



Catharina C Boehme, Mark P Nicol, Pamela Nabeta, Joy S Michael, Eduardo Gotuzzo, Rasim Tahirli, Ma Tarcela Gler, Robert Blakemore, William Worodria, Christen Gray, Laurence Huang, Tatiana Caceres, Rafail Mehdiyev, Lawrence Raymond, Andrew Whitelaw, Kalaiselvan Sagadevan, Heather Alexander, Heidi Albert, Frank Cobelens, Helen Cox, David Alland, Mark D Perkins

OPEN ACCESS Freely available online

PLOS ONE

LED-Fluorescence Microscopy for Diagnosis of Pulmonary Tuberculosis under Programmatic Conditions in India

Lord Wasim Reza^{1*}, Srinath Satyanarayana¹, Donald A. Enarson², Ajay M. V. Kumar¹, Karuna Sagili¹, Sujeet Kumar¹, Levi Anand Prabhakar¹, N. M. Devendrappa¹, Ashish Pandey¹, Nevin Wilson¹, Sarabjit Chadha¹, Badri Thapa¹, Kuldeep Singh Sachdeva³, Mohan P. Kohli^{1,3}

Quasi-experimental studies

Strengths & Weaknesses

Strengths

- possible to conduct as new tests are being implemented
- may be conducted rapidly using available routine data
- may reflect “real-life” use and practice

Weaknesses

- greater risk of bias (confounding, selection bias)
- therefore more complex analysis and greater difficulty in making strong inference
- temporal trends

Diagnostic Before-after Studies

Types & Features

- each patient is his own control
- outcome: **clinical decision-making**
 1. pre-test management plan
 2. post-test management plan
 3. actual management
- prospective: ask directly
- retrospective: look at what was done

Diagnostic Before-after Studies

Study Design Examples

INT J TUBERC LUNG DIS 15(5):641-646
© 2011 The Union
doi:10.5588/ijtld.10.0195

Does solid culture for tuberculosis influence clinical decision making in India?

N. Stall,* T. Rubin,[†] J. S. Michael,[‡] D. Mathai,[§] O. C. Abraham,[§] P. Mathews,[§] K. Thomas,[§] M. John,* P. Daley[§]

Use of Drug-Susceptibility Testing for Management of Drug-Resistant Tuberculosis, Thailand, 2004–2008

Eugene Lam, Sriprapa Nateniyom, Sara Whitehead, Amornrat Anuwatnonthakate, Patama Monkongdee, Apiratee Kanphukiew, Jiraphan Inyaphong, Wanlaya Sitti, Navarat Chiengsorn, Saiyud Moolphate, Suporn Kavinum, Narin Suriyon, Pranom Limsomboon, Junya Danyutapolchai, Chalinthorn Sinthuwattanawibool, and Laura Jean Podewils

INT J TUBERC LUNG DIS 15(3):405-407
© 2011 The Union

SHORT COMMUNICATION

Implementation of liquid culture for tuberculosis diagnosis in a remote setting: lessons learned

P. Hepple,* J. Novoa-Cain,[†] C. Cheruiyot,[†] E. Richter,[‡] K. Ritmeijer[§]

Effect of Nucleic Acid Amplification for *Mycobacterium tuberculosis* on Clinical Decision Making in Suspected Extrapulmonary Tuberculosis*

Renda Soylemez Wiener, MD; Phyllis Della-Latta, PhD; and Neil W. Schluger, MD, FCCP

ORIGINAL STUDIES

Evaluation of the Impact of Interferon-Gamma Release Assays on the Management of Childhood Tuberculosis

Daphne I. Ling, MPH,* Claire A. Crépeau, BScN,[†] Marieke Dufresne, RN,[‡] Shazia Khan, BA,[‡] Caroline Quach, MD, MSc,*[‡] Nandini Dendukuri, PhD,[§] Kevin Schwartzman, MD, MPH,*[¶] Dick Menzies, MD, MSc,*[¶] Larry C. Lands, MD, PhD,[†] and Madhukar Pai, MD, PhD*[¶]

Diagnostic Before-after Studies

Strengths & Weaknesses

Strengths

- avoids confounding since comparisons are within-person
- may allow straight-forward inference if link between management decision and patient outcome is clear

Weaknesses

- prospective: often difficult to carry out in practice
- retrospective: pre-test management decision may not be available
- reported pre-test management plans may not match what truly would have been done

Hypothetical Trials

Types & Features

- each patient is his own control
 - accuracy study + assumptions
 - **“formal extrapolation”**
- Analytic methods
 - simple hand-calculations
 - decision-analytic modelling

Hypothetical Trials

Study Design Examples

Test and Treat: A New Standard for Smear-Positive Tuberculosis

To the Editors:

THE PROBLEM

Despite its limited sensitivity, sputum smear microscopy is the principal method of tuberculosis (TB) diagnosis in high-burden settings around the world. It will likely remain so for many years because more cost-effective and affordable diagnostic techniques are unlikely to be developed soon. Optimizing smear microscopy by improving rates of testing, completion, and reporting is

Supported by National Institutes of Health Grants K23HL094141 (to A. Cattamanchi) and K23AI080147 (to J. L. Davis); National Center for Research Resources Grant NIH KL2RR024130 (for the technical support by UCSF Clinical and Translational Sciences Institute Career Development Program).

The authors have no conflicts of interest to disclose.

J. L. Davis and D. W. Dowdy contributed equally to this work.

Correspondence to: Adithya Cattamanchi, MD, MAS, San Francisco General Hospital, Room 5K1, 1001 Potrero Avenue, San Francisco, CA 94110 (e-mail: acattamanchi@medsfgh.ucsf.edu).

Integrated Strategies to Optimize Sputum Smear Microscopy

A Prospective Observational Study

Adithya Cattamanchi^{1,2}, Laurence Huang^{1,3}, William Worodria^{4,5}, Saskia den Boon⁶, Nelson Kalema⁶, Wincelaus Katagira⁶, Patrick Byanyima⁶, Samuel Yoo⁶, John Matovu⁴, Philip C. Hopewell^{1,2}, and J. Lucian Davis^{1,2}

Impact of GeneXpert MTB/RIF® on Patients and Tuberculosis Programs in a Low-

Burden Setting: A Hypothetical Trial

J. Lucian Davis^{1,2}, L. Masae Kawamura³, Lelia H. Chaisson¹, Jennifer Grinsdale³, Jihane Benhammou⁶, Christine Ho⁵, Anna Babst⁴, Houmpheng Banouvong³, John Z. Metcalfe^{1,2}, Mark Pandori⁴, Philip C. Hopewell^{1,2}, Adithya Cattamanchi^{1,2}

Clinical Infectious Diseases Advance Access published January 14, 2014

Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact

Hojoon Sohn^{1,2}, Abebech D. Aero^{1,2}, Dick Menzies^{1,2}, Marcel Behr^{1,2}, Kevin Schwartzman^{1,2}, Gonzalo G. Alvarez^{1,3}, Andrei Dan¹, Fiona McIntosh^{1,2}, Madhukar Pai^{1,2*}, Claudia M. Denking^{1,4*}

Hypothetical Trials

Strengths & Weaknesses

Strengths

- quick and low cost (no follow up required)
- flexible in assessing different scenarios/ algorithms
- can be used to extrapolate findings to diverse settings (contextualize evidence)

Weaknesses

- strong reliance on modelling assumptions

Concluding Remarks

- important to look **beyond accuracy**
- **epidemiologists, HTA units and economists** provide important insights on methodology
- use **causal diagrams** to inform choice of outcomes, design and consideration of contextual factors
- **context** matters and should be carefully considered from the outset
- a range of **designs** can be used, depending on the research question