HIV Modelling Consortium

Proposal for the Model Validation Exercise

Aim

To use the forthcoming Community Randomized-Control Trials (CRCTS) to determine if mathematical models can usefully anticipate the short-term impact of the interventions and improve validation.

Background

Mathematical models are often used to anticipate the impact of HIV prevention interventions but their ability to do this reliably has never been objectively tested. Several CRCTs will be initiated in 2013 that will measure the impact of HIV prevention interventions. Mathematical models have been used to predict the magnitude of these specific interventions and those calculations have informed the design of the studies. It should therefore be possible to store those predictions, and the models that were used to generate them, so that they can be compared with actual empirical observation at the end of the trials. Although this provides only a limited test of the capability of models -- in particular, it will only be possible to examine the short-term impact of the intervention -- we believe it is a valuable opportunity to challenge mathematical models.

Models are increasingly being used in the biomedical arena where validation standards are currently less rigorous and a methodology to quantify credibility is greatly sort after (1). The most important objective of this exercise is to begin to understand how model projections can succeed or fail in projecting the impact of interventions in CRCTs rather than to identify a 'winning' model.

Hypotheses to Test

In order to gain an understanding of how and when models projections can succeed or fail, we propose to test the following hypotheses.

1. *Do mathematical models provide the correct guidance about the impact of the intervention in the design phase (and interim phase, if appropriate) of the trials?*

   This is the hardest test and challenges the model that was actually developed under real constraints at the beginning of a trial and making assumptions often in the absence of the best data. It is a test of the "modelling process" in usefully informing trial design.

2. *Can the models predict the short-term impact of interventions once they are updated with information about particular aspects of the system that was not known*

when the models were first developed? For example, coverage of interventions, effects of intervention, of epidemiological data for the setting.

These are softer tests that allow the original model to be updated as new information becomes available. They test whether the essential "model structure" that was employed was suitable for making the relevant predictions.

3. *Can our very best, fully-updated models, predict the short-term impact of interventions?*

These relax the requirement to use the same model structure that was used at the beginning of the trial, recognizing that that choice will have been at least partly driven by time constraints. Here analysts can use a 'best' model, fully updated with information about the epidemiological context and the intervention coverage of the trial. It is a test of the state-of-the-field of HIV impact modelling.

Methods

The key to the process is to formally define the "models" that were used at the beginning of the trials (or could have been used at the beginning of the trials). The model protocols (that define all aspects of the model, parameters and analysis) will be documented and archived in a secure online site that is managed by a third party.

It is also necessary to define strict boundaries in the definition of the model between the "structure", the "parameters" (epidemiological and biological assumptions) and the "intervention assumptions" (see Table 1).

Mathematical Model Component	Description
Model Structure	The basic architecture of the model (e.g. differential equations, simulation algorithms) and the process used to generate a result (e.g. run the model 20 times and take average).
Epidemiology Parameters	Values that describe aspects of the epidemiology of the setting (e.g. HIV prevalence) or the biology of HIV (e.g. transmission probabilities).
Intervention Coverage Assumptions	Values that describe the extent to which the intervention is scaled-up in a community (e.g. coverage of testing, proportion virally suppressed etc.)

Mathematical Model Component	Description
Intervention Effect Assumptions	Values that describe the expected effect of particular interventions (e.g. behavior change following HIV testing).

Table 1: Four components of the mathematical model.

Any mathematical modeller can submit their model protocol to this exercise. It will be incumbent upon modellers who take part in this exercise to propose how their model can be split along these lines. The model protocol must also precisely describe how the model's parameters would be updated with new data (e.g. would refit to X using parameter Y, based on maximum likelihood as described). A working group will be convened to review these proposals (for clarity, specificity and exhaustiveness) before the protocols are finally archived.

Analysis Plan

We would then be able to conduct the following analyses, by keeping the "frozen" version of the model and updating different parts (Table 2).

Analysis	Question	Model Structure	Epidemiology Parameters	Intervention Coverage Assumptions	Intervention Effect Assumptions
1	Did mathematical models provide correct guidance about the impact of the intervention in the design phase (and interim phase, if appropriate) of the trials?	Frozen	Frozen	Frozen	Frozen
2a	Can the models that were used predict the short-term impact of interventions when they are updated with information about the intervention coverage that was attained?	Frozen	Frozen	Updated	Frozen

Analysis	Question	Model Structure	Epidemiology Parameters	Intervention Coverage Assumptions	Intervention Effect Assumptions
2b	Can the models that were used predict the short-term impact of interventions when they are updated with information about the intervention coverage that was attained and any new information about the effect of those interventions?	Frozen	Frozen	Updated	Updated
2c	Can the models that were used predict the short-term impact of interventions when they are updated with any new information about the coverage attained in the intervention, or the effect of the interventions, or the underlying epidemic conditions or the biology of HIV?	Frozen	Updated	Updated	Updated
3	How well can our best models predict the short-term impact of incidence?	Updated	Updated	Updated	Updated

Table 2: Proposed Analysis Plan. "Frozen" means that the model's original values (as of end of 2012 are used; "Updated" means that the model can be updated (in the manner specified in advance by the analyst) with new data.

No decision has been reached about the way in which "agreement" of the model will be quantified and articulated, as this will inevitably be an iterative exercise with multiple stages as we seek to understand the factors that allow or prevent models to predict the impact of interventions.

Process

The HIV Modelling Consortium will take responsibility for publicizing this exercise among mathematical modellers, convening a process of review of submitted protocols and archiving the protocols. In addition, the Consortium will assist in the assembly of a writing committee of those participating in the model validation exercise to take the lead in developing publication outputs relating to the task.

Timeline

- Early December: Announcement of this exercise.
- 1st March: Submission of model protocols.
- Late March (date TBC): Convening of a working group to review protocols.
- Mid April (two weeks post WG meeting): Revised protocols submitted.
- 29th April: Formal "lock-down" of the model protocols.

Contact:
HIV Modelling Consortium Project Manager, Ms Ellen McRobie: e.mcrobie@imperial.ac.uk

References:

(1) Sainani K. Getting it Right: Better Validation is the Key to Progress in Biomedical Computing. *Biomedical Computation Review* 2012. Available at: http://www.biomedicalcomputationreview.org/content/getting-it-right-better-validation-key-progress-biomedical-computing (last accessed Dec 2012)