INNOVATIONS in pharmacy

Volume 8 | Number 1 Article 22

3-9-2017

Intelligent Design and Imaginary Worlds in Cost-Effectiveness Claims: An Overview of Commentaries in INNOVATIONS in Pharmacy from July 2016 to February 2017

Paul C. Langley
University of Minnesota, langley@maimonresearch.com

Intelligent Design and Imaginary Worlds in Cost-Effectiveness Claims: An Overview of Commentaries in *INNOVATIONS in Pharmacy* from July 2016 to February 2017

Paul C Langley, PhD, Adjunct Professor, University of Minnesota

Abstract

Over the past 30 years, thousands of modeled cost-effectiveness studies have been published. Whether this effort has been worthwhile is debatable. For supporters of modeled cost-effectiveness studies, this effort has been worthwhile because the modeled claims are intended to inform health care decision makes. To opponents, the effort is seen as largely a waste of time because the claims are typically non-evaluable. Rather than subscribing to the standards of normal science in its focus on hypothesis testing through experimentation and replication, opponents have viewed practitioners in health technology assessment as being content to develop imaginary worlds with no thought to their role in practical decision-making. Given this ongoing commitment to constructing imaginary worlds, the purpose of this commentary is bring together the various commentaries, together with supporting publications that have appeared in INNOVATIONS in Pharmacy since mid-2016. This is in response to requests from faculty and graduate students at the College of Pharmacy, University of Minnesota for an overview of the arguments presented to date in support of a new research program that rejects imaginary constructs in favor of credible, evaluable and replicable claims to support formulary decision making.

Keywords: imaginary worlds, intelligent design, reference case, credibility, evaluability, replicability

Introduction

Over the past nine months, 22 commentaries published in *INNOVATIONS in Pharmacy* have pointed to the lack of credibility in modeled or simulated claims for the clinical benefit and cost-effectiveness of pharmaceutical products and devices. This lack of credibility stems from the construction of long-term or lifetime models of interventions in chronic disease states. The hallmark of such models is that the claims made are impossible to evaluate. The exemplars are the so-called 'reference case' models, which mandate the construction of lifetime cost-per-quality adjusted life year (QALY) claims for hypothetical treatment cohorts. Claims for competing products which capture a time horizon extending for twenty, thirty or more years, are impossible to evaluate; they are immune to failure. Models which fail to generate evaluable claims have been characterized in these commentaries as imaginary worlds [1] . The health care decision maker has no idea whether these claims are right or even if they are wrong, and, of course, they will never know.

Corresponding author: Paul C Langley, PhD
Adjunct Professor
College of Pharmacy, University of Minnesota
Director, Maimon Research LLC
5061 North Apache Hills Trail, Tucson, AZ 85750
Tel: 520-577-0436; Email: langley@maimonresearch.com
Web:www.maimonresearch.net

Clearly, decision makers must be aware that such modeled claims are impossible to assess. Even so, guidelines issued by number of agencies mandate or suggest manufacturers submit non-evaluable modeled or simulated imaginary cost-effectiveness claims. These guidelines have been reviewed in the commentaries and supporting articles:

- National Institute for Health and Care Excellence (NICE) in the UK[2]
- Health Quality and Information Authority (HQIA) in Ireland [3]
- Pharmaceutical Benefits Advisory Committee (PBAC) in Australia [4]
- Haute Authorite de Santé (HAS) in France [5]
- Dutch National Healthcare Institute (ZN) [6]
- Pharmaceutical Management Agency (PHARMAC) in New Zealand[7]
- Institute for Quality and Efficiency in Health Care (IQWiG) in Germany[8]
- European Network for Health Technology assessment (EUnetHTA) [9]
- Academy of Managed Care Pharmacy (AMCP) in the US[10]

Indeed, the Canadian Agency for Drugs and Technologies in Health (CADTH) has gone so far in its latest draft for the 4th edition of it guidelines released in October 2016 to make explicit its commitment to the construction of imaginary worlds and to put the scientific method to one side: 'Economic evaluations are designed to inform decisions. As such, they are

distinct from conventional research activities, which are designed to test hypotheses' [11].

This notion of 'informing decision makers' through the construction of imaginary claims for competing pharmaceutical products can, in fact, be traced back over the past 30 years with a number of publications supporting cost-effectiveness analyses, to include the seminal Drummond et al textbook on health technology assessment [12] [13] [14]. Unfortunately, the notion of 'informing' decision makers through constructing evidence to support cost-effectiveness claims opens the floodgates for competing imaginary worlds. Decision makers, even if they engage external reviewers to examine the 'reality' of competing models, will be faced with competing claims with each model justified in terms of its conformity to the model builders (or the commissioned model builders) perception of a future reality. The result is that over the past 30 years hundreds (if not thousands) of non-evaluable claims for product cost-effectiveness have been published. This is not, however, a blanket condemnation of health technology assessment, rather a rejection of those activities which have supported the construction of decision model claims which are non-evaluable. To separate the wheat from the chaff requires a commitment to science as an 'investigation of nature, based on the construction of empirically verifiable theories and hypotheses'.

This commitment to the construction of imaginary claims for competing pharmaceutical products looks set to continue. There is no evidence that health technology assessment agencies or professional groups such as the AMCP, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) or the EUnetHTA are prepared to reject imaginary non-evaluable cost-effectiveness claims, irrespective of whether these claims are expressed in cost-per-QALY terms or other measures of health impact [15] [16]. These groups seem prepared to continue to be 'informed' through the construction of imaginary worlds. It could of course be argued that decision makers are aware of the constructed nature of the cost-effectiveness claims that are presented yet put them to one side, preferring instead to focus on the evidence from randomized clinical trials (RCTs) as the critical input to formulary positioning and pricing. If so, then it is difficult not to conclude that the effort put into promoting the construction of imaginary worlds over the past 30 years has been largely wasted.

This commentary has been written in response to requests from faculty and graduate students in the Program in Social and Administrative Pharmacy at the University of Minnesota for an overview of the 22 commentaries published since July 2016. Other commentaries will follow. Given the material covered, it was considered to be useful to look back and consolidate the arguments made in support of the scientific method for cost-effectiveness claims. This focus is reflected in the use of the term 'intelligent design' in the title to this commentary. The intent here is twofold: on the one hand to emphasize the importance of distinguishing science from pseudoscience (or, as Pigliucci would describe it, bunk) and on the other as a caution that claims for 'realism' in model structure and assumptions that accompany the construction of long-term or lifetime models are no more 'intelligent' that those claims for intelligent design (a.k.a. creationism). Unlike Dawkins' 'blind watchmaker', the various of Paley's 'skilled watchmakers' are known; they not only publish but their activities are endorsed by professional groups such as the AMCP and ISPOR in their creation of unevaluable claims through modeled or simulated imaginary worlds[17] [18] [19]. The concern is that after 400 years of the acceptance of the scientific method, it is a moot point as to whether or not the many advocates of constructing imaginary worlds to inform decision makers recognize the disconnect between their advocacy of imaginary worlds and the standards of normal science.[20].

The Invention of Science

There is no lack of evidence for the claim that the late 16th century and the 17th century (from Brahe's nova of 1572 to Newton's *Opticks* of 1704) witnessed a profound transformation; a change that Wootton has described as the 'invention of science'. The new science of the 17th century, what we may now call the Scientific Revolution was characterized, again as described by Wootton, by a preoccupation with the systematic employment of the test of experience, observation and experimentation. For our present purpose, the key point is that the revolution of the 17th century presented a challenge to and eventual overthrow of Aristotelian authority. Experimentation replaced an appeal to authority as witnessed in the motto of the Royal Society (founded 1660; Royal Charter 1662): *nullius in verba* (take no man's world for it').[21]. As stated on the Royal Society website, this motto 'is an expression of the determination of Fellows to withstand the domination of authority and to verify all statements by an appeal to facts determined by experiment'[22].

Whether a scientific revolution actually occurred has been challenged by what has been described as the relativist school, in particular the strong program and the advocacy of equivalence in maintaining that the 'content' of science, not just the organization, values and activities of scientists, can be explained sociologically through the 'principle of symmetry'. Whether knowledge claims are sustained or not, relativists would maintain that it is illegitimate to prefer one explanation over another just because one is supported by evidence. Experimentation and hypothesis testing are not the only way of coming to grips with reality; a position presumably

subscribed to by CADTH. Evidence, in Wootton's summary of the relativist position, is never discovered 'it is constructed within a social community … the success of a scientific research program thus depends on not its ability to generate new knowledge but to mobilize the support of a community'. Science is about 'rhetoric, persuasion and authority'.

For our present purposes, the relativist position, in particular its subscription to constructed evidence is important in the advocacy and acceptance of models or simulations that reject the scientific method and support non-evaluable product claims as decision criteria. For those readers who remain unconvinced then a Wikipedia reference 'scientific method' is a useful starting point [23]. In the commentaries published to date this relativist position (it might even be described as postmodern) has been assessed, first, in the various commentaries that have reviewed the health technology assessment guidelines published in the US by the AMCP, the UK, New Zealand, Australia, The Netherlands, France, Germany and those proposed by EUnetHTA for the European Union [24] [25] [26] [27] [28] [29] [30] [31]. A separate review of the Irish guidelines was published in *Current Medical Research and Opinion* [32]. Reference should also be made to a series of papers on the status of modeled claims commissioned by the *Journal of Medical Economics* in late 2015 [33] [34] [35] [36] [37]. The purpose of these last commissioned papers was to argue that the only acceptable modeled claims for costs and outcomes are those that are testable and can be validated in a timeframe that is acceptable to a formulary committee. The four papers explored the methodological issues in validation, the UK experience with NICE, the questions a formulary committee should ask of modeled claims, and the role of Big Data in validating modeled claims.

The common theme in these commentaries is that manufacturers and others making formulary submission are asked to create imaginary worlds to support, typically in chronic diseases, competing clinical and cost-outcomes claims. In all of these cases, the exemplar is the NICE lifetime cost-per-QALY 'reference case'. The key requirements are a modeled or simulated claim that includes: (i) all direct health effects; (ii) a cost-utility-analysis with full incremental analysis; (iii) a time horizon long enough to reflect all important differences in costs or outcomes between the technologies being employed; and (iv) health effects to be expressed in QALYs with the EQ-5D as the preferred measure. In practice, this is a blueprint for constructing (outside of short-term acute interventions or possible end-of-life interventions) long-term or lifetime models with outcome claims that are not intended to be open to evaluation and replication.

Demarcation: Pseudoscience

There is an ongoing debate in the philosophy of science literature over the question of demarcation; how do we distinguish science from non-science, including between science and pseudoscience? For Popper, the issue was clear cut. The necessary and sufficient condition for demarcation is the ability to falsify claims [38] [39]. Unlike verification, 'statements or systems of statements, in order to be ranked as scientific, must be capable of conflicting with possible or conceivable observations'. This does not mean that single instances where a theory is falsified would lead to its rejection; scientific progress is not as neat as that. However, the point remains: theories must be judged empirically. If they fail the test of generating evaluable claims then they are either non-scientific or fall into the pseudoscience category. Unfortunately, as lifetime cost-per-QALY or life years saved models demonstrate, there is still a willingness to accept pseudoscience and, as Pigliucci argues, superstitious belief in general.

Unfortunately, despite assertions that the construction of claims through imaginary worlds can 'inform' decision makers, the fact that the models lack evaluable claims put them in the category of pseudoscience. As such, lifetime-cost-per-QALY models share the platform with intelligent design rather than with natural selection. While we might agree that there is a continuum between 'hard' and 'soft' sciences, the common feature is the construction of empirically verifiable theories and hypotheses'. Pigliucci describes this as a 'trinity' that underlies all scientific activities: a commitment to dealing with natural phenomena and processes; a commitment to dealing with coherently conceptual constructs in the form of theories or hypotheses; and a commitment to assessment through experimentation and systematic observation. These three elements are not present in the construction of imaginary worlds. Rather, models and simulations that are considered to be sufficient in their representation of reality that the claims made necessarily follow [40].

Relevance: Quality Adjusted Life Years

As well as pointing to the uncritical acceptance of the reference case as the standard for modeled or simulated claims, the commentaries have also pointed to the limitations, both methodological and practical, of focusing on QALYs as the exemplar outcome measure in its ability to capture both morbidity and mortality [41]. Despite the popularity of QALYs as the 'gold standard' outcome measure among academic audiences, professional groups and a number of single payer health care systems, there is no evidence that health care systems have put in place policies to collect systematically utility scores to support, if presented in evaluable terms, cost-per-QALY claims either through experimentation or observation.

The commentaries also point out that in the US there has been little effort put towards agreeing on the need for, let alone the

construction, of a reference standard for QALYs. First, there is no agreement on whether or not it is even possible to agree on a reference standard given the diversity and incompatibility of the various QALY measures; a situation that is no different to the plethora of patient reported outcomes instruments that characterize various disease states. Second, it is doubtful, even within a health care system whether agreement could be reached on the choice of QALY instrument and the preferences for the defined health states. Third, unless a QALY metric is established as a process or outcomes measure for quality assessment, there is little chance that any health care system would invest resources in capturing QALYs from a specific instrument on a regular basis. Fourth, the chance that a QALY quality metric would be mandated in the US is effectively zero. This is made abundantly clear in the *Affordable Care and Patient Protection Act (2010)* which requires that the Patient Centered Outcomes Research Group (PCORI) exclude discounted cost-per-QALY or similar discounting measures and threshold values for priority setting in health care [42].

Overall, the commentaries conclude that it is doubtful, that the great expectations for QALYs could ever be realized outside of reference case imaginary worlds, or the willingness of decision makers to suspend belief in the standards of normal science, and accept lifetime cost-per-QALY claims as decision criteria. Unless, therefore, a case can be made for short-term and evaluable QALY claims, there seems little scope for QALYS, and associated cost-per-QALY claims, as inputs to formulary decision making. Perhaps, as Pip says to Estella, it has been 'a vain hope and an idle pursuit'[43].

One exception to this rejection of QALYs in the US are the activities of the Institute for Clinical and Economic Review (ICER), a research organization. The ICER has adapted the NICE model to a US environment and has produced a succession of reports that have applied a reference case standard for lifetime (or long term) cost per QALY models to threshold willingness-to-pay value based pricing claims. Two of these reports have been reviewed in these commentaries pointing to their lack of scientific credibility in generating unevaluable simulated claims [44] [45] [46] [47]. The first of these reviews explored alternative modeled claims for Entresto, an angiotensin receptor neprilysin inhibitor, versus the standard of care with an ACE inhibitor in patients with chronic heart failure. Two models were compared: a lifetime cost-per-QALY model and a 3-year cost and budget impact model. The primary reason for this review was the puzzling feature that for a product which is over 120 time as expensive compared to the standard of care (Entresto $380 per month vs. ACE inhibitor $3 per month) the modeled claim can be made that the product is, in willingness to pay terms, cost-effective. The analysis illustrates that, perhaps not surprisingly, different models can generate quite different perspectives on the presumption of 'cost-effectiveness'. The commentary concluded that if modeled claims are to be useful for formulary decision making, then we need to eschew 'black box' models with non-evaluable claims in favor of those models that yield credible, evaluable and replicable claims that can support defensible product placement and pricing decisions. The second and more recent of these reviews is the ICER January 2017 report on disease modifying therapies (DMTs) in multiple sclerosis. The review concluded that the claims presented for the competing DMTs are not credible, evaluable or replicable.

The objection to lifetime cost-per-QALY models applies also the models that have put QALYs to one side in favor of measures that have focused on more 'realistic' endpoints such as life years saved or other long-term disease specific clinical criteria. Unless these claims are evaluable they face the same objections raised in respect of QALYs: the claims are not credible and should be rejected. This is pointed to on the review of the German IQWiG guidelines where an efficiency frontier that attempts to capture lifetime costs and benefits has no more scientific merit than the application of the NICE reference case [30].

Absence: Compliance and Pricing Strategies

In these commentaries the question has been raised as to why, in the construction of imaginary worlds, simulations have typically failed to take account of (i) patterns of adherence and persistence to medications and (ii) a commitment by pharmaceutical manufacturers to a policy of sustained price increases. Irrespective of the lack of scientific merit in generating unevaluable claims, both issues undermine the lifetime perspective.

Evidence points overwhelmingly to the lack of adherence and persistence with medications. Within as short a period as two years there are a number of disease areas where less than one-third of patients continue to be fully adherent to their medications or have discontinued therapy altogether. If this is the case then modeled claims that do not accommodate this perspective, or which understate the lack of long-term compliance, are of little interest to decision makers. This, unfortunately, opens Pandora's box: would manufacturers support modeled claims and the underwriting of a protocol to assess those claims when the evidence would suggest that for the majority of patients the perceived clinical benefits are only short term or transitory in nature? From the manufacturer's perspective, of course, it is not compliance behavior *per se* that is of interest but maintaining and increasing market share primarily by encouraging patients and physicians to initiate therapy. Manufacturer supported policies to improve compliance behavior are, it appears, of secondary interest. As long as new entrants at least balance out those discontinuing

therapy, market share is maintained. Any question of continuing benefits to patients is put to one side.

Policies by manufacturers for sustained annual or semi-annual price increases, with the percentage increase often in the double digits, also casts doubt on lifetime models. These models, while discounting lifetime costs and outcomes, take no account of long-term price increases and their contribution both to annual drug costs and other direct medical costs. Failure to consider the implications of policies by manufacturers to regularly increase the prices of pharmaceuticals on a semi-annual or annual basis means that claims for cost-effectiveness based on current prices are out-of-date within a few months. Indeed, if the strategy is to publish cost-effectiveness lifetime non-evaluable claims, then the claims are out of date even before the commissioned 'marketing' paper is published. This issue has been raised in the commentaries, again with specific reference to the two ICER evaluations [44] [45]. In the case of multiple sclerosis, reference was made to a recent study by Hartnung et al, of the trend in annual drug costs for nine disease modifying therapies (DMTs) from 1993 to 2014 [48]. Apart from the fact that DMT costs are two to three time bigger in the US than other countries, the principal finding is that DMT costs have accelerated well beyond inflation and substantially above rates for drugs observed in a similar biologic class. Annualized change in the cost of the nine DMTs in the evaluation ranged from 35.7% for glatiramer acetate to 7.9% for fingolimod. Four of the DMTs had annualized cost increases above 20% and four with annualized price increases between 13.0% and 16.8%. Natalizumab, for example, although being withdrawn briefly from the market between February 2005 and June 2006, increased in cost from $25,850 in 2004 to $64,233 in 2013 or an annualized increase of 16.2%. In terms of the overall 'costs' of care for commercially insured MS patients, a comparison of charges between 2006 to 2011 pointed to the continuing significant impact on total costs of the charges associated with drug costs (52.6% in 2011) [49] . At the same time, the increase for DMTs far outstripped the charges for other medical services (95.7% vs. 32.4%).

Progress: Replication of Imaginary Claims

If competing claims for pharmaceutical products are non-evaluable then it puts to one side one of the critical features of progress in science: the ability not only to empirically assess claims but to replicate those claims across, in this case, target patient populations. This is true both of clinical claims, in particular comparative clinical claims based upon indirect comparisons, as well as cost-outcomes and budget impact claims.

In health technology assessment the evidence base is, all too often, somewhat scanty. Guidelines for formulary submission typically put the question of replication of clinical claims to one side even though there is considerable disquiet over the inability to replicate phase 2 and phase 3 RCT claims [50] [51] [52]. This omission is compounded when non-replicable claims are embedded in models or simulations to support cost-outcomes claims, extrapolating over the long-term or the lifetime of an imaginary treatment cohort. The only way to overcome this question of 'belief' in comparative clinical claims is, as suggested here, to put claims in evaluable terms.

Publications: Imaginary Worlds

In commentaries published to date, the willingness of editors of three of the leading health technology assessment periodicals to publish non-evaluable claims, typically lifetime cost-per-QALY models, has been evaluated. The three journals are *Value in Health*, *Pharmacoeconomics* and the *Journal of Medical Economics* [53] [54] [55]. In each case a systematic review was undertaken of articles publish in the period January 2015 to December 2015. Each of the studies was judged against four criteria: (i) Is the model capable of generating evaluable claims? (ii) did the author(s) attempt to generate evaluable claims? (iii) did the author(s) suggested how the claims might be evaluated? and (iv) did the author(s) caution readers as to the implications of generating non-evaluable clams for the credibility of the analysis? In the case of *Value in Health,* 16 papers were identified. Of these 14 presented a cost-per-QALY analysis, with 9 presenting their claims for comparative effectiveness in a lifetime cost-per-QALY framework. With the focus of the assessment on whether or not these studies generated testable claims, none of the studies met this standard. The result for *Pharmacoeconomics* was the same. A total of 31 studies were evaluated, including 14 research articles, 8 systematic reviews and 9 reviews. Although the majority of the studies met recommended standards for cost-effectiveness analysis, none met the standards of normal science. They were best categorized as imaginary worlds or thought experiments. The same conclusion was reached for the *Journal of Medical Economics*. A total of 32 studies were identified. None of the studies presented their claims or projections in an evaluable form and none suggested how they might be evaluated. None met the standards of normal science. The claims made for cost-effectiveness were either impossible to verify, or if potentially verifiable, were not presented in an evaluable form. Indeed, in the case of the *Journal of Medical Economics* of the 32 studies, 30 were apparently funded by pharmaceutical manufacturers and 29 of the constructed simulations favored the manufacturer's product.

It is proposed to continue these evaluations of published studies and to extend the analysis to other journals as part of this ongoing commentary series. A major focus will be the extent to which unevaluable modeled comparative claims for

pharmaceutical products should be seen as contributions to the marketing and formulary acceptance of products when the study is underwritten by the manufacturer.

As well as these reviews of leading journals, a separate assessment was also undertaken of cost-effectiveness studies and their assumptions in diabetes mellitus and atrial fibrillation [56] [57]. In the first commentary reviewing diabetes mellitus models claims presented are all too often either immune to failure or are presented in a form that is non-testable. As such they fail to meet the key experimental requirements of falsification and replication. The second commentary the focus was on whether not the modeled claims for new oral anticoagulants (NOACs) met the standards of normal science in supporting falsification and replication. A systematic and consensus review by the authors identified 23 cost-utility NOACs evaluations along with four single technology appraisals undertaken by the National Institute for Health and Care Excellence (NICE) in the UK. None of the studies presented projections or claims in a form suitable for empirical evaluation. None could support falsification or replication. They failed the standards associated with the scientific method.

Progress: Minnesota Proposed Guidelines[58]

The primary purpose of the Minnesota guidelines, with the first version published in July 2016 and a revised second version published on the College website in December 2016, and detailed in accompanying commentaries, was to revisit an earlier guideline for formulary submissions developed in 2005 for the WellPoint (now Anthem) health system [59] [60] [61] [62]. Under the guidelines manufacturers submitting proposals for either new products or for products being revaluated as part of disease area and therapeutic class reviews, were required to present evaluable claims together with a study protocol detailing how the claims were to be evaluated and reported within a meaningful time frame. The proposed timeframe was two years. Until the claims were reported back, any pricing and formulary position would be provisional and in line with the product that was most likely to be replaced in practice. Non-evaluable modeled or simulated claims were to be rejected.

The revised December 2016 version 2.0 of the Minnesota proposed guidelines took into account the genetic targeting of therapies under next generation sequencing (NGS). It was considered that, with the dearth of clinical trials to support targeting therapies for products, which were likely to be off-label, the requirement for protocols to support evaluable claims was important. The question of NGS and its implications for formulary evaluations was explored in two commentaries [63] [64]. The first of these commentaries considered (i) the evidentiary standards for the evaluation of an NGS test and comparator tests and (ii) identified questions that a formulary committee should address in submissions made for a test in health care systems. A critical issue was not only comparative claims for the test against the standard of care and comparator tests, but the assessment of test performance for the identified treatment pathways where mutations or variants are linked to recommendations for therapy options. The second commentary considered how NGS may disrupt not only the accepted process of drug development but also the hurdles a drug manufacturer would be expected to face in securing formulary approval and a possible premium price for new compounds. This was characterized as a process of creative destruction, where adoption of NGS in personalized medicine sets in train a process of ongoing product review. A mechanism driven by continuing modifications and extensions to NGS platforms as our understanding of the role of mutations and mutation load in therapy choice expands.

In proposing protocols for formulary claims evaluations, the Minnesota guidelines present a checklist for assessing whether or not the claims presented are potentially credible, evaluable and replicable. There are a number of guidelines proposed in the literature, such as the ISPOR sponsored CHEERS guidelines, which do not meet these requirements [65]. In a recent commentary the Assessment of the Validation Status of Health-Economic Decision Models (AdViSHE) checklist for decision models was evaluated [66]. This new checklist which is intended to address the perceived need for a tradeoff between confidence in a decision model and the need to allocate resources by developers and payers to validating the model, does not address the issue of credibility, evaluation and replication of claims. The checklist fails the standards of normal science. Apart from the absence of a commitment in the AdViSHE checklist to the modeling of claims that are evaluable and replicable, the validation check list makes no allowance for a product pricing strategy that may commit a manufacturer to regular and substantial annual or semi-annual product price increases. Indeed, product-pricing assumptions are conspicuous by their absence.

Next Steps: Towards a Credible Research Program

As well as presenting a critique of current standards and processes for formulary evaluation, the various commentaries also considered the key elements in developing a research program in formulary submission evaluations that met the standards of normal science. The key here was the development of credible, evaluable and replicable claims to support pricing and formulary placement decisions. After over 30 years, however, it is unlikely that the current commitment by academic groups, professional associations, consultants and manufacturers to the construction of imaginary worlds will be readily abandoned. These groups have a considerable investment in constructing lifetime cost-per-QALY, viewing these as an essential input to formulary decisions; or, at least,

a worthwhile marketing tool to demonstrate the superiority of a sponsor's product. Many will be affronted by the label 'pseudoscience', even if they recognize that the claims made are non-evaluable. Many will, no doubt, continue to argue that while constructed claims are not consistent with the standards of normal science, these imaginary projections have a role to play in 'informing' decision makers. Presumably, in adopting a relativist (indeed a postmodernist) position means accepting that a lifetime cost-per-QALY threshold claims is to be viewed as just as valid as a daily astrological forecast or an evaluable short-term cost-outcome claim in informing decision makers.

For relativists, as Wotton points out, there is no concept of progress in science [20]. Relativists view science as *entirely* a social construction. Knowledge is whatever we choose it to be. There is no basis for demarcating good from bad science; good 'information' from 'bad' information'. Evidence to support competing theories is irrelevant. Reality does not constrain belief [20]. A historical case for the 'invention of science' in the late 16th and the 17th centuries tracing out the growing acceptance, the path-dependence, of the role of experimental programs of testing, refining and rejecting theories is beside the point. As Wootton emphasizes, relativists (or those accepting cognitive egalitarianism) *cannot* think historically [20]. This does not mean, of course, that one would necessarily embrace a strictly (and naïve) realist position. Even so, there remains the key question of demarcation and the role of experimentation. It is experimentation that divides science from pseudoscience and science from non-science. For experimentation and systematic observation to play their role, we require claims that are empirically evaluable.

Unfortunately, as much as interested parties may mount a rearguard action in support of non-evaluable claims for clinical and cost-effectiveness outcomes, the fact remains that modeled or simulated claims can fail. The problem in health technology assessment is that the modeled or simulated claims are typically immune to failure. If this argument is accepted, then there appears no option but to abandon constructed, non-evaluable claims in favor of a research program that meets the standards of normal science. After 30 years, when many have devoted a considerable part of their academic life to developing standards, promoting and publishing modeled or simulated non-evaluable claims, many of increasing mathematical complexity, their lack of relevance for decision making may come as a blow. Unfortunately, the last 30 years have witnessed the willingness of practitioners, supported by professional groups such as the AMCP and ISPOR, together with many academic research centers, to put the standards of normal science to one side in encouraging the construction of imaginary worlds.

Conclusions

Unless claims are evaluable and replicable in target patient populations, we lose any notion of scientific progress in our ability to develop models or simulations to support a greater understanding of the anticipated impact of new products and claims for comparative cost-effectiveness. This sets the stage for a progressive research program. The process of conjecture and refutation is central to the scientific method. Our understanding proceeds through the continuous process of empirically evaluating claims. This process may lead to a modification, enhancement or even a rejection of the model. This feedback is an essential element and one, unfortunately, that is entirely absent if modeled or simulated claims are not evaluable. In this situation, the idea of progress is absent. Presumably, if we continue to subscribe to the current paradigm of modeled yet unevaluable claims, we can look forward over the next 25 years to an ever growing body of cost-per-QALY lifetime disease specific models that the authors believe 'inform' decision makers, yet which decision makers may quite reasonably reject in the absence of any predictive content.

Some readers may find it odd that it is necessary to remind analysts in health technology assessment of the need to develop claims for the cost-effectiveness of pharmaceutical products and devices that meet the standards of normal science. Irrespective of the widespread and uncritical acceptance of constructed lifetime models in health technology assessment by the community of professional groups in health technology assessment as a viable and informative research program, this position must be abandoned. Embarrassment aside, a commitment to developing evaluable cost-outcomes claims is not asking for a radical departure from standards that apply in other areas of health technology assessment. There is, to give one example, an established literature evaluating the determinants of adherence and persistence behavior. Another example would be the use of retrospective data to evaluate comparative clinical claims and patterns of resource utilization. There are a growing number of accessible databases capturing not only administrative claims but integrating these with electronic medical records as well as web-based software to monitor and track patient outcomes through linking on-line patient reported outcomes with clinical profiles. These data provide ample opportunities to evaluate claims. As evidence for this, we can point to the nascent commitments to comparative effectiveness analysis. Although it is, doubtful if this program will meet its objectives given the issue of funding and the understandable reluctance, or at least wariness, of manufacturers to support such a research program. Even so, there is no excuse for not moving to evaluable claims.

Mainstream analysts in cost-effectiveness analysis are asking health decision makers to put the clock back, to take 'their' word for it. Over the last 30 years, we have witnessed the willingness of practitioners in health technology assessment, supported by the appeal by professional groups such as the AMCP and ISPOR, together with academic 'modeling' centers, to encourage and provide training in the construction of imaginary worlds. An appeal that mirrors the appeal to Aristotelian authority, which after some 1,500 years, took a century to overturn. The commentaries published over the past 10 months in *INNOVATIONS in pharmacy* are an attempt to overturn the existing technology assessment paradigm; to reject relativist reference case standards for non-evaluable claims in favor of a progressive research program. The hallmark of this program is to be to one side non-evaluable claims in clinical and cost-outcomes claims, focusing instead on models or simulations, which can generate feedback to health care decision makers. Whether it is sufficient to overturn the present technology assessment paradigm is an open question.

References

[1] Langley PC. Supporting Formulary Decisions: The Discovery of New Facts or Constructed Evidence?. *Inov Pharm.* 2016;7(2): No. 15. http://pubs.lib.umn.edu/innovations/vol7/iss2/15

[2] United Kingdom. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. London: Nice, April 2013

[3]Ireland. Health Information and Quality Authority. Guidelines for the Economic Evaluation of Health Technologies in Ireland. Dublin: HIQA; 2014

[4] Australia. Department of Health. Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (PBAC). Version 5.0. Canberra September 2016

[5] Haute Authorité de Sante. Choices for Methods for Economic Evaluation. Paris, 2012

[6]Zorginstituut Nederland. Guideline for Economic Evaluation in Healthcare. Diemen: 2016

[7] New Zealand. Pharmaceutical Management Agency. Prescription for Pharmacoeconomic Analysis: Methods for cost-utility analysis (v 2.1). PHARMAC, 2012

[8] Institut für Qualität und Wirtschaftlichkeit im Gesundheitseesen (Institute for Quality and Efficiency in Health Care). General Methods. Version 4.2. IQWiG: April 2015

[9]European Network for Health Technology Assessment (EUnetHTA). Methods for health economic evaluations. May 2015

[10] Academy of Managed care Pharmacy (AMCP). Format for Formulary Submissions, Version 4.0, April 2016

[11] Canada. Canadian Agency for Drugs and Technologies in Health. Guidelines for the Economic Evaluation of Health Technologies: Canada. Methods and Guidelines. Draft Report. Ottawa: CADTH, October 2016

[12] Weinstein MC, Fineberg HV. Clinical Decision Analysis. Philadelphia: Saunders, 1980

[13] Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in Health and Medicine. New York: Oxford University Press, 1996

[14] Drummond MF, Sculpher MJ, Torrance GW et al. Methods for the Economic Evaluation of Health Care Programmes. 3rd Ed. Oxford: Oxfor University Press, 2005

[15] Caro JJ, Briggs A, Siebert U et al. Modeling good research practices – Overview: A report of the ISPOR-SMDM Modeling good practices task force – 1. *Value Health*. 2012 (15:796-803

[16] Eddy DM, Hollingworth W, Caro JJ et al. Model transparency and validation: A report of the ISPOR-SMDM modeling good research practices Task Force – 7. *Value Health*. 2012;15:843-850

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

[18] Paley W. Natural Theology or Evidence of the Existence and Attributes of the Deity Collected from the Appearances of Nature. Oxford: J Vincent, 1826

[19] Dawkins R. The Blind Watchmaker. Oxford: Oxford University Press, 1986

[20] Wootton D. The Invention of Science: A New History of the Scientific Revolution. New York: Harper Collins, 2015

[21] https://royalsociety.org/about-us/history/

[22] https://royalsociety.org/

[23] https://en.wikipedia.org/wiki/Scientific_method

[24] Langley PC. Modeling Imaginary Worlds: Version 4 of the AMCP Format for Formulary Submissions. *Inov Pharm.* 2016;7(2): No. 11. http://pubs.lib.umn.edu/innovations/vol7/iss2/11

[25] Langley PC. Sunlit Uplands: The Genius of the NICE Reference Case. *Inov Pharm.* 2016;7(2): No. 12. http://pubs.lib.umn.edu/innovations/vol7/iss2/12

[26]Langley PC. He Ao Pohewa: The PHARMAC Prescription for Pharmacoeconomic Analysis in New Zealand and the Standards of Normal Science. *Inov Pharm.* 2016;7(2): Mo. 13. http://pubs.lib.umn.edu/innovations/vol7/iss2/13

[27] 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. http://pubs.lib.umn.edu/innovations/vol8/iss1/5

[28] Langley PC. Denkbeeldige Wereld: The New Dutch Guidelines for Economic Evaluations in Healthcare. *Inov Pharm.* 2017;8(1): No. 4. http://pubs.lib.umn.edu/innovations/vol8/iss1/4

[29] Langley PC. HAS Should Not Be NICE: Rejecting Imaginary Worlds in the French Technology Assessment Guidelines. *Inov Pharm.* 2017;8(1): No. 8. http://pubs.lib.umn.edu/innovations/vol8/iss1/8

[30] Langley PC. Imaginary Worlds and Efficiency Frontiers: Should We Abandon the IQWiG Health Technology Assessment Model?. *Inov Pharm.* 2017;8(1): No. 13. http://pubs.lib.umn.edu/innovations/vol8/iss1/13

[31] Langley PC. Imaginary Worlds: The European Network for Health Technology Assessment (EUnetHTA) Recommendations for Health Economic Evaluations. *Inov Pharm.* 2016;7(3): No. 3. http://pubs.lib.umn.edu/innovations/vol7/iss3/3

[32] Langley PC. Na domhain shamhlaíochta: formulary submission guidelines in Ireland and the standards of normal science. *Curr Med Res Opin*. 2016;32(5): DOI:10.1080/03007995.2016.1190699

[33] Langley P. The status of modeled claims. *J Med Econ.* 2015;18(12):991-992

[34]Langley P. Validation of modeled pharmacoeconomic claims in formulary submissions. *J Med Econ.* 2015;18(12):993-999

[35] Schommer J, Carlson A, Rhee G. Validating pharmaceutical product claims: questions a formulary committee should ask. *J Med Econ.* 2015;18(12):1000-1006

[36] Belsey J. Predictive validation of modeled health technology assessment claims: lessons from NICE. *J Med Econ.* 2015;18(12):1007-1012

[37] Wasser T, Haynes K, Barron J, Cziraky M. Using 'big data' to validate claims made in the pharmaceutical approval process. *J Med Econ.* 2015;18(12):1013-1019

[38] Popper KR., *The logic of scientific discovery*. New York: Harper, 1959.

[39] Popper KR. *Objective knowledge* (Rev. Ed.). Oxford: Clarendon Press, 1979.

[40] Grim P, Rosenberger R, Eason RE et al. How simulations fail. *Synthese*. 2013;190(12):2367-2390

[41] Langley PC. Great Expectations: Cost-Utility Models as Decision Criteria. *Inov Pharm.* 2016;7(2): No. 14. http://pubs.lib.umn.edu/innovations/vol7/iss2/14

[42] Persad G. Priority setting, cost-effectiveness and The Affordable Care Act. *Am J Law Med*. 2015;41:119-166

[43] Dickens C. Great Expectations. London, 1861

[44] Institute for Clinical and Economic Review. CardioMEMS™ HF System (St. Jude Medical, Inc.) and Sacubitril/Valsartan (Entresto™, Novartis AG) for Management of Congestive Heart Failure: Effectiveness, Value, and Value-Based Price Benchmarks. ICER: December 2015

[45] Institute for Clinical and Economic Review. Disease-modifying therapies for relapsing-remitting and primary-progressive multiple sclerosis: Effectiveness and Value. Evidence Report. ICER: January 26, 2017

[46] Langley PC. Cost-Effectiveness and Formulary Evaluation: Imaginary Worlds and Entresto Claims in Heart Failure. *Inov Pharm.* 2016;7(3): No. 6. http://pubs.lib.umn.edu/innovations/vol7/iss3/6

[47] Langley PC. Multiple Sclerosis and the Comparative Value Disease Modifying Therapy Report of the Institute for Clinical and Economic Review (ICER). *Inov Pharm.* 2017;8(1): Article 12. http://pubs.lib.umn.edu/innovations/vol8/iss1/12

[48] Hartung D, Bourdette D, Ahmed S et al. The cost of multiple sclerosis drugs in the US and the pharmaceutical industry. Too big to fail? *Neurology*. 2015:84:2185-2192

[49] Carroll C, Fairman K, Lage MN. Updated cost-of-care estimates for commercially insured patients with multiple sclerosis: retrospective observational analysis of medical and pharmacy claims data. *BMS Health Services Res*. 2014;14:286

[50] Editorial. *Nature*. 2016;529:256.

[51] Camerer CF, Dreber A, Forsell E et al. Evaluating replicability of laboratory experiments in economics. *Science*. 10.1126/science.aaf0918 (2016)

[52] Baker M. Is there a reproducibility crisis? *Nature*. 2016;533:452

[53] Langley PC. Imaginary Worlds: Modeled claims for cost-effectiveness published in PharmacoEconomics January 2015 to December 2015. *Inov Pharm.* 2016;7(2): No. 9. http://pubs.lib.umn.edu/innovations/vol7/iss2/9

[54] Langley PC, Rhee T. Imaginary Worlds: A Systematic Review of the Status of Modeled Cost-Effectiveness Claims Published in the Journal of Medical Economics from January 2015 to December 2015. *Inov Pharm.* 2016;7(2): No. 16. http://pubs.lib.umn.edu/innovations/vol7/iss2/16

[55] Langley PC, Rhee T. Imaginary Worlds: The Status of Modeled Economic Evaluation Claims Published in Value in Health January 2015 to December 2015. *Inov Pharm.* 2016;7(2): No. 18. http://pubs.lib.umn.edu/innovations/vol7/iss2/18

[56]Langley PC, Rhee T. Imaginary Worlds: The Status of Modeled Quality Adjusted Life Year Claims for New Oral Anticoagulants in Atrial Fibrillation Published Between January 2012 and February 2016. *Inov Pharm.* 2016;7(3): No. 7. http://pubs.lib.umn.edu/innovations/vol7/iss3/17

[57]Langley PC, Rhee TG. Imaginary Worlds: The Status of Simulation Modeling in Claims for Cost-Effectiveness In Diabetes Mellitus. *Inov Pharm.* 2016;7(2): No. 17. http://pubs.lib.umn.edu/innovations/vol7/iss2/17

[58] University of Minnesota, College of Pharmacy. Program in Social and Administrative Pharmacy. Guidelines for Formulary Evaluation: Proposed. Version 2.0, December 2016 https://www.pharmacy.umn.edu/sites/pharmacy.umn.edu/files/minnesota_guidelines_december_2016.pdf

[59] Langley PC. Nullius in verba: The University of Minnesota Social and Administrative Pharmacy Program proposed Guidelines for Formulary Evaluations. *Inov Pharm.* 2016;7(2): No 10 http://pubs.lib.umn.edu/innovations/vol7/iss2/10

[60] Langley PC. Nullius in verba: Version 2.0 of the University of Minnesota, School of Social and Administrative Pharmacy Program, Proposed Guidelines for Formulary Evaluation. *Inov Pharm.* 2016;7(4): No. 16 http://pubs.lib.umn.edu/innovations/vol7/iss4/16

[61] Sweet B, Tadlock CG, Waugh W et al. The Wellpoint outcomes based formulary: enhancing the health technology assessment process. *J Med Econ*. 2005;8:13-25

[62] Langley PC. Recent developments in the health technology assessment process in Fulda TR, Wertheimer AI (eds). Handbook of Pharmaceutical Public Policy. New York: Pharmaceutical Products Press. 2007.

[63] Langley PC. Next Generation Sequencing Claims in Precision Medicine: Questions a Formulary Committee Should Ask. *Inov Pharm.* 2016;7(4): No. 11 http://pubs.lib.umn.edu/innovations/vol7/iss4/11

[64] Langley PC. Creative Destruction: Next Generation Sequencing in Drug Development, Formulary Evaluations and Pricing. *Inov Pharm.* 2016;7(4): No. 13 http://pubs.lib.umn.edu/innovations/vol7/iss4/13

[65] Husereau D, Drummond M, Petrou S et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) – explanation and elaboration: A report of the ISPOR Health economic Evaluation Guidelines Good reporting Practices Task force. *Value Health*. 2013;16:231-50

[66] Langley PC. Validating Imaginary Worlds? The AdViSHE Assessment Tool. *Inov Pharm.* 2017;8(1): No. 3. http://pubs.lib.umn.edu/innovations/vol8/iss1/3