MAIMON WORKING PAPER No. 13 JULY 2024

REVISITING THE COMMITMENT TO BULLSHIT: ASSUMPTION DRIVEN SIMULATED REFERENCE CASE CLAIMS IN HEALTH TECHNOLOGY ASSESSMENT

Paul C. Langley, Ph.D., Adjunct Professor, College of Pharmacy, University of Minnesota, Minneapolis, MN and School of Pharmacy, University of Wyoming, Laramie, WY

ABSTRACT

The long-standing commitment in health technology assessment (HTA) and cost-effectiveness analysis (CEA) to the construction of assumption driven models to create non-evaluable claims for cost-effectiveness raises an intriguing question: is this belief system science, pseudoscience or bullshit. With non-falsifiable claims the presumption must be that the models fail the standards for demarcation between science and non-science. This leaves a revised question: are these models a variety of pseudoscience or just plain bullshit. This is an important distinction. Science is typically viewed, in its normative sense, as a systematic and critical investigation to acquire the best possible understanding of the workings of nature, people and society. Put simply, and this applies to the social sciences as well, normal science provides the framework for determining which beliefs are epistemically warranted through provisional fact-finding practices. This leads to the criteria for demarcation: falsification, which, in Popper's view, is the necessary and sufficient condition for distinguishing science from non-science or pseudoscience. Accepting Popper's criteria would relegate assumption driven simulations to a niche category in pseudoscience; but is this sufficient? Is there the possibility that we overlook the category of bullshit? The ability to distinguish HTA-CEA as pseudoscience rather than bullshit or vice versa raises a number of questions. These follow from the OED definition of pseudoscience as a pretended or spurious science involving a collection of beliefs that are mistakenly regarded as being based on scientific method with the status that scientific truths have. This would characterize reference models and cost-effectiveness analysis (CEA). It also characterizes the concept of preference scores and the QALY; which as far as the US is concerned have to be abandoned with the prohibition amendments to Section 504 of the Rehabilitation Act of 1973. Care has to be taken in considering how pseudoscience differs from nonscience, bad science and science fraud, although to lie has an important contribution to the concept of pseudoscience where there is deliberate deception in the propagation of falsehoods. Bullshit, once we consider deception as a key element in pseudoscience, stands apart with a complete lack of concern with the truth. While it is possible for a lie to be exposed, bullshit resists such exposure because it makes no definite claims; it lacks epistemic conscientiousness. The purpose of this brief note is to consider the question of bullshit: is the HTA-CEA meme with its focus on assumption driven modelled simulations promoting a program of bullshit. The argument presented here makes clear that with the blatant disregard for the standards of normal science, the standards of fundamental measurement and the rejection of the problem of induction, the HTA-CEA meme in the invention of mind-dependent evidence is nothing but bullshit. A naïve belief we

must put behind us. Finally, while use of the term 'bullshit' might disturb the more genteel and sensitive reader, it is a term that is accepted currency in the philosophy of science.

INTRODUCTION

The health technology assessment or cost effectiveness (HTA-CEA) meme presents a unique yet puzzling, if not disturbing, characteristic: the belief in and commitment to the invention of evidence [1]. This is seen in the commitment over more than 30 years in HTA-CEA to imaginary reference case models [2]. At product launch, where data on product performance are necessarily limited, the recommended approach is to fill evidence gaps by constructing assumption driven modeled simulations to support pricing and access decisions. The most widely known actor with its global imitators is the reference case framework of the National Institute for Health and Care Excellence (NICE) in the UK and its junior imitator the reference model of the Institute for Clinical and Economic Review (ICER) in the US [3 4]

In order to meet the requirements of reference case modeling, there is a blatant disregard for the standards of normal science, the standards of fundamental measurement and the problem of induction [1]. Rather than focusing on the discovery of new yet provisional facts to meet evidence gaps, this strategy rejects any notion of the evolution of mind-independent objective knowledge [5]. Instead, we have a mind dependent analytical framework based on the model builder's choice of structure and the assumptions to populate the model. The entire exercise, which creates claims which are by design fantasy and not empirically evaluable, is an analytical dead end.

At no stage in the creation of the HTA-CEA reference case meme is there any justification for abandoning the unique characteristics that have defined the physical sciences and the more mature social sciences in the 400 years since the scientific revolution of the 17th century [6]. Rather than recognizing the critical importance of the motto of the Royal Society, *nullius in verba* (1662) or take no one's word for it, we are asked to take a model builders word for it with claims for cost-effectiveness in a mind dependent fantasy or, more accurately, a figment of the imagination labeled approximate information.

Instead, there is a meme (not a paradigm) that is probably best viewed as, by default, a framework of analysis that is epistemically relativistic; or anything goes, where normal science has no special claim to knowledge and the acceptance of a particular set of belief is essentially sociological [6]. In other words, there is no special claim for the search for intersubjective knowledge or an observed mind independent reality; there is no meaningful distinction between science, pseudoscience and bullshit [7 8 9]. Rather, the focus of HTA-CEA is on creating or inventing approximate information to support cost-effectiveness claims; the concept of the discovery of new facts, the evolution of objective knowledge has no role.

The purpose of this commentary is to assess the epistemological status of assumption driven simulations to create approximate information that are the core framework for the HTA-CEA belief system with the pivotal role of the impossible QALY [10] .The question of interest is whether we view the endorsement of assumption driven modelled claims as just a type of pseudoscience or whether there are more fundamental issues, not just the open door to deception and fraud, but of

the entire corpus of HTA-CEA belief as bullshit; epitomizing a conscious commitment to the absence of epistemic virtue [11] [12].

JUSTIFICATION

Reviewing the HTA-CEA literature provides no reasoned justification of the blatant disregard of the standards for normal science, standards for fundamental measurement or the problem of induction. While it may be that the leaders in HTA-CEA promoting reference case models were actually unaware of these standards, it is unlikely that even with a limited background in the philosophy of science that they were unaware of the question of demarcation between science and non-science [13]. A more believable reason is possibly more prosaic: they could not provide a plausible case for abandoning these standards. The overriding objective was to create a global and unassailable metric to support resource allocation in health care systems; if this meant developing an analytical framework that would be immediately rejected if the standards were applied, then the least said the better.

The fact that seeking the Holy Grail of a universal multiattribute metric to drive resource allocation fails to meet the key standard that all measures must relate to single attributes with linear, interval and ratio properties has an undeniable Pythonesque quality. In the social sciences, including health technology assessment, where the focus is often on non-physical latent constructs and the measurement of a manifested attribute (e.g., quality of life where the attribute of interest may be needs fulfilment) we require the application of the unique necessary and sufficient Rasch rules to transform observations to measurement [14] [15]. These rules have been recognized and applied globally since the 1950s to the counting of observations. With a few exceptions, the Rasch rules have not been applied to PROs in HTA-CEA. To all intents and purposes, they don't exist. So, there is no needed justification for abandoning them. Ignoring them in the futile Pythonesque search for the Holy Grail metric yields only the mathematical dead end of a multiattribute ordinal preference scale.

The blatant disregard of the standards for fundamental measure, most obvious in the cases of illusory multiattribute scores and the hundreds of PRO instruments is that if it cannot be demonstrated that a 'proposed' measure has, following the Rasch rules, a unidimensional, linear, interval and invariant properties, the default must be that it is an ordinal scale [16] If this is the case then the observations that make up the ordinal scale can only be evaluated with non-parametric statistics. Rasch measurement is the only option for transforming observations to measurement. Rasch is unique in fitting the required data or items to the Rasch model: it is confirmatory and predictive. Item response theory and traditional statistical analysis fit a model to the data which have primacy: they are exploratory and descriptive [14] .

Supporting the transformation by application of Rasch rules is not necessarily unduly complex, given there are a number of software packages available for some 40 years to achieve this transformation, but it is important to understand what the Rasch transformation means. The Rasch model is patient centric, probabilistic and disease or target patient population specific: the likelihood of a respondent successfully responding to a questionnaire item is a function of the difference between the ability of the respondent and the difficulty of the item. If the items selected

for the questionnaire fit the requirements of the Rasch model, then we can have confidence that the questionnaire has the required interval property and can be transformed to a ratio scale with a true zero.

If we seek the Holy Grail of a global metric to support resource allocation, the convenience of abandoning the standards of fundamental measurement throws the baby out with the bathwater: failing to recognize that only a patient centric and disease specific Rasch transformation can support an interval scale means it is the only path to applying standard parametric statistical techniques to gage response to therapy. The fact that the standards of normal science have been rejected as well is a saving grace as it would be impossible, and embarrassing, to apply them to a universe of ordinal scales.

Not only did the various multiattribute instruments create, by design (if unwittingly) nothing more than ordinal preferences, but this ensured in turn that the QALY is an impossible construct. Multiplication requires numerical values that have a meaningful and consistent numerical relationship where multiplication can be applied. Ordinal scales lack this because they do not have equal intervals or a true zero point. Therefore, operations like multiplication, addition, or division do not make sense or cannot be meaningfully applied to ordinal data. Not only was the scale (e.g., EQ-5D-3L preferences) impossible to manipulate to assess therapy response, with the icing on the cake of states worse than death, but it invalidated attempts through reference case models to apply incremental cost-per-QALY claims and cost-per-QALY thresholds. Not surprisingly respondents who are classified by their reference score as in a state worse than death deny that they would prefer to be dead [17]. To this should be added the eugenic abliest or discriminatory implications of preference scores in deciding who is to receive therapeutic support [18] . Most recently, due to these concerns, these preference scores and the QALY have been prohibited in the US under the Section 504 amendments to the Rehabilitation Act of 1973; taking effect in July 2004 [19] .

DEMARCATION

While postmodernism in is various guises, including epistemic relativism, has long ceased to be taken seriously in its failure to recognize the core of what makes science meaningful, the necessary and sufficient condition of falsification, there is still debate over what may be considered the 'fuzzy' demarcation frontier [20]. Popper thought that in invoking falsification he had also solved Hume's problem of induction and the circularity of claims for confirmation [5] . Unfortunately, Popper (to many) oversimplified: do scientists actually behave as falsificationists? If scientists are reluctant to be true falsificationists, dispensing with a theory as soon as a claim is falsified, then we fall back on the question of when is a theory rejected or, more often, modified. To what extent does the belief held by a scientist or team, one that they want to confirm rather than falsify, lead to efforts to reverse the negative results?

Add to this the question of the reluctance by many to reject a theory, is the question of the techniques applied to support a claim for falsification or non-falsification. There are a range of techniques, many specific to disciplines, that need to be evaluated to separate opinion from knowledge. Plus, of course, the inevitable attraction of fraud [21]. The willingness of a small number of scientists to engage in outright fraud is well established, including the invention of patient data

and the modification of images. But fraud can take more subtle forms; including the choice of assessment measure, the trimming of data to eliminate outliers and the suppression of results in a well-crafted peer-reviewed paper. The 'science fictions' as Ritchie describes them can include the employment of hand-picked consultants to craft models and the contracting with paper mills, as well as successful and less successful fraud on unsuspecting colleagues and journal editors.

In the case of HTA-CEA, all of the above apply as there seems no obvious effort to address the issue of epistemic knowledge. The problem is, however, much wider: HTA-CEA is not committed to falsification because it has justified and scrupulously followed for over 30 years the creation of imaginary claims which are designed to avoid any possibility of being falsified; the gold standard of approximate information. This denial is best exemplified in the standard textbook which gives a detailed framework for creating imaginary claims; reinforced by the CHEERS 2022 guide to creating 'acceptable' imaginary reference case model claims for submission to leading journals [2] [22] [1].

It is important for our assessment of the HTA-CEA belief in imaginary modelled claims to recognize that Popper continued to maintain that falsifiability is both the necessary and a sufficient criterion for demarcation; a sentence or a theory is empirical-scientific if and only if it is falsifiable [23]. For practical purposes this is the core of demarcation whether you take a rigid single assessment standard or a more sophisticated accumulation of evidence framework. This is important because we are dealing in HTA-CEA with deliberately designed non-falsifiable claims which reject the entire corpus of belief in the scientific method, as described by Popper, which has been focused since the 17th century on discovering a provisional mind-independent external reality. In contrast, the HTA-CEA meme, in rejecting the standards of normal science, denies demarcation to ensure all reference model claims are non-science. The approximate information goal is a mind-dependent imaginary fantasy.

ASSUMPTIONS

It must, from those better versed in the philosophy of science, seem bizarre to place any credibility on reference case models which not only disregard standards for normal and fundamental measurement, but which are driven entirely by assumptions. In the context of scientific inquiry, assumptions are deemed valid primarily if they support the generation of testable and falsifiable hypotheses. This approach emphasizes empirical testing and the continuous refinement of theories, aligning with Popper's formulation of problems and their solution, the evolution of objective knowledge through conjecture, refutation, error elimination and feedback to grasp, albeit provisionally, a mind-independent external reality. While the problem of induction highlights the inherent uncertainty in projecting future outcomes based on past observations, the focus on testability and falsifiability provides a pragmatic pathway for advancing knowledge and understanding in science. This is in direct contrast with HTA-CEA. In simulation modelling, assumptions are the first and last defense of a reference case model.

Unfortunately, if we accept the problem of induction, raised first by Hume 1748 [24], the problem is that it cannot be proved or disproved, the fact that an assumption has been 'true' in the past, does not mean it is true for the future. The future may, in retrospect, consider that assumption choice a

deliberate lie; future futures failing to resemble past futures. The point is made by Russell: *The man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that more refined view on the uniformity of nature would have been useful to the chicken* (p.33) [25].

The implications of this blatant disregard for accepted standards are profound for how we view the HTA-CEA meme together with the corpus of ratio-based cost-outcomes models and claims... In rejecting the notion of demarcation and the criteria for science, the HTA-CEA meme falls clearly in the category of non-science. The question then becomes: do we categorize this non-science as pseudoscience or as bullshit. This is not a trivial distinction given the position taken by Frankfurt that bullshit is the greater enemy of truth than lies [9].

DEFINING PSEUDOSCIENCE

The Oxford English Dictionary (OED) defines pseudoscience as: *A pretended or spurious science; a collection of related beliefs about the world mistakenly regarded as being based on scientific method or as having the status that scientific truths now have.* The emphasis by a number of writers, as noted by Hansson, is that pseudoscience is non-science posing as science; accepted beliefs masquerade as genuinely scientific ones [26]. In other words, pseudoscience is seen to involve a sustained effort to promote standpoints different from those that have scientific legitimacy; it deviates from the quality criteria of science: reliability, fruitfulness and practical usefulness.

When the standards of normal science are summarized, it is to consider credible claims, empirical evaluation and replication (and reproduction). The first question is, therefore, the status of a credible claim: what distinguishes a credible claim from pseudoscience? Hansson proposes two criteria to classify a claim as pseudoscience, where the second criteria can take a narrow (ii) or wider form (iii):

(i) *It is at variance with the most reliable knowledge about its subject matter that is currently available; and*

(ii) *it is part of a non-scientific doctrine whose major proponents try to create the impression that it is scientific or*

(iii) *it is part of a doctrine whose major proponents try to create the impression that it represents the most reliable knowledge on its subject matter.*

Criteria (i) is a revised version of an earlier criteria which simply stated that the activity is not scientific. The revised version follows from consideration that demarcation recognizes the quality of science; the function of science as a process of discovery or fact-finding that has the objective of providing the most reliable current information. The process of discovery follows rules; pseudoscience while often attempting to mimic science, fails at this basic level. This does not, it should be emphasized, imply, following Popper a rigid application of the test for falsifiability. We could opt for a more sophisticated falsification as detailed by Lakatos or adopt a multi-criterion as opposed to a mono-criterion approach, a check list, to label a theory or analytical framework as pseudoscience practice [27]. Elements of such a list could include a belief in the authority of leaders

in a field (particularly apt for HTA-CEA), non-testable claims, the process by which claims are created, nonreplicable or nonreproducible claims and the rejection of refutation. A major problem with the multi-criterion approach is which criteria do we choose, in particular in the assessment of clinical claims and the measurement basis for those claims, where it is suggested that the majority of such claims fail replication. If a systematic review is the basis for claims, has the review excluded dubious claims? Should this corpus of pivotal trial knowledge be considered false? If so, basing assumptions on such data with no support from replication or reproduction is, even with the imaginary HTA-CEA mindset, foolhardy.

In terms of criteria (ii) and (iii) above, the question is the extent we wish to cast a narrower or wider net for a definition of activities that we characterize as pseudoscience. The narrower net is to consider activities characteristic of individuated sciences or belief systems that are seen as specific branches of knowledge. The wider net (iii) sees science as activities which share a common focus with the individual sciences as merely examples.

As Hansson makes clear, there is a critical distinction between science and pseudoscience [22]. In the case of science there is agreement on the essential unity of the objectives of scientific enquiry and the process of creating and evaluating credible or falsifiable claims; ultimately, the determination of provisional beliefs that are epistemologically warranted. Pseudoscience lacks any unifying theme or agreement: there is a smorgasbord of specific pseudoscience activities ranging from the HTA-CEA contribution, to tarot card reading to intelligent design. Judged by Popper's standards, all are fairy stories.

Allied to the rejection or ignorance of the standards for fundamental measurement leaves us with no doubt that reference case models meet the criteria for pseudoscience under (i) above. In respect of criteria (ii) this non-scientific doctrine of approximate information was promoted with the intention of creating a scientific aura for the less well informed. In respect if criteria (iii) this approximate information was given guidelines for assembly in order to represent it as the most reliable knowledge that could populate a Markov model or similar decision framework. The problem of induction was overlooked, just ignored or not understood.

There is a factor that we consider in more detail in Frankfurts distinction between pseudoscience and bullshit: the question of lying, by commission or omission. Hansson's (and the OEDs) definition of pseudoscience does not imply that in pseudoscience, there is a deliberate attempt to lie, to inject falsehoods to avoid the truth. However, it might seem obvious, that in order to lie the truth must be known. If lying is part of assumption choice in modelled simulations then the 'true' assumption must be known.

Focusing on the HTA-CEA belief in assumption driven simulations the principal question is why were the reference case assumption driven simulations developed? The answer is quite clear, rather than a research program to meet evidence gaps, possibly driven some variant of value of supplementary information, it is easier, and more lucrative for consultants, to invent information. Necessity is the mother of invention: fill gaps with always justified assumptions from the literature. The problem of induction, if recognized, is ignored; the model hinges on the claim by the model builder that the assumptions are 'reasonable'. This focuses on the commitment to 'justified'

approximate information. The truth-value of which is far from obvious. After all, any set of assumptions can be justified in reference to any number of competing sets. It is a game or, more realistically, a marketing exercise where imaginary claims justify the sponsor's product. To put this is terms of needed approximate information is sheer nonsense. The term 'reasonable' is meaningless. If one assumption is reasonable or true for the future, then we must be aware of other assumptions that are not reasonable or true for the future.

The supplementary question is why do we need these simulation models presented in terms of cost effectiveness analysis? Again, the answer is easy: to justify HTA-CEA as a policy relevant activity, it must provide a framework for allocating resources. With imaginary claims built on assumptions yet non-evaluable, the easier it is to determine whether to shut the access door to new therapies or restrict access to certain groups; possibly subject to price negations to squeeze under a cost-per-QALY threshold.

To illustrate the gulf between fundamental measurement and the multiattribute measurement debacle, consider the assessment of mobility. The EQ-5D respondent is asked to tick he box that best describes your health today: in the case of mobility (one of 5 health dimensions) the options are: (i) I have no problems walking about; (ii) I have some problems walking about or (iii) I am confined to bed. These are observations, ordinal responses, which are weighted to as part of a preference algorithm. Response (i) has a weight of zero, response (ii) a weight of 0.069; and (iii) a weight of 0.314. These are numbers not measurement; we cannot infer the size of the differences. The EQ-5D preference scores may provide an illusion of measurement but they are not fundamental measures which are the only basis for measurement. The time tradeoff (TTO) technique yields only interval preferences.

The Rasch model starts from the position that mobility (like temperature) is a latent construct The first step is to manifest or invent a property of mobility as an object to be measured; the second step is to construct and evaluate an assessment instrument. We don't measure mobility directly but through its effect on other objects which can measure and infer a particular aspect of mobility status. This is standard practice in the physical and medical sciences. For psychological or PRO latent constructs, the same procedure is followed by applying Rasch rules to create an instrument which we infer has the required unidimensional, linear and interval and invariant properties. An example would be the latent construct quality of life with the manifest of interest the needs of patients. This is a long-recognized aspect of patent value and the value to patients of different health states [28]. As an attribute of interest manifested from quality of life as a latent construct it has been applied in a number of PRO instruments applying Rasch rules. As a fundamental unidimensional linear, interval and invariant measure, it can be transformed to a ratio measure to create what has been described as the need quality of life measure (N-QOL) [29] . This is of course a disease or target patient group specific measure of needs and the extent to which therapy implantation shifts the distribution of needs fulfilled. There are no global implications for a Holy Grail metric as needs are specific to a patient group.

To support therapy decisions in HTA-CEA, the intent was to construct claims based on mind-dependent approximate information, assumptions from the literature and limited data from pivotal clinical trials, not to create unidimensional, linear, interval and invariant measures to support

hypothesis testing but to subvert it. The EQ-5D instruments are not designed to respect the rules for fundamental measurement. They go no further than proposing latent constructs or health dimensions and assume that these can be captured by three (or five) response levels. There is no concept that what is required for measurement is to identify a meaningful manifestation of that latent construct and construct the appropriate instrumentation. We don't bundle together health state descriptions in a crude attempt to short-cut meaningful measurement by assigning subjective response levels.

It is also worth noting that a Rasch measure for a latent construct manifestation cannot to inferred *ex post facto* from existing PRO instruments. In order to apply Rasch rules to assess whether item reduction, for example will yield a sort-of Rasch scale involves a logical contradiction: in order to apply Rasch criteria with parametric statistical techniques we have to apply them to a measure which has the required interval standards. The exercise is pointless.

In this commitment to nonscience, the HTA-CEA leadership succeeded beyond their wildest dreams with the acceptance of modelled imaginary claims as the necessary first step in formulary pricing and placement; a first and, unfortunately, last step as there was no path forward from the choice of multiattribute health state community preferences. This was a non-starter. Certainly, gatekeepers could engage academic groups to act as reliable knowledge police to give a model or a revised version their good housekeeping seal of approval without recognizing the lack of the appropriate epistemic standards that are recognized outside of HTA-CEA in the sciences and other social sciences. The result is literally tens of thousands of published reference-type models that take advantage of this easy route via assumption based approximate information to support value claims. Marketing made easy when the audience, both in academia and the public sector health system decision makers, lapped it up. If they had any awareness of the scientific method or even, more unlikely, of fundamental measurement then this was easily dispensed with. The HTA-CEA imaginary claim meme dominated and continues to dominate as a parallel, relativistic universe.

LIES AND FRAUD

In the present context of the dominance of the HTA-CEA meme belief in the relevance as decision criteria of assumption driven non-evaluable simulations raises the question of whether these beliefs are epistemologically warranted. This leads to the distinction raised earlier between the category of pseudoscience where there is no intention to deceive and pseudoscience with lies which, with the intention to deceive, is simply fraud. It is this second category which is of particular interest in categorizing the HTA-CEA belief system and practice in creating assumption driven simulated imaginary claims. Even so, it is worth emphasizing Ladyman's contention that not all pseudoscience is necessarily science fraud [30]. While most of us at one time or another are lacking in epistemic conscientiousness one argument is that for a statement to be considered as bullshit as opposed to pseudoscience some minimal degree of unconscientiousness is required [31]. This seems weak; after all the essence of bullshit is the culpable indifference to truth; which applies to both the activity and the results in HTA-CEA simulation modeling.

While not perceived by an analyst as an indifference to truth, the belief that the particular choice of assumptions to populate a simulation has truth-value, is wrong. Putting intent and deception in

assumption choice to one side, any set of assumptions can be claimed to have truth value for a simulation as long as there is a willingness to ignore the problem of induction: a lack of understanding that the fact that past futures have resembled past pasts does not mean that future futures will resemble future pasts. The willingness to rest a case for imaginary and non-evaluable claims in HTA-CEA for cost-effectiveness on literally dozens of assumptions makes clear that there is no acceptance or even recognition of the problem of induction; a belief that relies on the denial of recognized standards for progress in science. Of course, assumptions have a role in theory construction and claims, but only if those claims are falsifiable in a meaningful time frame to meet the demands of decision makers.

The choice of assumption, deliberate or not, extends to measurement. Just as the problem of induction is sidelined in HTA-CEA so are the axioms of fundamental measurement. The willingness to challenge measurement standards and claim by ICER that ordinal scores are considered by health economists as ratio measures in disguise, together with the denial that measurement must refer only, not to composite algorithmic scores, but single attributes with linear, interval and invariant properties is not even discussed [32]. More than the denial of truth-seeking as justification for assumptions, this should be seen as a culpable indifference to measurement, notably Rasch modeling for PRO interval measures. A deliberate rejection of standards in place for over 60 years [33]. Whether this is tantamount to lying as opposed to deception is an open question. It can hardly be considered an unfortunate oversight. A key factor is the gold standard QALY as the driver for resource allocation. This fails the standard for fundamental measurement; it is an impossible mathematical construct [34]. Yet it continues to play a central role in promoting the relevance of the HTA-CEA meme.

FRAUD AND PSEUDOSCIENCE

There is a question of whether fraud should be characterized as pseudoscience given that fraudsters claim to commit to the standards of normal science for evidence creation and falsifiable claims. One argument is that academic fraudsters, operating within the science belief system, whether merely tweaking their data, inventing patient data or relying on a paper mill are still endeavoring to give the impression to colleagues and journal editors that they are scrupulously following the accepted standards of analysis. As such, while lies and consequent fraud may be present, they might not be considered as outside the pale of science; they are not obviously pursuing pseudoscience merely taking advantage of opportunities offered by imaginary simulated model claims.

An alternative view goes to intent to deceive. Certainly, there is a smokescreen of compliance with standards, but the object in HTA-CEA fraud is to produce a false or manufactured empirical claim that is supported at the required decision level. Fraudulent endeavors may persist for years, garnering accolades for the aspiring academic. A major problem is that fraudulent papers may be accepted, they may even be recognized as a seminal contribution, yet the final decision to retract a paper may be years in the making with those accused making every effort to obfuscate, often aided and abetted by colleagues and institutions. Yet these endeavors show no concern for the truth; as long as the false analysis passes muster with peer reviewers, journal editors and colleagues

who are asked to be joint authors. Intent to deceive is a judgement call; all too often the perpetrators escape judgement.

Fraud is typically at the individual or small team level; it is opportunistic and intentional. It operates within a framework which is scientific where the variety of assessment procedures for reporting a satisfactory p-value, allied with the reluctance of journals to undertake a detailed assessment of the process of data collection and analysis, makes fraud an attractive option. In other words, the fraudster is trying to insert a particular falsehood into an accepted system of beliefs. The intent is fraudulent, even if the financial rewards are limited, with the author(s) hiding under cover of an accepted set of standards, a doctrine, which can hardly be described as non-scientific (see [ii] above). At the same time, the false knowledge gained by fraud is typically viewed as consistent with the subject matter knowledge for a particular field of study; it is reverse engineered to be seen as consistent with progress and the discovery of new facts. A fraud can be committed to meet the standard for falsification and demarcation where the entire exercise has been created from a set if patient data that has been designed to produce the required results. Unless it is possible to make a definitive case that evidence has been fabricated, then the fraudulent paper remains, at least until exposed, within the scope of science. Where intent to deceive is present there is no excuse to give the fraudster an easy ride. From this perspective, fraud by academics is simply part of everyday experience with HTA-CEA pseudoscience and should be seen as an integral part of the approximate information meme in the promotion of lies and attempts to relegate disconcerting facts such as standards for fundamental measurement to oblivion.

DEFINING BULLSHIT

The term bullshit entered the lexicon of the philosophy of science in 1986 with Frankfurt's seminal paper *On Bullshit,* with wider currency in its publication, as a book, *On Bullshit* in 2005 and *On Truth* in 2007 [9 35] . The essence of bullshit is culpable unconcern with the truth or indifference towards the truth; the bullshitter is not concerned with truth or falsity; while capable of responding to reasons and argument, they fail to do so. They are epistemically unconscientious, but in two respects. In a narrow sense one may be indifferent towards the truth of a statement while in a wider sense one may care without taking care. While Frankfurt by and large endorses the narrow concept a more flexible or wider view is that those promoting bullshit exhibit a culpable lack of epistemic conscientiousness in adhering to a belief system that manifests a self-willed ineptitude, regardless of whether this manifests as indifference toward the truth.

This lack of concern with the truth of statements is, for Frankfurt, the essence of bullshit; bullshit is not false but phony. This sets bullshit apart from lying which is only possible if one knows the truth. The liar is someone who deliberately promulgates a falsehood. As such, the deception that characterizes lying is more insidious than the deception that characterizes bullshit where the former intent is to divert attention from what Frankfurt describes as a correct apprehension of reality. Liars are aware they are making false statements, including false assumptions. Bullshitters are not interested.

THE APPROXIMATE INFORMATION MASQUERADE

With sufficient scrutiny lies can be disproved, the claim for an ordinal score assumed to act as a ratio score is easily demolished if the actors are aware of the standards for fundamental measurement. But bullshit, which makes no claims at all cannot be demolished; it is unconcerned with the truth or, in our terms standards of normal science, fundamental measurement and induction. In this respect we might accept, at least provisionally, that the HTA-CEA meme is bullshit rather than pseudoscience. The intent in assumption driven simulations to avoid any opportunity to allow falsification of claims; which is straightforward as the outcomes and timelines involved effectively eliminate the possibility. It is not deception as the intent is clear. There is no apparent concern with truth value; only with convincing the intended audience that the imaginary claim has a singularly important role in decision making.

The distinction between lies and bullshit is in the intent and motivation behind a communication. If the communication is dependent on omissions for it to be accepted by a recipient such as a formulary committee or health system in therapy assessment, then the acceptance depends upon maintaining a masquerade of revealed truth. Those promoting and ensuring the transmission fidelity of the belief system, the relativist position that truth is consensus, maintained by rhetoric, persuasion and authority ensures that there is no appeal to superior evidence, the normal science standard.

However, if the HTA masquerade is promoted as the invention of reliable approximate information, the creation of imaginary evidence could be viewed as fraud where the creation of evidence is no different from contracting with a paper mill where the parties are unconcerned with what they are saying is true or false; just that it is accepted. The factual accuracy of claims takes second place to the need to be seen as knowledgeable, impressing the recipient. There is no coherent epistemic framework that can be justified.

A key point to note is that in the reference guidelines and in the promotion of the HTA-CEA belief system there are no qualifications presented that might cast doubt on the embrace of approximate information. The role of assumptions is taken at face value; the role of simulations is taken at face value. The belief system explicitly maintains that this is the necessary and sufficient framework for supporting formulary decisions with the QALY as the center piece. Yet omissions of substance are no different from lying. To admit to omissions regarding, for example, the failure to apply the recognized standards of fundamental measurement is put aside because it would destroy the entire analytical framework. Inconvenient critiques are ignored and, where possible never published. The leading textbook for the application of HTA-CEA makes no mention of the unique status of Rasch measurement in creating PRO value claims; where measurement is briefly mentioned, the discussion is a best confused and uninformative [2]. Is it reasonable to argue that deception is fundamental to the success of modeled HTA-CEA claims? Or is there no intention to deceive?

EMBRACING BULLSHIT

The hallmark of the HTA-CEA belief system and reference case simulation modeling is the invention of evidence or approximate information to support formulary decisions and the allocation of health care resources. This lack of concern with the truth is not inadvertent; it does

not reflect an unavoidable need to invent evidence. It is precisely what HTA-CEA is intended to accomplish. The rejection of the standards for normal science, fundamental measurement and induction all open the door to bullshit. With the creation of fantasy reference cases there is neither concern with how things really are as the crafting of a reference case for a hypothetical assumption driven future can be neither on the side of the true or the false (p.56) [9].

The HTA-CEA belief system in reference case simulation models denies the relevance of the standards of normal science, denies the role of fundamental measurement and ignores the problem of induction. This blatant disregard for standards that have been agreed since the scientific revolution of the 17th century is made the more absurd when there is no statement of why this rejection is essential. There is, apparently, no need to justify abandoning standards that have been developed to support discovery of new facts. The over-riding objective is to propose an impossible metric: cost-per-QALY outcomes for healthcare resource allocation. To achieve this, all common-sense standards are rejected; the end justifies the means.

At the same time, the accepted standards are not just rejected, the rejection is implicit not explicit. There is no statement of why it is necessary to reject the standards of normal science to create the allocative metric; neither is there a statement of why it is essential to ignore the standards for fundamental measurement and Rasch measurement to create the desired metric; and nor is there any statement of why the problem of induction is irrelevant to assumption driven model simulated claims.

It is against this background that we must address the question: does this unique character of HTA-CEA and simulation modelling point to the embrace of bullshit and not pseudoscience? In this context it is worth quoting Frankfurt: bullshit is a greater enemy of the truth than lies (p. 61) [9]. The argument is straightforward: *Someone who lies and someone who tells the truth are playing on the opposite sides …each responds to the facts as he understands them, although the response of the one is guided by the authority of the truth while the response of the other defies that authority and refuses to meet its demands. The bullshitter ignores these demands altogether* (p. 60-61) [9]. In other words, pseudoscience continues to recognize the standards for normal, science, fundamental measurement and assumptions to support evaluable claims. This is the context within which false claims or lies are made; paper mills do not reject the standards of science as they are essential to making a plausible or foolproof false case. The liar and someone who tells the truth agree on the character of the playing field. This goes to the point made by Frankfurt that the lies of pseudoscience are deliberate falsehoods that imply a concern with truth-values. The truth is the springboard for a well-crafted lie designed to *insert a particular falsehood at a specific point in a set of systems or beliefs* (p. 51) [9]. To invent a lie the perpetrator needs to know what is the truth.

However, lying is essentially a one-off action; it should not be construed as a program of repeated lying focused on a common set of non-truth values. Engaging with a paper factory to create, with well crated lies, a research paper to gain the approval of peers, is not a program of lies even if a number of papers may be involved. The fear of discovery is always present; lies have a low survival value. The bullshitter has no such concerns.

The supporters or leaders in HTA-CEA are not liars; they are not engaged in promoting pseudoscience, knowing the truth yet promoting knowingly the invention of false statements as a counterpoint or misrepresentation to inserting the truth. For HTA-CEA there is no concept of a false statement; there is no intention to deceive nor to report the truth or conceal it. There is no intention of describing reality or an endeavor to communicate the truth but a commitment to a well-crafted and focused program of producing bullshit. There is no attempt to meet accepted standards but a blatant disregard of those standards. Bullshit is disconnected from the truth; where the truth is of no value. There is no notion either of conveying something that is false; rather it is a phony activity, produced without concern for truth-values or to report the truth or conceal it (p, 55) [9] . This does not mean that bullshit is not carefully wrought; messages that are tailored with advertising are prime examples. It means that there is no reason why a carefully constructed bullshit presentation cannot be considered as sufficient to inform formulary decisions. After all, few recipients of such models bother to probe very far below the surface.

A blatant rejection of standards underpins the reference case model. But this rejection is continuing through high transmission fidelity supporting the HTA-CES meme to create an ongoing program of uncontested bullshit claims for pricing and access for new therapies. This has been promoted assiduously by groups such as ISPOR with a surprising success for now over 30 years. While it would be presumptuous to call this reinforcement of belief as promoting, following Dawkins, a mind virus, there is an unwavering, almost placid, acceptance of the *status quo* [36] . One example would be the focus on the extension to the EQ-5D-3L/5L with the EQ-Health and Wellbeing (EQ-HWB) index which merely continues the train wreck of the EuroQoL instruments in promoting an 'enhanced' ordinal score with a new vision of the Pythonesque Holy Grail to make more efficient and responsive tool to support invented social awareness in the allocation of health care resources [37] ejecting hypothesis testing in favor of approximate information and the implications that follow from this would appear to fit this interpretation of intent. It is as though the leadership in HTA-CEA is really unaware of the standards of normal science, fundamental measurement and induction; with a continuing willingness to pursue non-falsifiable and impossible cost-effectiveness claims. Whether they are aware in the case, for example, of the EQ-5D-3L ordinal multiattribute status of the preference score, and its implications, is a moot point. If they had been aware the instrument should never have been developed in the first place. It is a classic example of measurement bullshit. The EQ-5D-3L instrument is a singularly pointless endeavor that produces community preference scores for clinical bundles defining health states in the range unity to minus infinity; as there is no true zero. The resulting scale is, charitably, best described as a convenience multiattribute ordinal scale with nowhere to go. The neglect of the standards of fundamental measure has led to the production, not of a measure, but bunk

. To challenge and overthrow the HTA-CEA belief system requires a new paradigm that returns to the required standards of normal science, fundamental measurement and an appreciation of the problem of induction. Few seem prepared to do this. Bullshit in HTA-CEA gives the aura of reasonableness for a non-scientific program. For followers of a meme where the believer unconsciously and automatically rejects substantive criticism, the truth-value of statements is of no interest. A situation that Dawkins describes as the presence of a mind virus where the HTA-CEA advocate and practitioners *who are impelled by a deep inner conviction* that the approximate

information simulation framework *is true, or right or virtuous but which owes nothing to evidence or reason* [36]. It is as though the embrace of the mind virus had transformed pseudoscience to bullshit with no concern for the truth or falsity of statements. Belief in the ratio transformed ordinal preference score, the QALY and imaginary modelled cost-effectiveness claims may be more bizarre, but only if they are seen as something other than a mystery that should never be challenged because they are, in effect, so impossible as to be believable: following Tertullian *Certum est quia impossibile est* (It is certain because it is impossible) [36].

For HTA-CEA believers the possibility of a mind independent external reality is irrelevant as is any thought of discovery and epistemic conscientiousness. Claims are made to suit present purposes; the approximate information modelling is product marketing, a more complex form of advertising, with the carefully crafted bullshit model supporting the sponsors product. The indifference to truth in model construction and claims is of no concern. The model provides the aura of scientific pretension with an always justified choice of assumptions and impossible incremental cost-per-QALY claims defended as the master stroke for reliable cost-effectiveness claims. The model edifice supported by copious tables, tornado diagrams and sensitivity analyses, notably probabilistic sensitivity analysis which all rely on the mathematically impossible QALY. Textbooks and practice guidelines, the province of organizations such as the ISPOR that support the bullshit program ensuring impressive transmission fidelity.

Taken overall, the assumption driven model simulations can be viewed as bullshit; a complete indifference or lack of concern towards what we can call the truth or the provisional discovery of new facts. Indifference as bullshit should be distinguished from lies and fraud in the advocacy of assumption driven simulations although they are all captured under the umbrella of bullshit. Approximate information is totally divorced from the focus over the past 400 years with the invention of discovery as a hallmark of the scientific revolution [6]. The assumption that there are discoveries to be made with the focus on new, yet provisional facts is absent from HTA-CEA reference case modelling. A claim for inventing approximate information defies hypothesis testing and falsification. There is no concept of a search for provisional truth; a commitment to the evolution of objective knowledge supported by Popper's commitment to falsification through a process of conjecture and refutation. This falls by the wayside when the commitment in HTA-CEA is to creating approximate information. HTA-CEA have created the antithesis to the ideas and commitments of generations of scientists; a step back to the dark ages.

There is no intention with bullshit to report the truth nor to conceal it. Descriptions of reality, or what is conceived of as reality, are of no interest. The analyst might be dimly aware of the importance of transforming ordinal to interval scores to create a well calibrated measure but continues to use ordinal scores because it is impossible to transform multiattribute scales to single attribute, linear, interval and invariant measure. The truth is put to one side with a claim that when looked at through HTA-CEA spectacles, in the case of the Institute for Clinical and Economic Review (ICER) the belief that ordinal scales are ratio measures in disguise. Without, it might be noted, any justification for this position. If this is the case, then it should be considered an omission, but not necessarily an intent to deceive. If the analyst is completely ignorant of fundamental

measurement and the classification of measures, then the question of the difference between an ordinal and an interval measure is a foreign country.

As these claims have been designed to influence decision makers; particularly the more ill-informed or those without the requisite training, then the reference case and cost-per-QALY ratios become the standard for formulary decisions. A commitment to what sociologists in the philosophy of science call the relativist strong program is widely recognized, although not raised in HTA-CEA as a defense of the indefensible. For the strong program evidence is never 'discovered' but always constructed within a social community [6]. For HTA-CEA the success of bullshit depends on its ability to mobilize support in the invention of evidence; apostasy is frowned upon. Science is not about generating new knowledge but about rhetoric, persuasion and authority. Different communities are addicted to different games. Bullshit in these terms is an addiction; there is no concept or interest in an appeal to new evidence and the relegation of beliefs. A bullshit program is an analytical dead end but a program in HTA-CEA that has now existed for over 30 years.

CONCLUSIONS

The rejection by the leadership and membership in HTA-CEA of the standards of normal science, fundamental measurement and induction and their support of the invention of approximate, imaginary or phony evidence as a mind dependent reality, gives no option other than to classify the focus on assumption driven simulations as bullshit. Most importantly, the lack, intentional or otherwise, of an epistemic consciousness also relegates it to non-science. The creation of knowledge, the process and progress of the discovery of new yet provisional facts, the understanding that there is a mind independent reality not a mind dependent reality, point to the absurdity and danger of embracing the HTA-CEA belief system in real world decision making.

Instead, we have the paradoxical situation, where rather than lies being the greater enemy of truth, the HTA belief system as the enemy of truth is appropriately characterized as bullshit; indifference to truth in bullshit is the greater enemy of truth. Lies are irrelevant to believers in in the HTA-CEA meme; if lies are accepted (but not as seen as lies as such except to an independent observer) then bullshit takes the high ground. This leaves us with the approximate information modelling simulation belief system that has been lapped up by thousands. The point is well put by Ladyman: the way bullshit and pseudoscience disconnect us from the truth is more insidious than lying for we may end up with not just false beliefs but no beliefs at all [13] . We want to lose contact with a mind independent reality.

Both bullshit and pseudoscience produce epistemic noise with only a superficial resemblance to the truth; assertions are made that this epistemic noise is a scientific issue with modes of presentation to dress it up in a more 'serious' aura. Rather than conflate the terms pseudoscience and bullshit, the latter is the greater concern. Constructions such as increment cost-per-QALY and cost-per-QALY thresholds supporting modelled claims are pure and simple bullshit where claims are divorced from reality; unconnected from the truth. It is not, therefore, a question of what side of the demarcation divide you all on, but the fact that the notion of demarcation has no relevance whatsoever to modelling bullshit; continuing the relativist tradition. The HTA-CEA belief system in encouraging a program of non-factual and non-evaluable claims for an unforseeable future that

is focused on the endless production and publication of bullshit, is hardly a future to look forward to in health care decision making. One has to be singularly naïve to pay any attention to the HTA-CEA reference methodology where the claim can never be wrong. It may never be wrong, but what is the downside of relying on bullshit to allow or deny access to care. Are we in any position to audit what damage NICE in the UK has wrought through denial of access given the duty of care for patients.

As a final point it is important to consider again the quality of science, in particular the notion of fruitfulness where an assessment is made between studies to assess whether one is scientifically more valuable. Discovery as part of the evolution of objective knowledge is impossible within the HTA-CEA meme or belief system with the fulsome embrace of bullshit despite an aura of scientific pretensions. There is no concept of progress, the discovery of new yet provisional facts and new perspectives on existing assessments. The HTA-CEA belief in assumption driven simulations is a dead end; it is barren. A one-off non-evaluable cost-effectiveness claim to convince the less tutored decision makers. It rejects out of hand the standards of normal science, standards of fundamental measurement and the problem of induction. This should come as no surprise; bullshit is hardly a basis for the evaluation of the truth-value of claims as it is completely indifferent to the question. Those subscribing to the HTA-CEA meme or belief system show no motivation to arrive at the truth; no authentic motivation for provisional knowledge. Rather, they have chosen the comfortable and non-challenged rabbit hole of bullshit.

The case for a new start in HTA has been detailed in a recently issued Certificate Program from the School of Pharmacy, University of Wyoming. This is detailed below:

UNIVERSITY OF WYOMING CERTIFICATE PROGRAM

A NEW START IN HEALTH TECHNOLOGY ASSESSMENT

For those who are interested in following up the arguments presented here for Rasch standard patient centric value claims, the recently released on-line University of Wyoming Certificate Program: A New Start in Health Technology Assessment is recommended.

The Certificate Program is in three parts:

- Part I: Required evidentiary standards for product and therapy assessment
- Part II: The failure of approximate modelled information for therapy decisions
- Part III: Formulary submission value claims and protocols for a new start in product evaluation in health system management

The Certificate Program package includes extensive notes (overall for the 14 modules 85,000 words), audiovisual presentations and a short true-false and multiple-choice assessment for each module. The cost of the Certificate Program is $875 USD with 20.5 hours of ACPE credit. For those who do not need ACPE accreditation, the University of Wyoming will

provide a Certificate of Completion. Following interest already expressed, for those introducing the proposed new start standards for technology assessment there will be a program of one- and two-day workshops and on-line seminars to support course development and alternative program structures to meet local needs. There will also be a series of working papers to explore specific aspects of the new start program.

The link to register in the Certificate Program is: https://www.uwyo.edu/pharmacy/resources/certificate-program-a-new-start-in-healthtechnology-assessment.html

This program was developed by Dr Paul C. Langley, Ph.D. a health economist. He is currently resident in Tucson AZ. If there are questions regarding content please contact Dr Langley at: langleylapaloma@gmail.com

REFERENCES

[1] Langley P. Nothing to Cheer About: Endorsing Imaginary Economic Evaluations and Value Claims with CHEERS 22 [version 1; peer review: 2 approved]. *F1000Research* 2022, **11**:248 (https://doi.org/10.12688/f1000research.109389.1)

[2] Drummond M, Sculpher M, Claxton K et al. Methods for the Economic Evaluation of Health Care Programs (4th Ed.) New York: Oxford University Press, 2015

[3] National Institute for Health and Clinical Excellence. Health technology evaluation: the manual. Economic Evaluation. October 2023. https://www.nice.org.uk/process/pmg36/chapter/economic-evaluation-2

[4] Institute for Clinical and Economic Review. ICER's Reference Case for Economic Evaluations: Elements and Rationale (September 2023) https://icer.org/wp-content/uploads/2023/10/ICER_Reference-Case_For-Publication_Sept2023.pdf

[5] Popper K. Objective Knowledge: An Evolutionary Approach (Rev. Ed.). Oxford: Clarendon Press, 1979

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

[7] Pigliucci M, Boudry M. Philosophy of Pseudoscience: Reconsidering the Demarcation Problem. Chicago: University of Chicago Press, 2013

[8] Hansson S. "Science and Pseudo-Science", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), ZaltaE (ed.), https://plato.stanford.edu/archives/fall2021/entries/pseudo-science/

[9] Frankfurt H. On Bullshit. Princeton: Princeton University Press 2005

[10] Langley PC and McKenna SP. Measurement, modeling and QALYs [version 1; peer review: 2 approved]. *F1000Research* 2020, **9**:1048 (https://doi.org/10.12688/f1000research.25039.1)

[11] Langley P. Facilitating bias in cost-effectiveness analysis: CHEERS 2022 and the creation of assumption-driven imaginary value claims in health technology assessment [version 1; peer review: 3 approved]. *F1000Research* 2022, **11**:993 (https://doi.org/10.12688/f1000research.123709.1)

[12] Hursthouse R, Pettigrove G. Virtue Ethics, *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta E . Nodelman U (eds.)
https://plato.stanford.edu/archives/fall2023/entries/ethics-virtue/

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

[14] Bond T, Yan Z Heene M. Applying the Rasch Model: Fndamental Measurement in the Human Sciences (4th Ed). New York: Routledge, 2021.

[15] Andrich D, Marais I. A Course in Rasch Measurement Theory: Measuring in the Educational. Social and Health Sciences. Singapore: Springer, 2019

[16] McKenna S, Heaney A. Composite outcome measurement in clinical research: the triumph of illusion over reality. *J Med Econ*. 2020;73(10):1196-1204

[17] Benfort L, Gerdle B, Husberg M et al. People in states worse than dead according to the EQ-5D UK value set: Would they rather be dead? *Qual Life Res*. 2018;27(7):1827-33

[18] Schneider P. The QALY is ableist: on the unethical implications of health states worse than dead. *Qual Life Res*.2022;31:1545-52

[19] US Government. Health and Human Services Department. Nondiscrimination on the Basis of Disability in Programs or Activities Receiving Federal Financial Assistance: Final Rule. Federal Register May 9, 2020. 89 FR 40066 https://www.federalregister.gov/public-inspection/2024-09237/nondiscrimination-on-the-basis-of-disability-in-programs-or-activities-receiving-federal-financial

[20] Wright B, Linacre J. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil.* 1989; 70(12):857-60

[21] Ritchie S. Science Fictions: How Fraud, Bias, Negligence and Hype Undermine the Search for Truth. New York: Henry Holt, 2020

[22] Husereau D, Drummond M, Augustovski F, *et al.*: Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated reporting guidance for health economic evaluations. *ValueHealth.* 2022; **25**(1): 3–9

[23] Magee B. Popper. London: Fontana, 1974

[24] Hume D. An Enquiry Concerning Human Understanding, 1748

[25] Russell B. The Problems of Philosophy, 1912

[26] Hansson S. *Defining Pseudoscience and Science* in Pigliucci M, Boudry M. Philosophy of Pseudoscience: Reconsidering the Demarcation Problem. Chicago: University of Chicago Press, 2013

[27] Lakatos I, Musgrave A Criticism and the Growth of Knowledge. Cambridge: Cambridge University Press, 1970

[28] McKenna S, Wilburn J. Patient Value: Its nature, measurement, and role in real world evidence studies and outcomes-based reimbursement. *J Med Econ*. 2018;21(5):474-80

[29] Langley P, McKenna S. Fundamental Measurement: The Need Fulfillment Quality of Life (N-QOL) Measure. *InnovPharm*. 2021;12(2):No. 6
https://pubs.lib.umn.edu/index.php/innovations/article/view/3798/2697

[30] Ladyman J. *Towards a Demarcation of Science from Pseudoscience* in Pigliucci M, Boudry M. Philosophy of Pseudoscience: Reconsidering the Demarcation Problem. Chicago: University of Chicago Press, 2013

[31] Moberger V. Bullshit, Pseudoscience and Pseudophilosophy. *Theoria*. 2020; 86:595-611

[32] Langley P. Peter Rabbit is a Badger in Disguise: Deconstructing the Belief System of the Institute for Clinical and Economic Review in Health Technology Assessment. *InovPharm*. 2021; 12(2): No.20
https://pubs.lib.umn.edu/index.php/innovations/article/view/3992/2855

[33] Rasch, G. *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.* Copenhagen: Nielsen & Lydiche, 1960.

[35] Frankfurt H. On Truth. Princeton: Princeton University Press, 2007

[36] Dawkins R. A Devil's Chaplain. New York: Houghton Mifflin, 2004

[37] Brazier J, Peasgood T, Mukuria C et al. The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. ValueHealth. 2022;25(4):482-91