Meta-analysis of diagnostic studies

Karen R Steingart, MD, MPH
Francis J. Curry International Tuberculosis Center
University of California, San Francisco
Berlin, 12 November 2010
[karenst@uw.edu]
Disclosure

• I serve as co-chair of the Evidence Synthesis subgroup of Stop TB Partnership’s New Diagnostics Working Group

• I am a member of the GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group

Overview

• Describe key steps in a systematic review/meta-analysis of diagnostic test accuracy studies

• Describe standard methods of meta-analysis of data from diagnostic studies

• Identify key references and tools for performing systematic reviews of diagnostic studies
Challenges with meta-analysis of diagnostic studies

• Diagnostic accuracy cannot adequately be summarized by one measure

• Considerable between-study heterogeneity is the rule and models of meta-analysis must account for this

An individual study of the diagnostic accuracy of a test...

...estimates the ability of the test to distinguish between those with disease (condition) and those without disease

...compares results of the index test with best available reference for classifying patients as having/not having disease

• Most studies report pairs of sensitivity and specificity
A systematic review/meta-analysis of data from diagnostic studies…

…appraises the quality of primary studies
…calculates an overall summary; considers both dimensions of test performance
…looks for and investigates possible reasons for inconsistency in results (heterogeneity)
…evaluates the impact of quality and other study characteristics on diagnostic accuracy
…stimulates new research questions

Meta-analyses (pooling) can increase the precision of the overall result

<table>
<thead>
<tr>
<th>Measures of test performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Index test +</td>
</tr>
<tr>
<td>Index test -</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Sensitivity = TP/(TP+FN)
Specificity = TN/(FP + TN)
Positive predictive value = TP/(TP + FP)
Negative predictive value = TN/(FN + TN)
Likelihood ratio positive = Sensitivity/(1 – Specificity)
Likelihood ratio negative = (1 – Sensitivity)/Specificity
Prevalence (proportion of people with disease in population to whom the test has been applied) = TP + FN/(TP + FP + FN + TN)
Key steps in a systematic review of diagnostic test accuracy

1. Definition of the objectives of the review
2. Study identification and selection
3. Assessment of study quality
4. Data extraction, analysis, and presentation
5. Interpretation of results


The objectives of the review

Population
Intervention
Comparison
Outcome
+ Purpose of the test/strategy
+ Study design
+ Reference standard

Richardson et al. The well-built clinical question: a key to evidence-based decisions. ACP Journal Club 1995;A-12
Sensible clinical question

Population: In patients suspected of active TB
Intervention: do commercial serological tests
Comparison: compared with sputum microscopy
Outcomes: improve sensitivity and specificity

What is the purpose of the test?

• Triage
  – minimize use of invasive or expensive test
• Add-on
  – improve diagnosis beyond what is already done
• Replacement
  – replace test that is harmful or costly

Bossuyt et al. BMJ 2006
2. Study identification and selection

- Search MEDLINE, EMBASE, the Cochrane Register of Diagnostic Test Accuracy Studies (under development)
- Search related diagnostic test accuracy reviews (for example HTA database, DARE etc)
- Check references of relevant studies/reviews
- Use highly sensitive (broad) search strategy
- Reflect key concepts of the review (focus on index test and target condition) in search
- Use a wide variety of search terms, both text words and database subject headings (MeSH terms)
- Routine use of search filters should generally be avoided!

Does Bleach Processing Increase the Accuracy of Sputum Smear Microscopy for Diagnosing Pulmonary TB?

Medline search


3. Assessment of study quality
4. Data extraction, analysis, and presentation

- Extract paired estimates of sensitivity and specificity
- Visually examine results of individual studies
- Calculate overall summary estimates using HSROC/bivariate meta-analysis
- Look for and investigate possible reasons for heterogeneity

http://ims.cochrane.org/revman
Forest plots of sensitivity and specificity, anda-TB IgG for the diagnosis of pulmonary TB

- One row is displayed for each study
- Extracted data are presented: TP, FP, FN, TN
- Data shown in the graph are also displayed numerically
- Each study result is given a box for a point estimate
- Horizontal line = confidence interval (CI); measures how much we think the result of the study varies with chance
  - The wider the CI, the less confident we are in the result
- We can judge whether results are consistent depending if CIs overlap

Steingart 2010 unpublished
Calculating an overall summary

• The hierarchical approach to SROC (HSROC) has emerged as the standard method

The hierarchical approach to SROC (HSROC)

• Hierarchical model allows for both within and between study variability

• Random effects allows for heterogeneity between studies
Metandi in Stata

The Stata Journal (2009)
9, Number 2, pp. 211–229

**metandi**: Meta-analysis of diagnostic accuracy using hierarchical logistic regression

Roger M. Harbord
Department of Social Medicine
University of Bristol
Bristol, UK
roger.harbord@bristol.ac.uk

Penny Whiting
Department of Social Medicine
University of Bristol
Bristol, UK

Abstract. Meta-analysis of diagnostic test accuracy presents many challenges. Even in the simplest case, when the data are summarised by a $2 \times 2$ table from each study, a statistically rigorous analysis requires hierarchical (multilevel) models that respect the binomial data structure, such as hierarchical logistic regression. We present a Stata package, metandi, to facilitate the fitting of such models in Stata. The commands display the results in two alternative parameterisations and produce a customisable plot. metandi requires either Stata 8.0 or above (which has the new command xtologit), or Stata 8.2 or above with gllamm installed.

Keywords: metandi, metandiplot, diagnostic meta-analysis, sensitivity and specificity, hierarchical models, generalised mixed models, glamm, xtologit, receiver operating characteristic (ROC), summary ROC, hierarchical summary ROC

Summary ROC plots for Anda-TB IgG for diagnosis of TB: (A) smear+ and (B) smear- pulmonary TB patients. Red squares are pooled sensitivity and specificity values. Steingart unpublished

- X axis displays specificity
- Y axis displays sensitivity
- A circle for each study
- Width of the circles is proportional to # patients in each study
Meta-DiSc: a software for meta-analysis of test accuracy data

Javier Zamora1, Victor Abraira1, Alfonso Muriel1, Khalid Khan2 and Arri Coomarasamy2

Address: 1Clinical Biostatistics Unit, Ramón y Cajal Hospital, Madrid, Ctra. Colón km 3.030 Madrid 28034, Spain and University of Birmingham and Birmingham Women’s Hospital, Edgbaston, Birmingham, UK

Email: Javier.Zamora@hrc.es; Victor.Abraira@hrc.es; Alfonso.Muriel@hrc.es; Khalid.Khan@bham.ac.uk; Arri.Coomarasamy@bch.nhs.uk

* Corresponding author

Published: 12 July 2006
Received: 13 March 2006
Accepted: 12 July 2006

This article is available from: http://www.biomedcentral.com/1471-2288/6/31

© 2006 Zamora et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Select plot and characteristics

[Diagram showing sensitivity analysis with data points and confidence intervals.]
Heterogeneity

- Refers to variation in results among studies
- May be caused by variation in
  - test thresholds (unique to meta-analyses of diagnostic tests)
  - prevalence of disease
  - patient spectrum
  - study quality
  - chance variation
- *When significant heterogeneity is present, summary estimates from meta-analyses may not be meaningful*

Exploring heterogeneity

- Subgroup (stratified) analyses
- Meta-regression analysis
5. Interpretation of results

- What are the consequences of using the test in terms of the numbers of TP, FP, FN, and TN?
- How applicable are the results?
- To what extent were the primary studies biased? If serious study limitations were identified, could these impact the results?
- What are the implications for research?
Summary

• Described key steps in a systematic review/meta-analysis of diagnostic test accuracy

• Described HSROC/bivariate meta-analysis of data from diagnostic studies

• Identified key references and tools for performing systematic reviews of diagnostic test accuracy
Going beyond diagnostic accuracy…

- Link evidence on diagnostic test accuracy to clinical practice
- Frame review questions so that influences of the clinical context on test performance are identified
- Use results of systematic reviews of diagnostic test accuracy as inputs into decision analyses
- Go beyond summary ROC curves to describe test performance in terms of the expected downstream benefits and harms of using a test

Diagnostic Test Accuracy and Clinical Decision Making Ann Intern Med. 2008;149:904-906

References and tools for meta-analysis

- Zamora. BMC Medical Research Methodology 2006, 6:31
- Cochrane Diagnostic Test Accuracy Working Group http://srdta.cochrane.org/
- http://www.teachepi.org/ Dr Pai’s website for learning and teaching epidemiology
- http://www.tbevidence.org/ Evidence-based TB diagnosis
- RevMan http://ims.cochrane.org/revman
With special thanks to

• Mariska Leeflang
• Madhu Pai
• Many others

Workshop on Meta-analyses of Diagnostic Test Accuracy, Montreal, May 2009