

More Unnecessary Imaginary Worlds – Part 3: Cystic Fibrosis and the Institute for Clinical and Economic Review’s Draft Evidence Report

Paul C Langley, PhD

Adjunct Professor, College of Pharmacy, University of Minnesota

Abstract

On 20 February 2020, the Institute for Clinical and Economic Review (ICER) released its draft evidence report to establish the value of innovative therapies in the treatment of cystic fibrosis. Following its usual practice, ICER contracted with an outside group to construct a value assessment framework, in this case a microsimulation model, to generate value claims. The primary outcomes for value claims were incremental cost-per-QALY simulations for four target cystic fibrosis populations. The value assessment, in common with the same model applied earlier by ICER in cystic fibrosis, recommended substantial price discounts based on arbitrary threshold cost-per-QALY values. Unfortunately, the entire exercise, as detailed in previous commentaries in INNOVATIONS in Pharmacy is essentially a waste of time. Not only is the reference case model presented by ICER only one of a multiverse of other models, all driven by a selective application of model structure and assumptions, but the fact that the utilities that are applied to hypothetical time spent in different disease stages to modeled QALYs and lifetime cost-per-QALY claims fail to meet fundamental measurement axioms: they are ordinal manifest scores. Applied to target cystic fibrosis target patient groups, the modeled claims are meaningless. From the manufacturer’s perspective, in this case Vertex Pharmaceuticals who have developed all the cystic fibrosis therapies ‘modeled’ by the ICER contractor, their response to ICER claims should be to reject them out of hand; the constructs are imaginary and the outcome claims nonsense.

Keywords: imaginary worlds, cystic fibrosis, ICER, pseudoscience, nonsense claims, nonsense recommendations

Introduction

The construction of assumption driven imaginary worlds to support incremental cost-per-QALY claims for pricing and access recommendations is the hallmark of the Institute for Clinical and Economic Review’s (ICER) business model. ICER has issued two evidence reports on cystic fibrosis. The first report, a final evidence report was released in 2018; the second report, a draft evidence report, on 20 February 2020^{1,2}. The focus of this commentary is on the second report. Even so, this analysis of the manifest failings of the second report apply equally to the first report: neither the economic model nor the recommendations for pricing and access should be taken seriously. Indeed, it is worth noting that following the release of the first evidence report Vertex Pharmaceuticals, whose drugs featured in ICER’s modeled ‘analysis’, Kalydeco (ivacaftor), Orkambi (lumacaftor/ivacaftor) and Symdeko (tezacaftor/ivacaftor and ivacaftor), issued a strong condemnation of the ICER methodology, ‘a flawed scientific methodology’, and a ‘sham’ evaluation process, where conclusions were agenda-driven and pre-ordained³. In the second report it once again it is only Vertex cystic fibrosis therapies that are considered. These include the therapies considered for modeling in the first report but with the addition of the Trikafta (elexacaftor/tezacaftor/ivacaftor) combination.

The ICER model for the 2020 draft evidence report was an updated version of their previous imaginary microsimulation model. The modeling objective was to estimate the hypothetical lifetime effectiveness and cost-effectiveness of CFTR modulator treatments plus best supportive care for cystic fibrosis patients in the specific imaginary model framework. The primary health outcome was quality adjusted life years (QALYs) utilizing EQ-5D-3L utilities. Other lifetime horizons outcomes were life years, equal value of life years gained and lifetime number of acute pulmonary exacerbations. There was no expectation that any of these fabricated outcomes could be evaluated empirically given the lifetime framework. These imaginary outcomes were discounted, along with costs, in the base case model at 3%.

The lifetime microsimulation model created an imaginary value assessment for four possible therapeutic options for four cystic fibrosis populations: (i) patients who are candidates for Kalydeco monotherapy; (ii) patients who are homozygous for the F508del mutation; (iii) patients who are heterozygous for the F508del mutation; and (iv) patients who are heterozygous with minimal function mutation. The primary model variable was FEV₁%, modeled as a continuous variable to capture the effect of the CFTR modulator drugs. Imaginary patients in the model are simulated to accumulate life years, QALYs and costs. EQ-5D-3L ordinal utility values were assigned to ppFEV_i status to generate QALYs per year and cumulate them over the lifetime of the patient.

The base case incremental cost-per-QALY ‘results’ (against best supportive care) for each of the patient models are, of course, imaginary and not for empirical assessment (i.e., hypothesis testing). We have to take them at face value; or reject them

Corresponding author: Paul C Langley, PhD
Adjunct Professor, College of Pharmacy
University of Minnesota, Minneapolis, MN
Director, Maimon Research LLC; Tucson, AZ
Email: langley@maimonresearch.com
Website: www.maimonresearch.net

accordingly. For each population simulated estimates are presented for total QALYs, total life years and equal value life years. Imaginary QALY gains are modest, ranging from 4 to 7 years. Incremental simulated lifetime costs are substantial ranging from \$5.5 million to \$7.0 million dollars, set against best supportive care. Consequent estimates of imaginary cost-per-QALY gained are, unsurprisingly, substantial. These range from \$818,000 for Trikafta plus best supportive care to \$1,060,000 for both Kalydeco plus best supportive care and Symdeko plus best supportive care. Inevitably, these lead ICER to conclude that while the therapies clearly have substantial clinical benefits (which we knew anyway), this can only be realized if there are substantial price discounts. According to the imaginary microsimulation, a discount of 35 – 51% would be necessary to reach a cost-effectiveness threshold of \$500,000/QALY for Kalydeco and Symdeko (as they classify as a rare disease target) while discounts of 65 – 66% would be necessary for Trikafta to reach a cost-effectiveness threshold of \$200,000/QALY. Again, caution is in order as these are entirely assumption driven imaginary claims; alternative models could lead to quite different imaginary conclusions.

Nonsense on Stilts

A recent commentary in INNOVATIONS in PHARMACY reviewed the latest ICER VAF to be applied over the period 2020 to 2023⁴. The commentary concluded that the VAF failed to meet the standards of normal science; it was considered pseudoscience. The principal reason for the ICER VAF failing the demarcation criteria between science and pseudoscience (or pure bunk) was its rejection of modeled claims that allow empirical evaluation. At the same time, the ICER VAF fails the standards for fundamental measurement⁵. It applies utilities which are manifest scores, they have ordinal rather than interval scale properties. This means that the consequent QALY and cost-per-QALY estimates are meaningless. The consequences are, for products in cystic fibrosis, that conclusions regarding estimated value and proposals for lifetime cost-per-QALY, with consequent recommendations for price discounting are unsupported. They are an unnecessary illusory distraction, irrespective of the degree of precision presented.

The purpose of this commentary is to build on the analyses and arguments presented in previous commentaries, notably the review of the ICER 2020-2023 VAF, to make the case for rejecting the draft evidence report on cystic fibrosis. The commentary starts with a brief restatement of the role of normal science as a process of discovering new facts; not recycling old assumptions. This is followed by a rejection of ICER's fabrication of an imaginary future sickle cell treating environment as simply one of any number that could be created with varying assumptions. After all, the accepted framework in health technology assessment which ICER accepts is to reject hypothesis testing in favor of 'approximate information' (whatever that means for an unknown future)⁶.

Given ICER's emphasis on QALYs and the fabrication of incremental cost-per-QALY claims, the next step is to point out the failure of ICER to grasp that if utility scores are to be applied to estimate time in disease states, then they have to meet the fundamental axioms of measurement theory: invariance of comparisons and sufficiency. They have to reflect an underlying construct of relevance to the target patient population with unidimensional and interval scoring properties. The EQ-5D utilities that ICER relies upon in fabricating claims in cystic fibrosis do not meet these standards. The QALYs are nonsense.

Finally, we point to the importance of constructing claims for treatment response within disease areas; not utilizing a generic health related quality of life construct (the EQ-5D) that captures a limited number of symptoms and ordinal responses. The case made here is for a latent needs-based construct, where the instrument reflects the needs of patients in cystic fibrosis and, if required, a separate instrument for the needs of caregivers. These should meet Rasch standards⁷. This is an unexceptional requirement that has been in place for 60 years in the application of Rasch Measurement Theory (RMT) to instrument development.

The Standards of Normal Science

The requirement for testable hypotheses in the evaluation and provisional acceptance of claims made for pharmaceutical products and devices is unexceptional. Since the 17th century, it has been accepted that if a research agenda is to advance, if there is to be an accretion of knowledge, there has to be a process of discovering new facts. By the 1660s, the scientific method, following the seminal contributions of Bacon, Galileo, Huygens and Boyle, had been clearly articulated by associations such as the Academia del Cimento in Florence (1657) and the Royal Society in England (founded 1660; Royal Charter 1662) with their respective mottos *Provando e Riprovando* (prove and again prove) and *nullius in verba* (take no man's word for it)⁸.

By the early 20th century, standards for empirical assessment were put on a sound methodological basis by Popper (Sir Karl Popper 1902-1994) in his advocacy of a process of 'conjecture and refutation'^{9,10}. Hypotheses or claims must be capable of falsification; indeed, they should be framed in such a way that makes falsification likely.

Although Popper's view on what demarcates science (e.g., natural selection) from pseudoscience (e.g., intelligent design) is now seen as an oversimplification involving more than just the criteria of falsification, the demarcation problem remains¹¹. Certainly, there are different ways of doing science but what all scientific inquiry has in common is the 'construction of empirically verifiable theories and hypotheses'. Empirical testability is the 'one major characteristic distinguishing science from pseudoscience'; theories must be tested against data. We can only justify our preference for a theory by continued evaluation and replication of claims. This applies in cystic fibrosis just as it does in other therapies. Constructing

imaginary worlds, even if the justification is that they are 'for information' is, to use Bentham's (Jeremy Bentham 1748-1832) memorable phrase 'nonsense on stilts'. If there is a belief, as subscribed to by ICER, in the sure and certain hope of constructing imaginary worlds, to drive formulary and pricing decisions, then it needs to be made clear that this is a belief that lacks scientific merit.

Assumptions

The ICER claim to fame is the ability to construct or fabricate an imaginary world that sets the stage for value impact over 10, 20 or 30 years in the future. In the cystic fibrosis model of the therapies offered by Vertex Pharmaceuticals, the number of assumptions made to support the microsimulations across the four patients groups is truly awesome; some come from the literature, others are pure guesswork. Unfortunately, even if an assumption driving the imaginary value assessment framework is defended by appealing to the literature (including pivotal clinical trials) the effort is wasted.

The point, and this goes back to Hume's (David Hume 1711 – 1776) induction problem, is that we cannot ask clients in health care to believe in models constructed on the belief that prior assumptions will hold into the future. It is logically indefensible: it cannot be *'established by logical argument, since from the fact that all past futures have resembled past pasts, it does not follow that all future futures will resemble future pasts'* ¹².

Utilities and QALYs

Quality adjusted life years (QALYs) can only survive if the measure is credible, evaluable and replicable. The QALY constructed by ICER in the cystic fibrosis model meets none of these criteria. The concept of a QALY is not new; it goes back some 40 plus years with the notion of combining time spent in a disease state with some multiplicative 'score' on a required interval scale of 0 to 1 (death to perfect health), Combining the two, multiplying time by utility is assumed to produce a QALY. In the ICER imaginary cystic fibrosis world these are combined to produce QALYs for the modeled life span.

Unfortunately, creating a QALY is a mathematically meaningless operation. An ordinal scale does not have the required fundamental measurement properties. The implications are obvious: any attempt to generate an incremental cost-per-QALY model makes no sense.

HRQoL versus QoL

It should be remembered that the EQ-5D is a health related quality of life (HRQoL) measures. It comprises 5 symptoms (select by an expert panel) and 3 response levels for each symptom (the EQ-5D-3L) or five response levels (the EQ-5D-5L). The latter is considered more responsive, but is for all intents and purposes a different system as it yields quite different 'utility' or manifest score profiles.

The fundamental weakness is that HRQoL symptom/response instruments have only a tenuous, if any, link to a latent construct. They, in fact, represent a mishmash of possible health constructs, where each symptom category might represent a latent construct. It is an operational measure, not one that bears even a limited relationship to a patient-centric measure of QoL.

Since the early 1990s, increasing attention has been given to QoL versus symptom and response characterizations of HRQoL in evaluating the benefit to patients of innovative therapies: the needs fulfilment model ¹³. Rather than imposing a series of potential symptoms and responses, the needs approach starts from a simple premise: the focus should be on QoL as a single latent construct; one which is disease specific (and specific to patients and caregivers) which is defined in terms of the needs of target groups. If health and disease status are the principal driver of 'good health' in QoL, then we have to identify those needs and devise an instrument that captures those needs in a single index with interval properties for target patient and caregiver groups. This is achieved by the application of Rasch measurement theory (RMT) in instrument development to give a single index with unidimensional properties; an index that meets standards for fundamental measurement ¹⁴.

HRQoL in Cystic Fibrosis

This does not mean that ICER is not aware of disease specific instruments in cystic fibrosis. Reference is made to probably the most widely used instrument, the Cystic Fibrosis Questionnaire (CFQ) ¹⁵. In the revised version (CFQ-R) there are three versions: adolescents and adults, children and parents. The instrument, adolescent and adult version, has 44 items on 12 generic and disease specific scales with a composite single score (0 – 100 scale).

Unfortunately, while the CTQ meets classical test theory criteria, it does not exhibit interval scaling properties and does not consider a broader QoL needs framework in instrument development. The scale only captures manifest scores. Once again, we find an instrument that is widely used but where those using it are not aware of its lack of fundamental measurement properties. This criticism applies across the board to the majority of patient recorded outcome (PRO) instruments. Unless there has been a review of Rasch standards with possible recalibration, including exclusion, of items, then a PRO instrument must be assumed not to meet RMT standards.

Creating Ordinal Utilities for Approximate Information

The starting point for the ICER imaginary QALY creation is a study by Schector et al which 'creates' utility scores for health states defined by FEV₁% ¹⁶. The basis for this exercise is a study by Bradley et al which reported for a UK cystic fibrosis population ¹⁷. This study produced three EQ-5D scores corresponding to: (i) 0.85 no current pulmonary events ; (ii) 0.79 mild (no hospitalizations) pulmonary exacerbation; and (iii) 0.60 severe (hospitalization) pulmonary exacerbations.

Apparently, these correlate well with the ordinal scores of the CTQ. Once again, these EQ-5D-3L 'scores' lack interval measurement properties; we can only rank them as the difference between the 'scores' has no meaning.

Nevertheless, in defiance of the axioms of fundamental measurement in a successor study, Tappenden et al translated the Bradley et al results and generated estimates of EQ-5D based on FEV₁%. FEV₁ > 70% EQ-5D = 0.864; FEV₁% 40 – 79% = EQ-5D 0.810; and FEV₁% < 40 = EQ-5D 0.641¹⁸. Despite the fact that these EQ-5D 'estimates' are still manifest scores, Schector et al went one step further and undertook a linear interpolation to 'predict' EQ-5D manifest scores for nine FEV₁% intervals. These scores were used to populate the ICER cystic fibrosis model and create imaginary QALYs.

According to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), the leaders in health technology assessment have laid down that modeled incremental cost-per-QALY claims are the gold standard, not for testing hypotheses but to fabricate 'approximate information'⁶. Apart from the obvious point that, if a formulary committee is faced with competing 'for information' modeled claims in cystic fibrosis it will have concerns over choosing one set of claims over another, it might also have an issue with the meaning of the term 'approximate information' defies common sense.

Should we be concerned that claims for innovative therapies in a disease such as cystic fibrosis should be the result of fabricated and mathematically indefensible QALYs and cost-per-QALY value assessments? After all, it is only for 'approximate information' not for testing hypotheses. However, as Tennant et al point out in their contribution to a special supplement in the ISPOR house journal *Value in Health* (2004): *As long as primitive counts and raw scores are routinely mistaken for measures by our colleagues in social, educational and health research, there is no hope of their professional activities ever developing into a reliable useful science*¹⁴.

Going Forth

Certainly, products in cystic fibrosis are considered expensive, but this has not stopped manufacturers such as Vertex from entering into market access programs in a number of countries

and health jurisdictions. The ICER model, with its bizarre cost-per-QALY claims is an unnecessary distraction. It short circuits rational, evidence based discussions in formulary decision making; discussions which could point to the role of needs-based patient centric instrumentation to assess claims for target cystic fibrosis groups, revised protocols for ongoing clinical and observational studies and, overall, a commitment to the discovery of new facts.

Science does not advance through the fabrication of imaginary worlds built on assumptions, real or imagined. This is echoed by Newton (Isaac Newton 1642-1727) with Descartes as his target (René Descartes 1596-1650) in saying '*hypotheses non fingo*' (I do not feign hypotheses). Descartes in Newton's view had 'produced fantastic and untestable ideas, then assumed them to be true and used them as building blocks of his philosophy'¹⁹

Of course, ICER is caught between the proverbial rock and a hard place. It has to use a generic utility measure (flawed as it is) to support the fabrication of imaginary QALYs; absent a generic utility measures (or at least a manifest score on a 0 – 1 metric), QALYs cannot be modeled and the exercise collapses. ICER has no option; it has to use the EQ-5D HRQoL instrument even though it may have no relation to the needs of patients and caregivers. ICER is then open to continuing criticism that its models for value assessment are just 'nonsense on stilts'. This applies to the cystic fibrosis reports as well as to its other evidence reports. The escape for ICER is to propose, after the imaginary migraine evidence model has been constructed and pricing and access recommendations presented, to suggest a move to *real* world evidence (from *imaginary* world evidence) and a possible disease specific research program. The first step seems redundant; but it is the ICER business model.

At launch, we may only have clinical data. This may be sufficient to propose cost-outcome claims, but these must be credible, evaluable and replicable. If these data points are not to hand then, rather than creating an imaginary modeled world, the focus should be on claims assessment and feedback to a formulary committee. Not, it must be emphasized, a retreat to a medieval world where the search for new facts is discouraged; subsumed in an acceptance of imaginary and unsupported claims for therapy impact and value.

References

- ¹ ICER. Modulator Treatments for Cystic Fibrosis: Effectiveness and Value. Final evidence report and meeting summary. 7 June 2018 https://icer-review.org/wp-content/uploads/2017/10/CF_Final_Evidence_Report_06082018.pdf
- ² ICER. Modulator Treatments for Cystic Fibrosis: Effectiveness and Value. Draft Evidence Report. 20 February 2020 https://icer-review.org/wp-content/uploads/2019/09/ICER_CF_Draft_Report_022020.pdf
- ³ <https://s3.amazonaws.com/assets.fiercemarkets.net/public/005-LifeSciences/Vertex+to+ICER+May+3+Response+FINAL.pdf>
- ⁴ Langley PC. Nonsense on Stilts – Part 1: The ICER 2020-2023 Value Assessment Framework for Constructing Imaginary Worlds. *InovPharm*. 2020;11(1):No. 12 <https://pubs.lib.umn.edu/index.php/innovations/article/view/2444>
- ⁵ Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: Time to end malpractice. *J Rehabil Med*. 2012;44:97-98
- ⁶ Neumann P, Willke R, Garrison L. A health economics approach to US value assessment frameworks – Introduction: An ISPOR Special Task Force Report (1). *Value Health*. 2018;21:119-25
- ⁷ Bond T, Fox C. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 3rd Ed. New York: Routledge, 2015
- ⁸ Wootton D. The Invention of Science: A new history of the scientific revolution. New York: Harper Collins, 2015.
- ⁹ Popper KR., The logic of scientific discovery. New York: Harper, 1959.
- ¹⁰ Lakatos I, Musgrave A (eds.). Criticism and the growth of knowledge. Cambridge: University Press, 1970.
- ¹¹ Piglucci M. Nonsense on Stilts: How to tell science from bunk. Chicago: University of Chicago Press, 2010)
- ¹² Magee B. Popper. London; Fontana, 1973
- ¹³ McKenna S, Doward L, Niero M et al. Development of needs-based quality of life instruments. *Value Health*. 2004;7(1 Suppl 1):S17-S21
- ¹⁴ Tennant A, McKenna S, Hagell P. Application of Rasch analysis in the development and application of quality of life. *Value Health*. 2004;7(1 Suppl 1):S22-S26
- ¹⁵ Quittner A, Buu A, Messer M et al. Development and validation of the Cystic Fibrosis Questionnaire in the United States: a health-related quality-of-life measure for cystic fibrosis. *Chest*. 2005;128(4):2347-54
- ¹⁶ Schechter M, Trueman D, Farquharson R et al. Inhaled aztreonam versus inhaled tobramycin in cystic fibrosis: An economic evaluation. *Ann Am Thoracic Soc*. 2015;12(7):1030-1038
- ¹⁷ Bradley J, Blume S, Balp M et al. Quality of life and healthcare utilisation in cystic fibrosis: a multicenter study. *Eur Respir J*. 2013;41:571-77
- ¹⁸ Tappenden P, Harnan S, Uttley L et al. The cost-effectiveness of dry powder antibiotics for treatment of *Pseudomonas aeruginosa* in patients with cystic fibrosis. *PharmacoEconomics*. 2014;32:159-72
- ¹⁹ Briggs R. The Scientific Revolution of the seventeenth century. London: Longman, 1971