MAIMON WORKING PAPER No. 27 DECEMBER 2020

LET A THOUSAND IMAGINARY MODELS BLOOM: ICER OPENS THE FLOODGATES TO PSEUDOSCIENCE

Paul C Langley, Ph.D., Adjunct Professor, College of Pharmacy, University of Minnesota

Abstract

It has been noted on numerous occasions that modeled claims for cost-effectiveness, if driven by assumption for the lifetime of a hypothetical patient population, can be easily 'gamed' to create a required claim. These marketing exercises to support product entry are all too common in the literature. The institute for Clinical and Economic Review (ICER) in its launch of the ICER Analytics platform has provided a framework to support precisely these activities. Following the mainstream methodology in health technology assessment, the ICER Analytics platform is to facilitate the creation of approximate information to support formulary decisions. This is an odd development because it undercuts ICERs belief that it is a key arbiter in health technology assessment, setting the stage for pricing and access recommendations. With the release of the ICER Analytics platform, others can now customize the 'backbone' ICER model in a disease area (i.e., change assumptions) to develop alternative value assessments. The problem is, of course, that without a reference point, there is no basis for comparing modeled claims other than challenging assumptions. Indeed, ICER has made this easy by reducing barriers to lifetime model building so that is easy for competing claims to be created. ICER will then become one of a multitude of competing voices for the attention of formulary committees and other health decision makers. Let a thousand imaginary models bloom.

INTRODUCTION

Demarcating science from pseudoscience rests on a simple premise: the ability of claims made to be credible, empirically evaluable and replicable [1]. Health technology assessment fails this test. Since the early 1990s it has focused on inventing claims for cost-effectiveness based on lifetime simulation models. Hypothesis testing has been rejected [2]. The reason for this denial of the standards of normal science is clear: it is easier, at product launch to drive claims for cost-effectiveness, to fill evidence gaps, with assumptions [3]. The alternative, to agree a research program to meet evidence gaps is too time consuming. It is far easier to create claims, from a lifetime simulation, that have no possibility of ever being empirically evaluated.

The Institute for Clinical and Economic Review (ICER) has accepted the approximate information meme. Its claims for pricing and product access are pseudoscience. But more interestingly is the extent of its failure to meet those standards. Not only do its claims lack credibility from the perspective of empirical assessment, but the claims themselves are mathematically impossible [4]. This

is because ICER has failed to appreciate the limitations imposed by the axioms of fundamental measurement. It holds to the belief, without a shred of evidence to support this position, that utility scales have ratio properties [5]. This allows the creation of quality adjusted life years (QALYS); the cornerstone of their reference case modeling. Unfortunately the utility scales were never designed to have ratio properties; they are ordinal scales. An ordinal scale cannot support multiplication; QALYs are therefore an impossible construct. Unfortunately, the misuse is even more egregious: the utility scale (such as the EQ-5D-3l) is a multiattribute scale. This means it lacks dimensional homogeneity. It cannot support a single score because the symptoms covered are each dimensionally unique. If any evidence was needed for the EQ-5D-3L failing to meet ratio standards all we have to note is that the utilities can take negative values; there is no true zero. Nor, it might be added does the EQ-5D-3L have interval properties. It cannot make any claims for response to therapy.

The deficiencies of I-QALY modelling are well known; yet ICER perseveres. After all, it is their business model. ICER refuses to accept the standards of normal science. Yet ICER has gone one step further; it has established ICER Analytics, a cloud based platform that allows those interested in imaginary worlds to modify an existing ICER model (i.e., change assumptions) to create alternative modeled claims [6]. More to the point, for those interested, ICER has all but eliminated barriers to building hypothetical lifetime simulations. Anyone can develop model variants to include both those who support modeled claims as well as though who may wish simply to demonstrate the key weakness: irrespective of how closely one observes 'standards' for model claims, it is always possible to 'create' a countervailing case.

OPEN SEASON FOR IMAGINARY MODELS

It would be fair to say that ICER has shot itself in the foot; as far as can be judged there are no constraints on ICER Analytics use. Open access to the model (all ICER reference case models) ensures the emergence of a multiplicity of models each claiming to be anchored in the ICER 'backbone' model. Pseudoscience proliferates. Manufacturers, dissatisfied with the ICER 'house' model for an evidence report may be quite willing to underwrite alternative modelled assumptions for the same or similar products to arrive at competing claims. The waters would not be muddied; they would be churned up. Published in peer reviewed journals, competing models would present decision makers with an interesting choice.

Unlike the UK and Australia, for example, the problem with an open season for simulated model claims is that there is no referee to 'judge' the contestants [7] [8]. With single payer health systems, modeled approximate evidence claims submitted for evaluation are subject to external review by academic groups schooled in assessing imaginary claims (don't ask). In the US, formulary committees subject to conflicting model claims based on the ICER 'backbone model' have no referee to assess their competing merits as imaginary constructs. The committees will not have the skills to unravel competing black boxes. They may just reject the applications which would be a wise move given that the models are an analytical dead end. In the US, the absence of a referee for competing imaginary

claims would be the equivalent of a football match without a referee. Welcome to ICER's imaginary playing field.

We should not underestimate the ability of consultants to develop alterative scenarios. The most obvious would be to choose an alternative multiattribute utility score to create 'competing' QALYs. Allied to an alternative definition of disease states and time spent, the result would be a quite different lifetime incremental QALY picture with an alternative classification and estimate of costs. Differential discount rates could also be applied. The options are wide open; as they will all be based on assumptions from the literature (or just guesses), all can be justified. The fact that whatever exercise is undertaken it will still fail the standards of normal science. Perhaps it will be beast viewed as an opportunity to support marketing claims; an exercise to convince the unschooled.

VALE COST-EFFECTIVENESS

A fundamental mistake (among many) in health technology assessment is to think that a 'single metric' can support claims for cost-effectiveness. A belief, unfortunately, shared by health system decision makers who are looking for a 'one size fits all' solution. Decision making is somewhat more complex with formulary committees required to consider a range of product attributes. While this might degenerate with checklists and weights into multiple criteria decision analysis (MCDA), it is not the purpose to pre-empt (as the ICER model does) the need for identifying the attributes of interest in therapy evaluation. This is a decision for the formulary committee. Imaginary claims from an ICER lifetime simulation are of little consequence; a view that is likely to be reinforced once competing ICER 'gold standard backbone' imaginary modelled claims are presented 'for approximate information' to the bemused members of committees all claiming the seal of approval from the ICER Analytics package

The question then becomes: why did ICER venture down this path? Not to put too fine a point on it, the choice seems suicidal. ICER is not the most admired agency with the QALY having even fewer friends, so why present manufacturers and others with the opportunity to discredit the ICER creation of imaginary claims through comparative simulations? Certainly ICER has been accused of being less than transparent in developing its models, but this seems an over-reaction. As promoted by ICER, the platfor4m is intended, it must be emphasized to support formulary development and adoption, internal assessment (?) and P & T committee prep, long-term value and budget impact modeling, and development of outcomes based agreements. This last point is intriguing: how can value assessment contracts be built on imaginary and mathematically nonsensical constructs if the intent of the contract is validation of empirical claims? It gets even more ambitious: ICER Analytics (hopefully) can be used to formulate pre-market research and pricing strategies (again with the lifetime impossible or I-QALY) while many others 'will find their own goals advanced … by patient groups seeking a full seat at the table …. to discuss pricing and access'; discussions presumably based on comparing imaginary pseudoscientific simulated clams without any comprehension that the 'backbone' ICER Analytics model is bankrupt. The proof, presumably, will be in the pudding: come back in 12 months to evaluate the and application of this platform.

A SIMPLE ALTERNATIVE

Hopefully, the ICER Analytics platform will bring home the absurdity of creating imaginary evidence to support nonsensical cost-outcomes claims. Fortunately there is an alternative that meets the requirements of normal science: version 3 of the Minnesota proposed formulary guidelines [9]. There are four key principles:

- **Claims must meet the standards of normal science**
  - **Claims must be credible, empirically evaluable and replicable**
  - **Claims must meet the standards for fundamental measurement**
  - **Claims must be dimensionally homogeneous or unidimensional**
- **Claims must be for single attributes defined for clinical outcomes, quality of life and resource utilization**
- **Claims must be specific to target populations within disease areas**
- **Claims must be accompanied by a protocol detailing how they might be evaluated or how they have been evaluated**

The implications of applying these standards is that generic, multiattribute claims and the subsequent construction of the mathematically impossible QALY are of no interest; nor is there any interest in broad 'cost-effectiveness' claims based upon imaginary incremental lifetime cost per QALY models. Thresholds are also of no interest along with value assessments of a 'fair' price. These are all mathematically impossible.

Ensuring that an instrument is designed to capture single attributes with measurement on either an interval scale or a ratio scale is critical. Unfortunately, the majority of disease specific PRO instruments fail these standards. They are dimensionally heterogeneous and lack construct validity. There are a handful of instruments, particularly in needs based quality of life that meet, for interval response assessment, the required Rasch Measurement Theory (RMT) standards [10].

Claims must be evaluated within a timeframe agreed with the formulary committee. Where possible an evidence base should be proposed to evaluate the claims and set the scene for ongoing disease area and therapeutic class reviews. This does not mean an evidence base for each formulary committee; it would be sufficient to report on a single evidence base for a number of committees (e.g., registry).

CONCLUSIONS

ICER presents and tries to sell a bankrupt analytical framework. The modeled evidence reports are a chimera; this has been demonstrated on multiple occasions Rather than trying to expand the market for imaginary information, ICER should either withdraw or attempt to meet the standards common in the physical sciences and the more advanced social sciences such as education, psychology and

education. Otherwise we face a slow and tenacious withdrawal from a technology assessment meme that should have been smothered at birth.

REFERENCES

[1] Piglucci M. Nonsense on Stilts: How to tell science from bunk. Chicago: University of Chicago Press, 2010

[2] Neumann PJ, Willke R, Garrison LP. A Health Economics Approach to US Value Assessment Frameworks – Introduction: An ISPOR Special Task Force Report. *Value Health*. 2018;21:119-123

[3] Langley P. The Great I-QALY Disaster. *Inov Pharm*. 2020;11(3): No. 7
https://pubs.lib.umn.edu/index.php/innovations/article/view/3359/2517

[4]Langley PC, McKenna SP. Measurement, modeling and QALYs [version 1; peer reviewed] *F1000Research* 2020, 9:1048 https://doi.org/10.12688/f1000research.25039.1

[5] Langley P. The Impossible QALY and the Denial of Fundamental Measurement: Rejecting the University of Washington Value Assessment of Targeted Immune Modulators (TIMS) in Ulcerative Colitis for the Institute for Clinical and Economic Review (ICER). *InovPharm*.2020;11(2): No 17
https://pubs.lib.umn.edu/index.php/innovations/article/view/3330/2533

[6] ICER. ICER Analytics Launched to Accelerate the Real World Application of ICER Reports, Helping Payers, the Lifde Science Industry and Health Care Policy Makers move from Insight to Action. Media Release. 1 December 2020.
https://icer-review.org/announcements/icer-analytics-launch

[7] Langley PC. Sunlit uplands: the genius of the NICE reference case. *Inov Pharm*. 2016;7(2): No.12

[8] Langley PC. Dreamtime: Version 5.0 of the Australian Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee (PBAC). *Inov Pharm.* 2017;8(1): No. 5

[9] Langley PC. Value Assessment, Real World Evidence and Fundamental Measurement: Version 3.0 of the Minnesota Formulary Submission Guidelines. *InovPharm*. 2020 11(4): No. 12
https://pubs.lib.umn.edu/index.php/innovations/article/view/3542/2613

[10] Bobd T, Fox C. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 3rd Ed.New York: Routledge, 2015