MAIMON WORKING PAPER No. 9 JUNE 2024

ARE THE PRACTICE GUIDELINES OF THE INTERNATIONAL SOCIETY FOR PHARMACOECONOMICS AND OUTCOMES RESEARCH (ISPOR) FIT FOR PURPOSE?

Paul C. Langley, Ph.D. Adjunct Professor, Graduate Faculty, College of Pharmacy, University of Minnesota, Minneapolis, MN

Abstract

Practice guidelines in health technology assessment are of no interest if they to recognize and meet the standards of normal science and the standards of fundamental measurement. Over the past 20 years the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) has published a series of guidelines that fail to recognize these essential requirements. This dereliction is readily explained by the relativist position of the current belief system or meme in HTA. The HTA belief system, in assessing comparative product effectiveness, is focus on approximate information, rejecting hypothesis testing. This means that ISPOR accepts the central role of reference case modelling to support formulary decisions. ISPOR is not alone in this acceptance; a belief system that has been actively promoted by academic health economists and the National Institute for Health and Care Excellence (NICE) in the UK since the late 1990s. The foundation for this belief system is the assumption driven modeled simulation of the reference case. This rejects, explicitly or otherwise, both the standards for demarcation in normal science and the standards for fundamental measurement. Given this belief system, the concern must be that the ISPOR practice guidelines are of limited if any importance, if the premises for the guidelines are not fit for purpose. The argument presented here us that because of the neglect of the standards of normal science and fundamental measurement, ISPOR needs an independent review of all practice guidelines to propose acceptance, amendments or outright rejection.

INTRODUCTION

To propose a review of practice guidelines that have been in place for many years risks, not only treading on senior toes, but finding a review team that is not already firm believers in the HTA meme. Challenging foundation beliefs is an invitation to academic oblivion. The opposition will be fierce, irrespective of the merits, with perceptions of winners and losers. Careers can be destroyed; academic infighting puts the Inquisition to shame.

Yet there must be a review that ISPOR welcomes. Not that that it is just long overdue but that there a number of substantial concerns that, for ISPOR, blur the line between science and non-science [1]. In some cases, unfortunately, ISPOR practice guideline are clearly non-science or bunk. This is seen in the long advocacy by ISPOR of the quality adjusted life year (QALY) which, by the accepted long-recognized standards of fundamental measurement, is a mathematical impossibility [2]. This leads to the question of multiattribute scores such as the EQ-5D-3L preferences which are nothing more than composite ordinal scores or observations which, again, fail the standards of fundamental [3]. A further concern is the embrace of approximate information and reference case simulation modelling, as opposed to hypothesis testing. to support cost-

effectiveness claims [4]. The reference case not only fails the standards for fundamental measurement but also the standards for normal science, which brings us back to the issue of demarcation and the unique contribution of Rasch measurement [5] [6].

STANDARDS FOR PRACTICE GUIDELINES

Health technology assessment (HTA) is viewed as embracing techniques that are the equivalent to those mandated in the physical sciences and the more mature social science such as positive economics and education. This means that in developing practice guidelines there must be a commitment to the standards of normal science and the standards of fundamental measurement. Although unnecessary in the physical sciences where these criteria are instinctively accepted, a clear statement by ISPOR would support the commitment to a new start in HTA guidelines and underwrite their credibility. Of particular importance in HTA are value claims for patient reported outcomes (PROs). ISPOR should endorse the unique nature of Rasch rules as the necessary and sufficient conditions for transforming observations to measurement [7] . In the case of disease or target patient population specific value claims, the probabilistic Rasch model looks to the successful response to an item as a function of the ability of the patient and the difficulty of the item. This is not new; it was developed by Georg Rasch in the 1950s and is globally accepted as a cornerstone of measurement [8]. ISPOR needs to make clear the implications of meeting these standards: the role of value claims that are for single attributes with the measure having linear, interval and invariant properties. Accepting the Rasch model means abandoning multiattribute generic instruments and the QALY; focusing instead on patient centric disease specific instruments.

As will be detailed below, adopting Rasch measurement with the demise of multiattribute instruments and the QALY means abandoning the long standing ISPOR support for assumption driven simulation models with their imaginary cost-effectiveness claims and threshold cost-per-QALY cutoffs. Needless to say, with the focus on imaginary claims, these models fail the standards of normal science. Indeed, it should be remembered that Stevens classic categorization of nominal, ordinal, interval and ratio scales refer only to single attributes [9]

A RELATIVIST MIND SET

The extent to which ISPOR in its belief system deviates from the standards of normal science and fundamental measurement raises the question of whether the thought leaders in HTA see themselves as inhabiting a parallel HTA universe as evidenced by leading textbooks which lack any discussion of the standards of normal science and fundamental measurement [10] Are we dealing with cultural relativity with the belief system or pseudoscience of HTA having no claim to being less valid than any other. In the UK, as Wootton details, the so-called strong program takes the position that the content and organization of science can be explained sociologically; we cannot say that one belief system is superior to another[11]. In confronting reality, evidence is constructed never discovered. The success of the HTA research program is not judged by its ability to create new knowledge but on its ability to generate the support of an HTA community and those dependent on it. HTA, in this view, is a non-science that depends on rhetoric, persuasion and

authority. There is no such thing as the scientific method. Truth is consensus; a position, as Wootton notes, at odds with one that looks to superior evidence as a basis for changing views [10].

While a relativist HTA perspective is at variance to a commitment to the standards of normal science and fundamental measurement, it provides an explanation of why organizations such as ISPOR, with its well-attended conferences and student chapters, supported by academic departments, have held so strongly over 30 or more years. The result for government agencies and health allocation gatekeepers such as NICE in the UK, the PBAC in Australia and CVZ in the Netherlands is to create not discover evidence through the application of value or reference case imaginary simulation models [12] [13] [14]. Yet, to apply mathematically impossible constructs such as the QALY, together with the ongoing commitment to approximate information, as a proposed meaningful basis for health system decisions, seems irrational.

Following Dawkins, the HTA belief system can be seen as a meme: a self-replicating element of culture ,passed on by imitation (OED) [15]. An informational mind virus travelling through generations with significant transmission fidelity *with the mind as a plausible candidate for infection*. A replication package we call the HTA belief system, where transmission fidelity is ensured by practice guidelines and the associated commitment of graduate programs and mutually reinforcing ideas and applications. Dawkins extends this memetic concept of a mind virus to the question of whether or not recipients realize they have been infected; even to the extent of denial. For Dawkins, accepting a mind-virus means a conviction that seems immune to evidence or reason; a belief that is the stronger because of the rejection of evidence or reasoned argument and the acceptance of mysteries.

While the characterization of the HTA belief system as a meme or mind-virus may seem far-fetched, it points to a key issue: how does a meme transform to a new paradigm? How easy will it be for current 'false' beliefs to be abandoned? Deeply held beliefs, as the Wars of Religion in the 16th and 17th centuries demonstrated, are surprisingly resilient.

MULTIATTRIBUTE SCALES AND QUALITY ADJUSTED LIFE YEARS

While it may seem as a sad farewell to an old friend, the quality adjusted life year (QALY) had no real existence from it is initial conceptualization as a mathematical construct. While we might describe the utility or preference scores created by multiattribute instruments as composite scores, it seems odd to characterize them as ordinal scores because they are composite. In any event, the fact that they are by design not unidimensional, linear, interval and invariant measures means they cannot support multiplication and hence QALYs. The QALY is a simply a will o'the wisp.

The rejection of the QALY after 30 or more years with over 26,000 references on PubMed will take some explaining, even though as ISPOR has been advised on many occasions that the QALY fails [16]. But ISPOR still holds to it tenaciously, as seen in the latest response to charges of discrimination and the HR 485 Protecting Health Care for All Patients Act 2023 which is in the Senate for consideration [17]. Unfortunately for the ISPOR defenders the final nail pre-empting HR 485 has been put in the QALY coffin with the amendments to Section 504 of the Rehabilitation Act of 1973 which take effect on July 8, 2024. These amendments make clear that there should be no value claim entertained for medical treatment decisions by those that receive Federal financial

assistance from the Department of Health and Human Services (DHHS) if based on (i) stereotypes about individuals with disabilities, (ii) judgments that an individual with a disability will be a burden on others or (iii) belief that the life of an individual with a disability has less value than the life of a person with disability [18].

Judged against these prohibitions, community preference scores for health states, the foundation for multiattribute instruments, are unacceptable. These prohibitions do not deny the application of value claims to support approval for new therapy interventions and the circumstances for that approval. Issues such as the quality of care and cost containment are relevant but only if the measures applied to support approval and access are non-discriminatory. In terms of prohibition (i) above, the various multiattribute generic instruments are clearly stereotypical in attempting to apply a minimalist global health state description (e.g., five health dimensions and three problem levels for the EQ-5D-3L) to classify winners from losers. In respect of prohibition (ii) value claims based on these instruments are intended to measure the community valuation of the burden of disease, assuming that a single composite metric is sufficient to distinguish a more burdensome disease state from one which is closer to the defined perfect health state, ranking health states in the process (243 health states for the EQ-5D-3L and 3,125 health states for the EQ-5D-5L). Finally, in respect of prohibition (iii) these instruments yield negative scores which is interpreted as the community valuation of a health state as worse than death, with the implication that this lesser value denotes a life not worth living. A conclusion that is not, apparently, shared by patients who are assigned to that category.

Abandoning the QALY also means abandoning all multiattribute instruments. They fail the standards of fundamental measurement for the simple reason that they are multiattribute; while measurement, in science, is focused in the first instance on single attributes. These may be combined to create other measures (e.g., body mass index) but the EQ-5D-3L and its associates are not measures. This means that the endless debate over value sets, reconciling the EQ-5D-5L with the EQ-5D-3L are storms in a non-existent measurement teacup.

This lack of a measurement strategy is seen in the time tradeoff (TTO) scores that community members attach to health state bundles for the EQ-5D-3L/5L instruments. These scores (which are supposed to be constrained to 0 to 1) actually have both negative and positive values. These are composite scores. They fail the first step in Rasch modelling in capturing not a single attribute but multiple attributes. This means we cannot proceed to a unidimensional, linear, interval and invariant measure by application of Rasch transformation rules. They remain observations.

The importance of fundamental measurement in PRO construction is made clear by Bond et al: the primacy of data over the theoretical model often accompanied by a parallel regard for the primacy of observations over substantive theory (p. 281) [5]. The question raised by Rasch analysis is: *are we to give such data primacy over a conceptual mathematical model of what good data should look like?* The Rasch model's quality control rules construct measures from the data that fit the model; an instrument of construct validation. The invariance property of the probabilistic Rasch model means it is predictive and inferential. In contrast to the traditional approach of fitting a model to as much of the observed data as possible which is simply descriptive of those data. A description of data is not measurement.

There are many instances where, there in an attempt to make a silk purse out of a sow's ear data which have been assembled to create a disease specific instrument, Rasch measurement rules have been applied, *ex post facto*, to claim the instrument has been 'Rasched' by eliminating items [5] . The result is not measurement although journals (such as *Value in Health*) are happy to publish the analyses; Rasch measurement is created where a linear, interval and invariant unidimensional measure is created from data that fit the Rasch model. Hopefully, we will not have an ISPOR good practices guideline on how to Rasch your PRO instrument.

The false belief that an existing instrument can be 'Rasched' is of particular interest with the instrument vacuum created by the Section 504 prohibition amendments. Now that the DHHS has effectively outlawed generic multivariate instruments and the QALY, together with the ICER reference case models and recommendations, we have to recognize that with the exception of a handful of Rasch disease specific measures, the overwhelming majority of disease specific instruments that have been developed and applied for over 45 years are redundant. There is a vast information vacuum to fill.

THE BEGUILING REFERENCE CASE

The is a certain Lego-set like attraction in creating imaginary cost-effectiveness claims from assumption driven model simulations. They can certainly while away a wet Sunday afternoon in the Hamptons, but the end product meets neither the standards of normal science nor the standards for fundamental measurement. The end-product, cost per QALY claims and cost-per-QALY thresholds are meaningless. They are based on scales that are ordinal and QALYs that are mathematically impossible. Even so, The Institute for Clinical and Economic Review (ICER) maintains that the multiattribute ordinal scales such as these created by the EQ-5D-3L algorithm are believed by health economists to be actually ratio measures in disguise; this is a delusion [19] . If for no other reason ICER does not appreciate that a ratio measure must have a true zero while the multiattribute scales yield health states with negative scores or states worse than death. While this may have intriguing religious implications, the fact is they fail fundamental measurement or Rasch standards; measurement requires single not multiple, bundled together attributes. This puts ICER in the position of promoting an analytical framework which is not only an analytical dead end but one that not have been proposed in the first place. The attempt to construct a metric to underwrite resource allocation in healthcare has nothing to offer.

There is a further point of logic that those promoting the reference model overlook: from the fact that all past futures have resembled past pasts it does not follow that all future futures will resemble future pasts. We cannot prove the validity of inductive procedures. Add to this, for those who might be inclined to argue (falsely) in probabilistic terms, a review of the assumptions in all too many of these imaginary creations points to how thin the material is from which the assumptions are taken. Despite the earnest efforts of the academic assumption police who are charged with reviewing reference models and their claims for agencies such as NICE and the PBAC, the fact is that reference models can be infinitely variable and manipulated to come to a cost-effectiveness claim of the sponsor's choice. They are just marketing devices.

The recent publication of the CHEERS 2022 guidance for creating imaginary assumption driven modelled imaginary value claims is presumably to reinforce the ISPOR message; to ensure ongoing transmission fidelity of the HTA belief system [20] [21]. The intent of the guidance, endorsed by a number of leading journals, is to make submission of imaginary reference type claims easier. Once again, there is no reference to the standards of normal science or fundamental measurement with the QALY taking its accustomed center place. Reinforced by a checklist of good imaginary practices for modeled value claims, the pathway to publication is smoothed with more imaginary and false claims joining the thousands already cluttering the literature [22].

The beguiling reference case has what many see as a dark side: it can disallow as well as allow access to health care. While this has unfortunate eugenic implications, it is why the multiattribute instruments were developed in the first place; they are designed to ration care in a fixed budget environment. The eugenic implications of QALYs, notably for states worse than death, are recognized but the ethical implications ignored [23]. The fact that the community, usually in ignorance of the health state implications, provides preferences is not a basis for allowing or denying care.

As an aside: the rejection of these key scales in HTA is valid but misses the point: if they had been challenged in terms of fundamental measurement they would never have emerged. Instead, we have 30 wasted tears and thousands of peer-reviewed and published QALY claims, many of which are deliberately false in their support of a sponsor's product [24] [25] .

GOOD RESEARCH PRACTICES IN MODELING: ISPOR TASK FORCES

There are two questions that should be applied to models in HTA. First, is the model focused on evaluating the impact of a specific therapeutic intervention on a credible and evaluable single outcome that meets fundamental measurement standards and, second, do the inputs to the model meet standards for fundamental measurement. There is no benefit from models that fail to meet these standards. A model which is characterized as providing approximate information or is one that is a policy decision model should be rejected as failing to meet the demarcation criteria between science and non-science. In simple terms: all modeled claims for therapy impact should be presented as falsifiable and reported within a meaningful time frame.

Judged against these standards, the modeling good research practices of ISPOR are of limited interest because they fail to consider the evidentiary and assessment standards of normal science and the importance of single attribute linear, interval and linear measurement. There is no concept of what Popper has described as the evolution of objective knowledge [26]. The various value simulation models are an analytical dead end.

A continuing major feature of the ISPOR commitment to imaginary modeled claims is the focus on model validation as shown in the ISPOR-SMDM modeling good practices task force [27] [28]. Certainly, we need to review the structure and provisional assumptions that drive the model, but validation should not extend to *a set of methods for judging a model's accuracy in making relevant predictions*. The applicability of a model's predictions to a given therapeutic environment is judged by whether or not the model predictions are falsifiable, not by criteria to say a model is valid and therefore must be believed. Good research practices, unlike the ISPOR beliefs, are not focused on

imaginary non-evaluable value claims, justified by the internal validated structure of the reference case model and possible comparisons with other imaginary model structures.

Predictive validation or confirmation of model claims is a curiously outdated term. Popper claimed to have killed logical positivism in the 1920s. Rather than predictive validation as the most desirable type, Popper's theory of falsification assesses the meaning of statements in terms of their falsifiability, whereas the verification principle determined the meaning of a sentence in terms of confirmation. Popper presents falsification as the antidote to verification. The verificationist criterion of *meaning* should be abandoned and replaced by a different criterion: one which differentiates between scientific and non-scientific, questions and answers. This criterion, according to Popper, should be the ability to refute claims. Instead of looking for arguments in support of scientific positions, they should be subjected to constant attempts at refutation. The theory which withstands the most (or even all) attempts at refutation should be provisionally accepted. Importantly, models or theories should be preferred if they produce claims that have the greatest likelihood of being falsified. There must be systematic attempts to refute all value claims. This position is entirely absent in the ISPOR belief system.

Verification or confirmation avoids the question of falsification. The failure to consider falsification, to escape from logical positivism, has two important consequences for HTA modeling: first, it gives a basis for rejecting all models that fail to have empirically falsifiable unidimensional claims that are evaluable and reportable in a meaningful time frame and, second, it points toward the question of progress in science or the evolution of objective knowledge which must rest on conjecture and refutation. The Good Practices Task Force fails in its neglect of the scientific method: discovery has no value when the focus is on creating imaginary evidence. A position which ISPOR doggedly retains[29] .

Certainly, we can follow the somewhat tedious check list for assessing a reference model structure; the so-called internal validation. But this misses the key point: we must reject all models that fail to produce a falsifiable outcome; the poverty of modeling in HTA. Value assessments of medical technologies must not rest on imaginary cost-effectiveness analysis constructs. Models must focus on supporting single value claims. Model inputs must respect the standards of fundamental measurement. Claims that a model can be validated without supporting an empirically falsifiable outcome claim are wrong. Focusing solely on internal validation of a model to demonstrate the model strengths, is immaterial; the intent must be to create a high information content, falsifiable value claim. This should be supported by a protocol detailing how the modeled value claim is to be evaluated and reported. A requirement that is necessarily absent if the value is imaginary. HTA has spent too long in imaginary make-believe worlds for formulary decisions.

If the standards for falsification are to be met, all value claims, whether modelled or not, must be for single attributes with linear interval and invariant properties. In some circumstances, they may be in ratio terms with a true zero. Unless these standards are met, any analysis with traditional statistical techniques is out of the question.

MAPPING HEALTH UTILITY ESTIMATES

One of the two top priorities for ISPORs Vision 2020 Initiative was to propose good research practice guidelines for mapping health-state utilities for cost-effectiveness analysis [30]. The recommendations include selection of data sets, selection of the statistical model, assessment of model performance, reporting standards and application of results [31]. While these are, no doubt, laudable objectives the entire exercise falls flat because no one addressed the fundamental question: what are the required measurement standards for mapping? Or, more to the point, does mapping from a non-preference-based outcome measure to estimate health utilities or preferences make any sense whatsoever?

If we accept the need to map from a patient reported outcome measure to estimate health utilities means mapping from an ordinal scale to another ordinal scale we have a problem. The overwhelming majority of patient reported outcome instruments fail the standards for Rash measurement. Both scales fail Rasch standards for fundamental measurement. No one asked the question: is the source instrument one that meets Rasch standards with unidimensional, linear, interval and invariant properties? Nor did anyone as why the target of the mapping should be a composite instrument that also lacks measurement properties?

It is not a question of overlap between the concepts encompassed by patient reported outcome to the multiattribute preference-based measure, but of the ability to map from an existing ordinal scale to a hypothetical (to be constructed) ordinal scale. The short answer is that it is impossible because ordinal scales, in the case here of the patient reported outcome measure, do not support traditional statistical analysis. The question of fundamental measurement is not raised in the recommendations. This is perhaps not unexpected as, following ICER, health economists have no need to raise the question of fundamental measurement as all scales, both for multiattribute preferences and patient reported outcomes are unquestionably ratio scales (albeit in disguise).

Despite the number of mapping studies that have appeared, the failure to appreciate the constraints of fundamental meal measure means they are all fatally flawed; they have no scientific credibility. It seems pointless, in the case for example of the Wailoo et al task force guidance, for ISPOR to propose a mapping methodology that has as its focus a health state utility score that has no meaning as a measure [31]. Yet, this is entirely consistent with the ISPOR belief system.

DENIAL OF MEASUREMENT

It is clear from ISPOR special task force reports that the standards of fundamental measurement are of no interest. Indeed, it they were addressed they would be seen to be the single most important failure in HTA. This is not an overstatement: take away multiattribute instruments and the edifice collapses. The concept of a single metric, the QALY, ceases to exist. The reason for this has been addressed above; the QALY is not a measure. The failure is as simple as that. Recommendations for expanding the scope of the QALY such as the EQ-Health and Wellbeing Index (EQ-HWB) are analytically bankrupt and of no interest [32].

If the QALY is declared impossible there is not much left in the endless discissions within ISPOR for the valuation of medical technologies. Recommendations for ‘improvements’ , to propose more

novel elements of value in economic evaluations to capture more patient centric considerations amount to nothing if the QALY prop is taken away. Indeed, it is more than a prop. With the devotion to cost-effectiveness analysis the QALY is the core article of faith.

Ignoring fundamental measurement means abandoning serious measurement. Focusing on composite measures, both generic and disease specific can lead to misleading results and adversely affect clinical practice. Composite measures are not reliable with claims for no improvement contradicted by patients and clinicians [33]. Unfortunately, with the failure to consider fundamental measure, the ISPOR good practices also fail to consider the measurement behavior of composite multiattribute and patient reported outcome instruments. Multidimensionality of instruments defeats accurate measurement. A conclusion that was obvious in early instrumentation (e.g., thermometer). More to the point: including patient or community reported outcomes from instruments that lack the required measurement standards may not only lead to incorrect and potentially dangerous conclusions in clinical practice but make a mockery of the gold standard for therapy choice of assumption driven modeled simulation where there is no Rasch quality check on outcomes claims included in the model.

The unanswered question is whether this complete absence in any of the special task force reports over the years of a mention of fundamental measurement reflects genuine ignorance or a belief that the standards of fundamental measurement and reliability of composite measures are irrelevant to the creation of imaginary valuations of health care technologies. As ICER maintains: perhaps health economists truly believe, have confidence in, the composite ordinal multiattribute preference scale as a unidimensional ratio measure in disguise. Belief in a measure because it is impossible is not unsurprising. Quoting Tertullian, Dawkins asks if 'mystery is a virtue': *Certum est quia impossible est* (It is certain because it is impossible).

CONCLUSIONS

There seems little doubt that the current practice of HTA with ISPOR as the champion faces some difficult questions. Why has ISPOR and the many supporters in academic institutions denied the importance of the standards of normal science and fundamental measurement, training students to think in non-measurable cost-outcome and imaginary terms? Why do graduate courses in HTA ignore fundamental evidence and fundamental measurement when the Rasch model is critical to PROs? Are they consciously inhibiting a parallel universe where these standards are irrelevant? Or is it an *ex post facto* cover-up when the neglect of these standards and their fatal implications are becoming apparent to a wider audience? Are we to reject a salient feature of the scientific revolution, as made clear in the motto of the Royal Society (1662): *nullius in verba* or take nobody's word for it? This is precisely what ISPOR is asking us to do. A position that is replicated in the leading textbook in HTA with its commitment to the QALY and imaginary cost-effectiveness claims.

Judged by Rasch standards for PRO fundamental measurement, ISPOR has driven itself into an analytical dead end. For a decade or more it has promoted the reference case value model with no regard for the standards of normal science and fundamental measurement. There is a superficial attraction to creating framework for resource allocation in health care determined by community

preferences at the expense of patient and caregiver needs. The QALY as a universal metric takes center stage. Take the QALY away and the entire analytical edifice collapses. To reject the existing HTA meme or belief system in favor of a paradigm for HTA that recognizes the standards for normal science and fundamental measurement is the only way out. Given the stranglehold the current HTA belief system has, it is doubtful if there is the will for change. After all, it must be comfortable being in a non-science relativist parallel reality with the support of so many ISPOR members.

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

REFERENCES

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

[2] 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)

[3] McKenna S, Heaney A. Composite outcomes measurement in clinical research: the triumph of illusion over reality. *J Med Econ*. 2020;23(10)1196-1204

[4] Neumann PJ, Willke RJ, Garrison LP: A Health Economics Approach to US Value Assessment Frameworks – Introduction: An ISPOR Special Task Force Report (1). *Value Health.* 2018; **21**(2): 119–123.

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

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

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

[8] Rasch G. Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. Nielsen & Lydich, 1960.

[9] Stevens S. On the theory of scales of measurement. *Science*. 1946;103:677-80

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

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

[12] National Institute for Health and Care Excellence. Technology Appraisal Guidance. Health Technology Evaluation Manual. Feb 2022 https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/changes-to-health-technology-evaluation

[13] Pharmaceutical Benefits Advisory Committee. Guidelines for Preparing Submissions to the PBAC. September 2016 https://pbac.pbs.gov.au/

[14] National Healthcare Institute. Guideline for Economic Evaluations in Health Care. 2024 Version. https://english.zorginstituutnederland.nl/about-us/publications/reports/2024/01/16/guideline-for-economic-evaluations-in-healthcare

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

[16] Langley P, McKenna S. Fundamental Measurement and Quality Adjusted Life Years [letter to the Editor] *ValueHealth*. 2021;24(3):461

[17] Wilkie R, Pizzi L, Rand L. The Value of Quality Adjusted Life Years. *Value Health*. May 01, 2024 https://www.valueinhealthjournal.com/article/S1098-3015(24)02350-7/abstract

[18] Langley P. The United States Outlaws Multiattribute Scores and the QALY. *Maimon Working Paper* No. 7 May 2024 https://maimonresearch.com/wp-content/uploads/2024/05/MAIMON-WORKING-PAPER-No.-7-MAY-2024-V21.pdf

[19] 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. *InnovPharm*. 2021;12(2): Article 20 https://pubs.lib.umn.edu/index.php/innovations/article/view/3992/2855

[20] 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

[21] Husereau D, Drummond MN, Augustovski F: Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 2022 Explanation and Elaboration: A Report of the ISPOR CHEERS II Good Practices Task Force. *ValueHealth.* 2022; **25**(1): 10–31

[22] 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)

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

[24] Xie F, Zhou T: Industry sponsored bias in cost-effectiveness analysis: registry-based analysis. *BMJ.* 2022;377.

[25] 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)

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

[27] Caro J, Briggs A, Siebert U et al. Modeling Good Research Practices – Overview: A Report of the ISPOR-SMDM Modeling Good Practices Task Force – 1. *Med Des Making*. 2012l;32:667-677

[28] Eddy D, Hollingworth, Caro J et al. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Practices Task Force – 7. *ValueHealth*. 2012;15:843-50

[29] Garrison L, Neumann P, Wilkie R et al. A Health Economics Approach to US Value Assessment Frameworks – Summary and Recommendations of the ISPOR Special Task Force Report [7]. *ValueHealth*.2018;21:161-65

[30] Wolowacz S. New ISPOR Recommendations – Mapping Methods for Estimation of Health State Utility (Editorial). *ValueHealth*. 2017;20:28-29

[31] Wailoo A, Hernandez-Alava M, Manca A et al. Mapping to Estimate Health-State Utility from Non-Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task Force. *ValueHealth*. 2017;20:18-27

[32] Brazier J, Peasgood T, Mukuria O et al. The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing. *ValueHealt*h. 2022;25(4):482-91

[33] McKenna S, Heaney A, Wilburn J. Measurement of Patient Reported Outcomes 2: Are current measures failing us? *J Med Econ*. 2019;22(6):523-30