

Predictive Validation and the Re-Analysis of Cost-Effectiveness: Do We Dare to Tread?

Jonathan Karnon¹  · Hossein Hajiali Afzali¹

Published online: 22 August 2017
© Springer International Publishing AG 2017

Predictive validation (also known as fourth-order validation) involves the comparison between model outputs and observed data collected after the initial analysis of a model. Given that the data might be collected years after the development of the model, this type of model performance evaluation is the rarest form of validation applied to cost-effectiveness models. In a review of model performance evaluation across 81 published cardiovascular cost-effectiveness models, predictive validity was not reported in any of the reviewed papers [1].

This is not a surprising finding—funding for the development and analysis of cost-effectiveness models and for the collection of prospective data rarely cover the period beyond the primary analysis of a model. However, predictive validation appears to be in short supply even when models and funding decisions are re-visited. The National Institute for Health and Clinical Excellence (NICE) has re-appraised a range of technologies between 3 and 8 years after the publication of original guidance [2]. The updated guidance for these appraisals makes no reference to the predictive validation of the original models.

What can predictive validation add to simply re-estimating a model to reflect the availability of additional data?

When new evidence becomes available, predictive validation can be used to identify sources of any inaccuracy in model extrapolations including structural aspects and model input parameter values. This is illustrated by the comprehensive set of analyses that were undertaken to

validate a relatively complex cost-effectiveness model of population-based screening for abdominal aortic aneurysms [3]. The model was originally developed to extrapolate the results of a clinical trial with a 4-year mean follow-up. The trial cohort were followed up for a further 3 years, at which point outputs from the original model were compared to observed outputs in the 3 additional years of follow-up. The fit of the original model was poor. The p values for ten of the 13 comparisons were ≤ 0.01 . The original model structure was revised to allow for the use of time-varying parameters and new model predictions derived by calibration produced a very close fit to the prospective data.

Kim and Thompson [3] conclude that their analyses provide evidence that cost-effectiveness models based on limited follow-up data “may be unreliable for decision-making”. This is important evidence, which supports the review of funding decisions on a more regular basis. NICE have reviewed only a small proportion of its technology appraisals. The Pharmaceutical Benefits Advisory Committee (PBAC) in Australia has only recently implemented a Managed Entry scheme to fund pharmaceuticals with highly uncertain clinical effectiveness whilst additional data are collected to inform a final funding (and pricing) decision. The more regular assessment of predictive validity may inform more accurate and efficient value-based pricing.

More widespread evidence on the predictive validity of cost-effectiveness models may also identify condition-, technology- and model-related factors that are associated with poor predictive validity. Such evidence would also inform the importance of efforts to assess model performance concurrently through assessments of internal and external validity. Kim and Thompson [3] reported

✉ Jonathan Karnon
jonathan.karnon@adelaide.edu.au

¹ University of Adelaide, Adelaide, SA, Australia

divergence between their analyses of internal, external and prospective validity. Is this finding generalizable or does internal and external validation improve predictive validity?

Series of validations across models for similar conditions may inform modeling approaches for subsequent analyses of treatments for similar conditions. An area in which such series of predictive validation may have a particularly important role is in the validation of cost-effectiveness models for systemic treatments for advanced cancers. There is generally significant uncertainty around the extrapolation of survival curves in these populations. The scheduled review and re-estimation of predicted survival curves will reduce uncertainty around the value of these high-cost drugs. In addition, series of predictive validations of applied approaches to estimating overall survival in advanced cancer may inform circumstances in which the available data are too immature to be usefully extrapolated. Analyses of predictive validity may also inform structural aspects of the extrapolation process, such as the choice of timepoint from which to extrapolate or credible rates of change in the predicted hazards.

Predictive validation is best informed by the planned prospective collection of relevant data beyond the initial development and analysis of a cost-effectiveness model. As noted, reimbursement authorities such as NICE and PBAC require continued data collection for a small number of new health technologies and so processes are in place to design and implement studies to collect relevant data.

The question of whether reimbursement bodies should revisit a higher proportion of funded health technologies remains. In Australia, the collection of prospective data as part of a managed entry is restricted to a small number of new pharmaceuticals with small eligible populations for which non-statistically significant estimates of clinical effectiveness are presented to the PBAC. Options for identifying other areas in which predictive validity might be assessed include the use of Expected Value of Information analyses to identify technologies with the greatest costs of uncertainty. It is also important to assess the costs and expected validity of the prospective data to be collected. The collection of data to inform clinical effectiveness may be costly and complex, whilst the collection of data to validate the extrapolation of observed clinical effects might be more feasible.

There are barriers to the expanded assessment of predictive validity and the associated re-assessment of the cost-effectiveness of new technologies. Despite the existence of processes to support managed entry and the collection of evidence to re-assess cost-effectiveness, the process involves complex negotiations and planning on a case-by-case basis. Such negotiations cover agreements

with respect to the data to be collected and the supply of the technology at the price implied by the revised cost-effectiveness analysis. There are also non-trivial costs associated with the planned collection of relevant additional data and the subsequent re-analysis of models.

Re-analysis with the potential for revising prices (either up or down) also increases uncertainty with respect to government budgets and company revenues. Politically, it may be costly to acknowledge decision uncertainty to the extent that continued data collection is required on a more regular basis. Such acknowledgements and the associated complexity of managed entry may support arguments to reduce reliance on cost-effectiveness modeling to inform reimbursement decisions.

Cost-effectiveness modelers wanting to investigate the potential value of predictive validation and the reappraisal of more new health technologies face the problem that such validation is difficult without the planned collection of relevant prospective data. That is our challenge. A recent study revisited the cost-effectiveness of an osteoporosis drug from an Australian perspective using data published since the PBAC funding recommendation, demonstrating that the drug was not cost effective [4]. The conduct of more studies such as this would begin to build an evidence base to inform the selection of new technologies for which predictive validity should be assessed with a view to the re-assessment of cost-effectiveness and the prices of new health technologies.

Compliance with Ethical Standards

Funding No funding relating to this editorial was received by the authors.

Conflicts of Interest JK is a member of the economic sub-committee of the Pharmaceutical Benefits Advisory Committee and HHA is a member of the economic sub-committee of the Medical Services Advisory Committee.

References

1. Afzali HH, Gray J, Karnon J. Model performance evaluation (validation and calibration) in model-based studies of therapeutic interventions for cardiovascular diseases. *Appl Health Econ Health Policy*. 2013;11(2):85–93.
2. National Institute for Health and Care Excellence. Guidance and advice list. <https://www.nice.org.uk/guidance/published?type=ta>. Accessed 7 Aug 2017.
3. Kim LG, Thompson SG. Uncertainty and validation of health economic decision models. *Health Econ*. 2010;19(1):43–55.
4. Karnon J, Shafie AS, Orji N, Usman SK. What are we paying for? A cost-effectiveness analysis of patented denosumab and generic alendronate for postmenopausal osteoporotic women in Australia. *Cost Eff Resour Alloc*. 2016;14(1):11.