INNOVATIONS in pharmacy

Volume 7 | Number 2 Article 13

7-8-2016

He Ao Pohewa: The PHARMAC Prescription for Pharmacoeconomic Analysis in New Zealand and the Standards of Normal Science

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

He Ao Pohewa: The PHARMAC Prescription for Pharmacoeconomic Analysis in New Zealand and the Standards of Normal Science

Cover Page Footnote
None

He Ao Pohewa: The PHARMAC Prescription for Pharmacoeconomic Analysis in New Zealand and the Standards of Normal Science

Paul C Langley
College of Pharmacy, University of Minnesota, Minneapolis MN

Abstract
In common with a number of other single payer health systems, New Zealand has, through the Pharmaceutical Management Agency (PHARMAC), established guidelines for formulary submissions by pharmaceutical manufacturers. A question that is important to health system decision makers is to whether or not guidelines for economic evaluations in countries like New Zealand are consistent with the standards of normal science. Do the guidelines require those making the submission to put their claims in the form of testable hypotheses that can support falsification and replication? Are post-listing evaluations of these claims ever carried out? The purpose of this review is to consider whether the 2012 PHARMAC guidelines meet these standards. The assessment argues that the guidelines do not meet the standards of normal science. Instead, from this perspective, they are best characterized as supporting the creation of he ao pohewa (an imaginary world). There is no requirement in the guidelines for claims to be expressed as testable propositions; as hypotheses for expected impact that can be evaluated and the outcomes reported as part of ongoing disease area and therapeutic class reviews. There is no commitment to the discovery of new facts though falsification and replication. Our review concludes with suggestions for a reworking of the guidelines to meet the standards of normal science.

Keywords: cost-effectiveness modeling, credibility, imaginary worlds, scientific method, pseudoscience

Introduction

Guidelines for the submission of clinical and cost-effectiveness claims to support formulary listing and pricing, the *Prescription for Pharmacoeconomic Analysis (PFPA)* were first introduced in New Zealand by the Pharmaceutical Management Agency (PHARMAC) in 1999. They were subsequently reviewed in 2007 and updated with minor changes in 2012 (version 2.1) [1]. PHARMAC's statutory objective is: *'to secure for eligible people in need of pharmaceuticals, the best health outcomes that are reasonably achievable from pharmaceutical treatment and from within the funding provided* (under Section 47(a) of the New Zealand Public Health and Disabilities Act 2000). The purpose of the PFPA guidelines document is to provide an overview of the methods PHARMAC uses when preparing submissions including a required cost-utility analysis. As stated in the PFPA guidelines 'Documenting of this methodology aims to ensure that cost-utility analyses performed by (and for) PHARMAC measure costs, benefits, time preference and uncertainty in a similar fashion; hence enabling comparison between the cost-effectiveness of different interventions and ensuring that the results of analyses are meaningful for decision making'. PHARMAC's objective is to ensure, with funding provided, 'the best possible health outcomes' for the target New Zealand population. This is to be achieved by maximizing the total health gains possible.

Corresponding author: Paul C Langley, PhD
Adjunct Professor
College of Pharmacy, University of Minnesota
Email: langley@maimonresearch.com

The purpose of this assessment is to consider whether or not the PFPA guideline methodology can support the objective of maximizing health gains. The key question is whether or not the PFPA methodology is consistent with the standards of normal science? Do the PFPA guideless support hypothesis testing, falsification and replication? If not, is their advocacy of models or simulated claims an acceptable alternative and a valid input to formulary and pricing decisions? Or should the PFPA guideless be seen as supporting the creation of *he ao pohewa* (false or imaginary worlds) to justify investment decisions [2 3]. Are the guidelines to be considered as science or pseudoscience? [4]

Standards of Normal Science

The standards of normal science, as they have been accepted since the 17th century in the natural sciences and, at least in economics in the social sciences, is that the discovery of new facts can only proceed through a process of falsification and replication [5]. In the case of formulary decision making where a committee has to judge the relative merits of competing products or a new product versus the standard of care, the applicable criteria are the evaluation of competing testable claims for clinical impact, cost-effectiveness and budget impact in target treating populations. If this objective is to be achieved, claims submitted by manufacturers (and internal assessments carried out by, in this case, PHARMAC staff)

should be in a testable form. While there is no necessary requirement that an assessment be carried out prior to an initial formulary and pricing decision, there should be the prospect that such an assessment can be carried out and reported back to decision makers in a meaningful time frame. In practical terms this might be expressed in terms of an assessment protocol to accompany a submission with a feedback within two years.

The requirement for hypothesis testing is both obvious and inviolate. Claims for cost-effectiveness should not be taken at face value. Judging a theory by the 'reasonableness' of its assumptions has long been rejected. Popper in his seminal contributions to the philosophy of science, points to the fact that claims must be capable of falsification; indeed they should be framed in such a way that makes falsification likely [6]. Commitment to falsification forces us to reconsider our models and the assumptions built into those models.

At the same time there is the need for replication. This has been brought home in the last few years by attempts to replicate clinical trial and laboratory claims. There has been increasing concern expressed over the presence of repetitive flaws and the need for guidelines to improve experimental reproducibility. As noted in a recent editorial in *Nature* applicants to the National Institutes of Health (NIH) are now required 'to explain the scientific premise behind their proposals and defend the quality of their experimental design' [7]. More recently, Camerer et al in their evaluation of laboratory experiments in economics find, of the 18 studies considered, an effect size in the same direction in only 11 replications with on average a replicated effect size of 66% of the original [8]. As the authors note: the deepest trust in scientific knowledge comes from the 'ability to replicate empirical findings' although rarely carried out in the social sciences.

The Pharmacy PFPA Guidelines

Under the PFPA guidelines the recommended technique to support investment decisions in pharmaceuticals when a significant funding proposal is involved is cost-utility analysis. Otherwise, an internal evaluation occurs. This has to take account of a pragmatic public policy/purchasing environment with constrained analytical capacity. There is, for PHARMAC, an inevitable trade-off between precision and timeliness of cost-utility analyses (leading to a threefold classification of preliminary, indicative and detailed cost-utility analyses). The PFPA guidelines apply to all classifications.

The cost-utility assessment is from the perspective of the funder (PHARMAC) with regard to the PHARMAC decision criteria. If there are anticipated appreciable impacts on non-healthcare sectors these should be assessed qualitatively. The target population, and possible sub-groups are those most likely to receive treatment. Sub-groups should be identified if there are anticipated significant differences in treatment effect. Comparators for the analysis should be the funded treatment that most prescribers or clinician's would replace. Alternatively, it could be the treatment given to the largest number of patients if this differs from the treatment most prescribers would replace. If treatment regimens differ substantially through New Zealand, then a range of comparators should be identified.

The preferred relative treatment effect claims are from randomized clinical trials (RCTs) and meta-analyses. Grades of evidence should be assigned with an assessment of external and internal validity. The methodology, limitations and any possible bias from extrapolating from clinical trials in the long-term (and generalizing results to a New Zealand setting) should be described clearly.

Models or simulations should be transparent and avoid undue complexity. In the case of discrete period events such as acute interventions without recurrence, a simple decision tree is recommended. Where the model is tracking the natural cause of a disease Markov models are recommended. In the majority of models, a lifetime horizon capturing all major clinical and economic outcomes, is required. If possible, adherence should be included in the modeled claim. The health benefits from interventions should be measured using QALYs. Options to QALYs are disability adjusted life years (DALYs) and healthy year equivalents (HYEs). The PHARMAC support for QALYs is that they are simple to calculate, have face validity and enable a cost-utility analysis to be performed. The recommended instrument for describing health states is the New Zealand EQ-5D Tariff 2 [9]. Where appropriate, costs should be estimated for (i) community and hospital pharmaceuticals; (ii) diagnostic related groups (DRGs) for inpatients; (iii) outpatients (e.g., healthcare professional costs); and (iv) direct patient health care (e.g. general practitioner visits). Costs incurred by patients living longer as well as indirect costs should not be included. The benefits and costs included in a simulation should be discounted at a rate of 3.5% (with rates of 0% and 5% in sensitivity analyses).

The results of the modeled or simulated cost-utility analysis should be reported as incremental utility cost ratios (IUCRs); the inverse of cost-per-QALY. These are interpreted as the opportunity cost of investment decisions under a fixed budget and are expressed as QALYs per $1 million invested. The IUCR should be presented as a point estimate as well as the plausible range over which QALYS per cost are expected to vary. To capture parameter uncertainty, univariate and multivariate sensitivity analysis together with probabilistic sensitivity analysis may be necessary. If there is a concern

with structural uncertainty, repeated analysis using alternative structures should be employed.

The Contribution of Simulations

The PFPA case for recommending cost-utility analysis in submissions is that it is practical, enables comparisons and supports the prioritizing of pharmaceuticals to support investment decisions. The PFPA document explicitly rejects cost-effectiveness thresholds. The PFPA case is that given the range of criteria that contribute to a formulary decision and the requirement that spending on pharmaceuticals is kept within a fixed budget, thresholds are inappropriate. As budgets may vary from year-to-year, there is no threshold below which a pharmaceutical is considered cost-effective.

With a commitment to lifetime (or at least long-term) cost-utility models to support interventions in non-recurrent or acute disease states, claims made for product impact are immune to failure. There is no way in which these claims can be evaluated empirically. At the same time, the proposed measures for reporting the results of the cost-utility simulation are, to all intents and purposes, also immune to failure. There is no possibility, unless interventions are reported for short-term, acute intervention decision models, that any of the claims made could ever be tracked and reviewed. The requirement for discounted QALYs, costs and savings makes the process of evaluation that more difficult.

If the standards of normal science are the benchmark, then PHARMAC's decision framework is clearly at variance. The required analysis fails the standards for hypothesis testing, falsification and replication. The focus on the lifetime modeling of disease states and the recommendation for QALYs as the preferred measure of health status effectively ensure that any post-market entry reassessment of the claimed benefits and costs of an intervention is impossible. In practical terms, therefore, it is impossible for PHARMAC to claim that, for a given budget period, their pharmaceutical investment decisions are consistent with securing 'the best health outcomes' with the funding provided. It is impossible for PHARMAC to look back over the past 15 years and review the long-term (or at least mid-term) impact of previous investment decisions against claims made (and accepted) for costs and benefits.

If we put the standards of normal science to one side, is it possible to justify the PHARMAC acceptance of modeled and simulated cost-utility claims as a meaningful approach to investment decisions? One possible argument could be that New Zealand, a small country with limited resources, is not in a position to set up standards that are at variance with those currently in place in other single payer developed economies such as Australia, Canada, Ireland and the UK [10 11 12 13]. In the UK, for example, the National Institute for Health and Care Excellence (NICE) has a staff of over 550 and an annual budget of over £60 million [14]. In the case of submissions to NICE, where the reference case is central to the process of review, the application of thresholds is best seen as a basis for price negotiations [15]. This is understandable given that the reference case cost-per-QALY claims are either untestable or immune to falsification. Indeed, it seems clear that it was never intended that claims could be evaluated. Matching imaginary claims to an arbitrary threshold of £20,000 or (in designated cases) £30,000 per QALY by modifying either the imaginary simulation itself or the unit price of the product is well accepted.

Arguing for a second best methodology is not consistent with the emphasis placed by PHARMAC on models and simulations. There is no recognition that their methodology may be considered second best. Rather, the case made is that cost-utility models or simulations are the 'global' standard in health care decisions. PHARMACs position is that by imposing a common methodology, analyses are undertaken in a similar fashion, enabling comparisons across different interventions. They are supported in this by practice standards issued by professional groups such as the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) [16 17].

The claim for a 'global' standard raises the possibility that, unwittingly or not, PHARMAC in common with agencies in other single payer systems and the US has adopted a relativist position [18 19 20]. Rather than subscribing to the position that the standards of normal science are the only standards to apply in health care decisions, the relativist position is that all perspectives are equally valid. In their advocacy of the equivalence or symmetry principle health care decisions are to be understood sociologically. No one body of evidence is superior to another. The success of a scientific research program rests not on its ability to generate new knowledge but on its ability to mobilize the support of this community. Decisions are based on the constructed evidence of models and simulations. Evidence is never discovered. Instead of coming to grips with reality this 'science' is about rhetoric, persuasion and authority. Truth is consensus.

A major objection to the relativist position is that competing evidence can be constructed and simulations can generate opposing claims. If the simulated input conditions and the simulated core mechanism correspond to 'reality', the sufficient condition character of the simulation assures us that the output is necessarily entailed. [21]. But simulations can be challenged; simulations can fail. Competing simulations can support alternative input conditions, core mechanisms and output predictions. Simulations can be 'reverse engineered' to generate competing claims yet still remain within 'agreed' guidelines. Any attempt to argue for the

superiority of one simulation over another in its correspondence to 'reality', in the absence of a commitment to a comparative experimental assessment of claims generated, simply lacks credibility.

These conclusions hold irrespective of how much we attempt to build up the appearance of being scientifically rigorous in the validation of models and simulations [22]. We can claim that they 'adequately reflect reality', we can apply deterministic and probabilistic sensitivity analyses, we can produce ICER cloud diagrams, we can apply thresholds and we can even introduce an analysis of the value of perfect information. Nevertheless, the fundamental objection still applies: there are no testable predictions and, as such, simulations for health care decisions fail the standards of normal science. It seems pointless to argue that the standards set by PHARMAC are acceptable because there is a community consensus, shared by NICE and other single payer systems together with the 'community of pharmacoeconomists', that models or simulations are acceptable as 'projections', even though the projected claims are patently untestable and, in all too many cases, we never intended to be testable. This misses the point entirely: claims for the comparative effectiveness of competing pharmaceutical products should meet the standards for evaluation, falsification and replication. Otherwise they are simply pseudoscience; they do not meet the standards of normal science, they share the platform with intelligent design rather than natural selection. To repeat the point made earlier: we have no idea if the claims are right or if they are wrong. To argue for a relativist position that truth is consensus is to put to one side the potential for discovering new facts and the role of evidence in clinical decision making.

The PHARMAC commitment to constructed evidence, to the advocacy of imaginary worlds, stands in contrast to recently introduced guidelines for formulary evaluation released by the College of Pharmacy, University of Minnesota [23]. These guideless reject untestable modeled or simulated claims, emphasizing the importance of testable claims, experimentation and replication. A key feature is the requirement for a claims assessment protocol to accompany formulary submissions. Manufacturers are asked to underwrite value claims assessment in a timeframe relevant to formulary decisions.

Conclusions

In the absence of a requirement for testable claims to accompany submissions for formulary listing and pricing, PHARMAC is in no position to argue that these submissions contribute materially to the objective of achieving, within budget constraints, the best health outcomes for target populations in New Zealand. Unless claims for outcomes expressed in cost-utility terms can be evaluated in a timeframe that is meaningful for decision makers, PHARMAC has to recognize that the claims presented are immune to failure. Unfortunately, QALYs are not reported on a regular basis as part of administrative health records or electronic medical records in health care systems that advocate cost-per-QALY claims.

A commitment to improving patient outcomes, and the increasing focus on the information needs and potential benefits of precision medicine, underscore the role of feedback for clinical and cost-effectiveness claims. Relying on simulated cost-utility models with a lifetime horizon to support investment decisions in pharmaceuticals is unacceptable. Rather than following the consensus 'pharmacoeconomic' view in modeling for health care decisions, PHARMAC should consider stepping back and focusing on the components of product claims that are evaluable in the short term. This does not exclude utility measures as a potential outcome or short term ICUR claims. The key point is that a commitment to short term evaluable claims commits PHARMAC to supporting the standards of normal science and a commitment to evidence based formulary decisions.

References

[1] New Zealand. Pharmaceutical Management Agency. Prescription for Pharmacoeconomic Analysis: Methods for cost-utility analysis (v 2.1). PHARMAC, 2012. https://www.pharmac.govt.nz/assets/pfpa-final.pdf

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

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

[4] Pigliucci M. Nonsense on stilts: How to tell science from bunk. Chicago: niversity of Chuicago Press, 2010.

[5] Wootton D. The Invention of Science. New York: Harper Collins, 2015.

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

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

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

[9] Devlin N, Williams A. Valuing quality of life: results for New Zealand health professionals. *N Z Med J*. 1999;112(1083):68-71.

[10] Australian Government. Department of Health. *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (version 4.4).* Canberra: July 2013.

[11] Canadian Agency for Drugs and Technologies in Health. *Guidelines for the Economic Evaluation of Health Technologies: Canada* (3rd ed.). Ottawa: CADTH, 2006.

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

[13] National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE, April 2013 www.

[14] National Institute for Health and Care Excellence. Annual Report and Accounts 2013/14, NICE, July 2014

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

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

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

[18] Langley PC. Sunlit Uplands: The genius of the NICE reference case. *INNOVATIONS in Pharmacy*. 2016;7(2): Article 12

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

[20] Langley PC. Modeling imaginary worlds: Version 4 of the AMCP Format for Formulary Submissions. *INNOVATIONS in Pharmacy*. 2016;7(2): Article 11.

[21] Grim P, Rosenberger R, Eason RE et al. How simulations fail. *Synthese*. 2011 DOI 10 1007/s11229-011-9976-7

[22] Ellis G Silk J. Defend the integrity of physics. *Nature*. 2014 (516): 321323; see also A Frank and M Gleiser. A crisis at the edge of physics. *New York Times*, 7 June, 2015.

[23] Langley PC. Nullius in verba. The university of Minnesota Social and Administrative Pharmacy Program proposed Guidelines for Formulary Evaluations. *INNOVATIONS in Pharmacy*. 2016.7(2): Article 11.