



[Overview](#)

[Checks](#)

[Comparison](#)

[Sensitivity](#)

[Generalizability & tests](#)

Validation

TB MAC/WHO Annual Meeting, Glion, Switzerland

Monday, 18 September 2017

Health Economics & Decision Science
School of Health & Related Research
University of Sheffield

Modelling should compare results to evidence not used in calibration

Pete Dodd



Overview

Checks

Comparison

Sensitivity

Generalizability & tests

Checklist items

- 1 Model results should be checked against local epidemiology, health service characteristics, and program budgets, as well as general TB epidemiology
- 2 Model sensitivity to assumptions should be checked
- 3 Results should be compared to other modelling results or empirical assessments where possible, or through consultation with stakeholders
- 4 Models should avoid broad claims of validity and actively test performance

Model results should be checked against local epidemiology, health service characteristics, and program budgets, as well as general TB epidemiology

My thoughts/interpretation

- Comparison with local data which may be difficult to formally include in a unified calibration target (different nature, reliability, etc.).
- Validity checking for derived/aggregated outputs, e.g. program costs.
- Reporting of meaningful and important statistics implicit in models: to check realism and assess drivers of conclusions.

Model results should be checked against local epidemiology, health service characteristics, and program budgets, as well as general TB epidemiology

Epidemiological example statistics (general, local, local):

- Lifetime disease risk (look out for Nick Menzies' work)
 - Empirical evidence exists; not infrequently mis-implemented.
 - Important driver of impact for many types of intervention.
- 'Proportion recent' - how much incidence is due to infection <2 yrs ago?
 - There is some expectation for what this should be in different settings.
 - Measure of the potential for indirect benefit from reduced transmission.
- Proportion of TB incidence in PLHIV

Others may depend on focus/setting
(e.g. proportion of DR-TB in previously treated cases etc.)

Overview

Checks

Comparison

Sensitivity

Generalizability & tests

Results should be compared to other modelling results or empirical assessments where possible, or through consultation with stakeholders

(Second clause needs wording tweak?)

My thoughts/interpretation

- Agreement of independent analyses adds confidence to conclusions
 - agreement doesn't mean correct; disagreement doesn't mean wrong!
 - understanding differences takes time and effort
 - unlikely to be a like-for-like comparison
- Consultation with stakeholders
 - buy-in and mutual understanding necessary for trust
 - consultation not at end, but with time to investigate/correct
- Proportion of TB incidence in PLHIV



My thoughts/interpretation

- Understanding that parametric sensitivity/uncertainty analysis covered by 'Evidence Synthesis' item.
 - (though rarely done well in infectious disease models)
 - here considering structural or non-parametric uncertainty
- Not checking the implications of an assumption known to be questionable = bad.
- Difficulty in identifying assumptions to check - can't check everything
 - Sensitivity analysis may help (but may not)
 - Consultation with stakeholders as a key method (expertise & perception)

[Overview](#)

[Checks](#)

[Comparison](#)

[Sensitivity](#)

[Generalizability & tests](#)

Models should avoid broad claims of validity and actively test performance

Pete Dodd



[Overview](#)

[Checks](#)

[Comparison](#)

[Sensitivity](#)

[Generalizability & tests](#)

My thoughts/interpretation

- First part ('claims'):
 - to avoid thoughtless application of results in different contexts...
 - ...and models designed/specialized for different purposes
 - Greater understanding of transferability between contexts (methods & necessity) useful
- Second part ('actively test'):
 - follow-up of post-decision data & retrospective evaluation?
 - testing/validity to understand the software development aspects of modelling?
(testing, debugging, version control, transparency)