

1-10-2017

## Denkbeeldige Wereld: The New Dutch Guidelines for Economic Evaluations in Healthcare

Paul C. Langley

*University of Minnesota*, [langley@maimonresearch.com](mailto:langley@maimonresearch.com)

Follow this and additional works at: <http://pubs.lib.umn.edu/innovations>

### Recommended Citation

Langley PC. Denkbeeldige Wereld: The New Dutch Guidelines for Economic Evaluations in Healthcare. *Inov Pharm*. 2017;8(1): Article 4. <http://pubs.lib.umn.edu/innovations/vol8/iss1/4>

*INNOVATIONS in pharmacy* is published by the University of Minnesota Libraries Publishing.

## Denkbeeldige Wereld: The New Dutch Guidelines for Economic Evaluations in Healthcare

Paul C Langley, PhD

Adjunct Professor, College of Pharmacy, University of Minnesota

### Abstract

*In 2016 the Dutch National Health Care Institute (Zorginstituut Nederland) published a new guidance for economic evaluations in healthcare to support reimbursement decisions. These Guidelines update and replace three previously published guidelines covering pharmacoeconomic evaluation, outcomes research and costing. The purpose of this commentary is to consider the merits of these new Guidelines from the perspective of modeled claims which meet the standards of normal science: credibility, evaluation and replication in the treatment of target patient populations. In evaluating the merits of the Guidelines the focus will be on the requirement for submissions to follow reference case standards where lifetime-cost-per-QALY claims are the preferred outcome measure. The assessment points out that in adhering to a reference case standard, the Dutch Guidelines, in common with those in the UK, Ireland and New Zealand, fail to address the fundamental question of claims assessment. Rather, in relying upon the reference case imaginary world (denkbeeldige wereld) to inform decision makers, the possibility of evaluating claims and generating feedback to decision makers on comparative effectiveness is put to one side. We have no idea as to whether the claims are right or even if they are wrong. Hopefully, future versions of the guidelines will address this issue and focus on a rigorous program of claims assessment.*

**Keywords:** Dutch Guidelines, economic evaluations, imaginary worlds, simulations, evaluable claims

### Introduction

Over the past 25 years, recommendations and standards for assessing the merits of competing health care interventions have focused on informing decision makers through the construction of imaginary worlds (denkbeeldige wereld). The new Dutch guidelines for economic evaluations in healthcare, published in 2016, follow in this tradition<sup>1</sup>. In common with countries such as the UK, Ireland and New Zealand, together with the proposed EUnetHTA guidelines for the European Union, the focus is on a reference case; a construct that mandates a lifetime perspective with claims for competing interventions expressed in cost-per-quality adjusted life years (QALYs)<sup>2 3 4 5</sup>.

Previous publications and formulary evaluation commentaries in this series have made clear that in putting to one side a commitment to the standards of normal science, where modeled claims or hypotheses are credible, evaluable and replicable, decision makers in health care have a limited evidence base for formulary decisions<sup>6 7</sup>. Rather than putting claims for competing or new interventions in a framework that supports evaluation and feedback in a meaningful time frame, groups

such as the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and, in the UK, the National Institute for Health and Care Excellence (NICE) have supported reference case cost-per-QALY modeling as the recommended submission standard<sup>8 9</sup>. A lifetime cost-per-QALY model is not designed to generate evaluable claims. It is a construct that is defended by its sufficient correspondence to reality. Validation focuses on the core model and its assumptions. Whether or not the model can support evaluable claims and whether or not these claims could ever be evaluated is irrelevant.

In rejecting the standards of normal science, advocates of models that are intended to 'inform' decision makers in health care systems (whatever that means) rather than establish a practical research program, effectively put to one side a commitment to standards that have been in place since the seventeenth century in favor of what may be described as pseudoscience: intelligent design rather than natural selection<sup>10</sup>. In an effort to avoid this characterization, guidelines have been proposed by the Program in Social and Administrative Pharmacy at the University of Minnesota that put to one side imaginary constructs in favor of credible, evaluable and replicable claims; claims which apply equally well to clinical outcomes as well as those for comparative cost-effectiveness and budget impact<sup>11</sup>. Formulary submissions are supported by protocols to detail how the claims are to be evaluated and reported.

The purpose of this commentary is to consider the merits of the new Dutch Guidelines from the perspective of modeled claims which meet standards of normal science: supporting

**Corresponding author:** Paul C Langley, PhD

Adjunct Professor, College of Pharmacy

University of Minnesota

Director, Maimon Research LLC

5061 North Apache Hills Trail, Tucson, AZ 85750

Tel: 520-577-0436; Email: [langley@maimonresearch.com](mailto:langley@maimonresearch.com)

Web: [www.maimonresearch.net](http://www.maimonresearch.net)

credibility, evaluation and replication in claims for treatment effect. In evaluating the merits of the new Guidelines the focus will be on the requirement for submissions to follow strictly reference case standards where lifetime-cost-per-QALY claims are the standard outcome measure. Given the limitations inherent in the reference case paradigm, recommendations are made for reformulation of the Guideline so that claims assessments conform to the standards of normal science.

### Structure of the New Guideline

The purpose of the new Guideline is to provide a framework to assess whether there is sufficient evidence to support reimbursement for a drug. The new Guideline comprises five core sections. These are:

- Framework of the Economic Evaluation
- Method: Analytic approach
- Input Data
- Reporting
- Budget impact analysis

### Framework of the Economic Evaluation

Before an economic evaluation is undertaken an evaluation framework and design should be agreed. The framework must encompass the aspects described below in reporting the evaluation with a justification for the choices made. The elements of the framework are:

- Objectives: a statement of the decision making problem the evaluation is to solve
- Users: the future users of the evaluation to include the National Health Care Institute, health insurers and health facilities
- Perspective: to include all relevant societal costs and benefits
- Research question: to be formulated in accordance with PICOT criteria: (i) Patient or target population; (ii) intervention to be assessed; (iii) comparator or control; (iv) relevant outcome measure(s)
- Relevant time span for measuring effects and costs

In the case of the control or comparator therapy this should be the standard of care in usual practice and must reflect most recent national or international guidelines or standards, to include palliative or supportive care. As far as outcome measures are concerned, QALYs are the standard measure. This does not mean that other outcomes such a life years lived or patient reported outcomes (PRO) are put to one side; merely that they are an adjunct to the core claims expressed as cost-per-QALY. For the Guideline it is 'imperative' to follow

the reference case (see below) which means that the perspective of the economic evaluation is societal, taking into account all relevant societal costs and benefits irrespective of who bears them or who benefits from them.

### Guideline Reference Standard

The principal objective of the new Guideline is to improve the comparability of economic evaluations in the Netherlands. This is accomplished by establishing a reference case which sets the minimum standard. While there may be some deviations from the reference case, alternative approaches cannot replace the required standards. These are:

- Submissions should adopt a societal perspective
- Submissions should focus on a Dutch population
- Interventions should be compared to Dutch standard or usual care
- Assessments should take a lifetime perspective
- Required analytical framework should be cost utility analysis
- Costs and effects should be discounted at 4% and 1% respectively
- Uncertainty and sensitivity should be evaluated through univariate, probabilistic sensitivity and scenario analysis
- Costs should include those appropriate to the healthcare sector, patient and family and other sectors
- Where appropriate productivity losses should be captured using the friction cost method
- Wherever possible reference prices should use the Manual for Cost Research
- Outcomes should be expressed as QALYs utilizing at least the EQ-5D-5L with Dutch valuation and, whenever relevant, life years gained<sup>12</sup>
- Results should be expressed as (i) total costs and effects; (ii) incremental costs and effects; and (iii) incremental cost-effectiveness ratios
- Uncertainty and sensitivity results should be captured by (i) univariate sensitivity analysis with a Tornado diagram and table; (ii) scenario analysis as a table; and (iii) probabilistic sensitivity analysis as a CE-plane and cost-effectiveness acceptability curve

### Decision Framework and Model Validation

A decision framework is central to the analytic approach mandated in the reference standard. While a developer may attempt to justify an empiric approach where all relevant costs and benefits are collected within one clinical study, the model-approach is clearly the preferred option. The choice of decision framework will be influenced by the research question and medical decision. The appropriate technique has to be justified

in the choice of decision trees, state-transition models, discrete event simulations or dynamic transmission models.

Once a decision framework has been selected, outcome and cost data are required to be discounted if the time frame is greater than 12 months. The next steps are: (i) to determine the degree of uncertainty for costs, outcomes and cost effectiveness ratios; and (ii), and to quantify the consequences of the uncertainty and the value of additional research to reduce uncertainty. The degree of uncertainty can be captured by uncertainty analyses as detailed in the reference case while the value of additional research can be assessed through value of information analysis.

The Guideline requires prior validation of the decision model. This is to ensure that the model results are 'usable, reliable and credible'. The validation should encompass the conceptual model, the input data, the software to support the model and the model outcomes. Optimal reporting on the validation process is achieved by: (i) using a purpose-designed checklist and (ii) by providing a systematic representation of the validation steps. The example given of a tool to achieve these objectives is the AdViSHE toolkit<sup>13</sup>. The limitations of this toolkit have been explored in a recent commentary in this series<sup>14</sup>.

#### **Model Input Data and Reporting**

Where an economic evaluation is undertaken in the context of a randomized clinical trial (RCT) or observational study, the effectiveness input data are derived from the study possibly supplemented by other relevant publications. For a model-based study the input data should be grounded on all relevant studies that meet required standards. In both cases literature searches and systematic reviews should conform to accepted standards, where search strategies, the ranking of individual studies and outcomes are documented together with reports on individual studies. Where necessary indirect comparisons should be undertaken to generate comparative effect claims.

All categories of societal costs should be identified. These should follow the guidance set down in the Manual for Cost Research and the reference prices for common units. Prices and volume of all cost components should be reported separately.

Quality of life should be consistently measured with the EQ-5D-5L with Dutch reference values. Alternative measures can be added to the reference case. Even if the EQ-5D-5L is considered to lack sensitivity in a target population it should still be used. If valuations of quality of life are derived from the literature the submission should report: the questionnaire or valuation method, nationality of respondents and patient or societal valuations.

#### **Budget Impact Analysis**

The budget impact analysis (BIA) should follow the ISPOR standards to estimate the difference in expected expenditure between a reference scenario and a scenario in which the new or optimized intervention is accepted and disseminated. Key elements are: (i) perspective, which in the Dutch case is the national government; (ii) the time horizon where a 3-year minimum is recommended; (iii) the dynamics of implementation; (iv) the items of expenditure, volume measurement and valuation; and (v) reporting by budget period.

#### **Further Applications**

The Guideline also considers additional applications outside of those for the economic evaluation of pharmaceuticals. Consideration is given to the economic evaluation of: (i) prevention interventions; (ii) diagnostics; (iii) medical devices; (iv) long-term care; and (v) forensics.

#### **Overview**

The intent of the Guideline is to ensure, through a prior validation of the reference model, that the decision model results are 'usable, reliable and credible'. At the same time, there is no requirement in the Guideline that the reference model results are 'credible, evaluable and replicable' in target patient populations. This is not an oversight. It is clear from the focus on the reference model standards, with the default standard of a lifetime model with outcomes expressed as incremental discounted cost-per-QALYs, that there is no intention that the claims generated by the model are to be evaluated. Rather the modeled or simulated claims is to be considered a 'sufficient' representation of the expected societal costs and benefits generated by competing pharmaceutical products. As long as the model meets the standards required then, in judging it sufficient, the outcomes claimed necessarily follow.

#### **Constructed Evidence**

Unfortunately, in the absence of claims that are evaluable and replicable, there is no assurance that they are 'usable, reliable and credible'. To argue that they meet these standards, in the absence of any empirical evidence to support the claims, is to adopt a relativist position. For a relativist evidence is never discovered, only constructed within a particular social community<sup>15</sup>. In a community of health economists that accept a reference case paradigm to support claims for competing pharmaceutical products they would reject any arguments that one body of evidence is superior to another. A research program is not seen as one that generates new knowledge through claims evaluation and replication but one that is judged on its ability to persuade and mobilize community support for invented facts. Such a research

program puts to one side any notion of the progress of science, of the process through which new evidence overturns consensus views, in favor of rhetoric and authority. Reference case modeling is seen as the 'gold standard'.

#### ***Evaluating Credible Claims***

Rejecting the reference case does not mean that claims expressed in quality of life terms using, for example, the EQ-5D-5L would be rejected out of hand because these data are not collected on a regular basis from patient encounters. In the Minnesota guidelines (and earlier guidelines proposed for the WellPoint (now Anthem) health system, the proposal was that when a manufacturer submitted an economic evaluation to support formulary listing, it should be accompanied by a study protocol that detailed how the claims for comparative product performance were to be evaluated and reported back to decision makers. The timeframe was assumed to be relatively short (under two years) with the product provisionally accepted for formulary listing until evidence for the validation of clinical and cost-effectiveness claims could be assessed.

#### ***Adherence and Persistence***

A puzzling feature of lifetime reference model claims is the neglect of adherence and persistence behavior. The new Guideline is no exception. Nowhere in the instructions for constructing the imaginary reference case world is there any guidance on including compliance behavior. This appears an odd oversight as the Guideline is explicit in the requirement for a default lifetime model. If this implies that the hypothetical patient cohort are persistent with therapy over their lifetime, this flies in the face of decades of accumulated evidence which shows that by the end of two years from an index prescriptions, probably less than one third of patients are persistent with therapy with an even smaller proportion maintaining an adherence at a clinically meaningful level. If this is the case, then to model competing therapies assuming full compliance over the lifetime of the patient cohort would seem pointless.

#### ***Lifetime Costs***

A further puzzling feature is the assumption that current costs (corrected for discounting) remain unchanged over the lifetime of the hypothetical patient cohort. Again, this seems an odd assumption. In the US, for example, there is ample evidence for pricing strategies by manufacturers for ongoing semi-annual and annual price increases both over the patent of a drug and beyond. The price increases being accompanied by co-payment waivers, coupons and other discounts to maintain market share. In the case of disease modifying treatments (DMTs) in multiple sclerosis a recent study of the trend in annualized drug costs for nine DMTs from 1993 to 2014 found changes ranging from 7.9% to 35.7%<sup>16</sup>. Four of

the DMTs had annualized cost increases greater than 20% and four in the range 13% to 16.8%. While these annualized changes were two to three times bigger than in other countries, the potential for annualized price increases together with possible price increases in direct medical costs should, presumably, be factored into reference case models. In the Dutch case, anticipated price increases should presumably be extended to other societal costs elements captured by the model.

#### ***Pipeline Competitors***

It is unlikely, over the lifetime of a patient cohort, that there will be no therapies entering the market place to compete with and replace existing medications. In the reference case model this is not the case. Patients are assumed to remain with the indicated drug over their lifetime. Again, this flies in the face of evidence for drug turnover in target populations where patients are switched to new therapies. This switching may reflect a lack of response to the index drug in the treatment arm or may involve moving to a combination therapy. Again, this appears an odd assumption but one that is, unfortunately, driven by the focus on constructing evidence to establish credibility rather than on a more practical perspective of evaluating claims for feedback to physicians and formulary committees.

#### ***Next Generation Sequencing***

The likelihood of competitor therapies and therapy combinations is also made more likely by the introduction of next generation sequencing (NGS) where assay platforms will recommend linking sub-groups of patients defined by mutation clusters to monotherapy or combination therapies<sup>17</sup><sup>18</sup>. An obvious application is in late stage cancer but there will be applications earlier in the treatment pathway. In these scenarios there will be a premium place on tracking evaluable claims and reporting in real time to clinicians and health system decision makers on clinical outcomes and resource utilization. If a disease area or target tumor group is characterized by a distribution of patients by mutation cluster than a simplistic reference case model is hardly a viable basis for therapy choices when multiple pathways are involved and patients are individually selected for an assay driven intervention.

#### ***Redrafting the Dutch Guideline***

The claim that the new Dutch Guideline represents a singular improvement ('From Good to Better') over the previous guidelines is debatable<sup>19</sup>. Clearly, if health decision makers in the Netherlands believe that the construction of reference case imaginary worlds provide a viable constructed evidence base to support comparative assessments then that is the end of the debate. Any criticisms will no doubt be put to one side

even though, as noted above, to claim that the reference model is a sufficient representation of the decision environment seems odd when it makes no accommodation for adherence, persistence, new therapies and potential price increases. Perhaps it should be made explicit that in accepting the reference case paradigm as the standard for economic evaluations any the claims presented are immune to failure. With the reference case as the 'gold standard' it should be admitted that it is impossible and was never considered feasible or necessary to attempt to evaluate and even replicate competing product claims.

If claims are to meet scientific standard for evaluation and replication, then there may have to be a reconsideration of the scope of the new Guideline. In attempting to take a societal perspective that encompasses all relevant social costs and benefits irrespective of who bears the cost or who benefits, the guideline may be overly ambitious. A more pertinent approach may be to ask what claims can be usefully evaluated and reported within a 2-year time frame? Should these claims be restricted to the healthcare system? It seems absurd to create a decision framework that attempts to encompass all relevant social costs and benefits when there are no data, apart from those generated by an evaluation protocol, which would support a credible evaluation. On the other hand, it might be possible, within an observational study protocol to evaluate claims for patient and family costs, individual absenteeism and presenteeism costs and even costs incurred outside the healthcare system by municipal services. Claims made should recognize data limitations as well as the potential costs of data collection to support claims evaluation. If data are limited, then observational prospective studies may offer the most appropriate avenue for claims validation.

Redrafting the Dutch Guideline would also give an opportunity to take explicit account of anticipated patterns of adherence

## References

<sup>1</sup> Zorginstituut Nederland. Guideline for economic evaluation in healthcare. Diemen: 2016

<sup>2</sup> Langley PC. Sunlit uplands: the genius of the NICE reference case *Inov Pharm.* 2016;7(2): No.12

<sup>3</sup> Langley PC. Langley PC. Na domhain shamhlaíochta: formulary submission guidelines in Ireland and the standards of normal science. *Curr Med Res Opin.* 2016;32(5):1641-1644

<sup>4</sup> Langley PC. He ao pohewa: The PHARMAC Prescription for Pharmacoeconomic Analysis in New Zealand and the standards of normal science. *Inov Pharm.* 2016;7(2): No. 13

<sup>5</sup> Langley PC. Imaginary Worlds: The European Network for Health Technology Assessment (EUnetHTA) Recommendations for Health Economic Evaluations. *Inov Pharm.* 2016;7(3): No. 3

and persistence with therapy. As noted above, over the proposed 2-year (or shorter) timeframe it is likely that up to possibly two-thirds of patients would discontinue or be non-adherent with therapy. These patterns need to be factored into claims for product effectiveness. They also give a baseline for evaluating comparative claims for persistence and adherence behavior as part of the assessment protocol.

Rejecting the reference case as the focus for modeling claims does not mean throwing the bathwater out with the baby. Many of the standards and processes detailed in the new Guideline will be relevant to the construction of short-term claims models. Just as the reference case modeled outcomes can be expressed as total costs and outcomes, incremental costs and outcomes and incremental cost-effectiveness ratios, so short-term claims can be expressed in these terms. The key is the specification of a claims assessment protocol detailing how the claims are to be validated and reported to decision makers. This does not mean that claims expressed in quality of life terms are necessarily excluded. It is up to the submission to make the case for quality of life in the assessment timeframe.

## Conclusions

From the perspective of normal science, constructed evidence for product impact claims is not acceptable. Regardless of how decision modeling is defended by the application of validation standards, the treatment of uncertainty and the application of value of information techniques, in the last resort the model stands or falls on its ability to generate credible, evaluable and replicable predictions. The new Dutch guideline fails to meet this standard. If the guideline is to be seen as credible then it must abandon constructing model evidence in favor of evaluable and replicable claims for comparative product performance.



- <sup>6</sup> Langley PC. Langley PC. Supporting formulary decisions: The discovery of new facts or constructed evidence? *Inov Pharm.* 2016; 7(2): No.15
- <sup>7</sup> Langley P. Validation of modeled pharmacoeconomic claims in formulary submissions. *J Med Econ.* 2015;18(12):993-999
- <sup>8</sup> 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
- <sup>9</sup> 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
- <sup>10</sup> Pigliucci M. Nonsense on stilts.: How to tell science from bunk. Chicago: University of Chicago Press, 2010
- <sup>11</sup> Langley PC. Nullius in verba: Version 2.0 of the University of Minnesota, School of Social and Administrative Pharmacy Program, Proposed Guidelines for Formulary Evaluation. *Inov Pharm.*2016;7(4): No.16
- <sup>12</sup> Versteegh MM, Vermeulen KM, Evers S et al. Dutch tariff for the 5 level version of EQ-5D. *Value Health.* 2016;19(4):343-52
- <sup>13</sup> Vemer PO, Ramos IC, van Voorn GAK et al. AdViSHE: A validation-assessment tool of health-economic models for decision makers and model users. *Pharmacoeconomics.* 2016;34:340-61
- <sup>14</sup> Langley PC. Validating imaginary worlds: The AdViSHE Assessment Tool. *Inov Pharm.* 2017;8. No. 3
- <sup>15</sup> Wootton D. The Discovery of Science. New York: Harper Collins, 2015
- <sup>16</sup> Hartung D, Bourdette D, Ahmed S et al. The cost of multiple sclerosis drugs in the US and the pharmaceutical industry. Too big to fail? *Neurology.* 2015;84:2185-2192
- <sup>17</sup> Langley PC. Next Generation Sequencing Claims in Precision Medicine: Questions a Formulary Committee should ask. *Inov Pharm.* 2016;7(4): No. 11
- <sup>18</sup> Langley PC. Creative Destruction: Next Generation Sequencing in Drug Development, Formulary Evaluations and Pricing. *Inov Pharm.* 2016;7(4): No. 13
- <sup>19</sup> Versteegh M, Knies S, Brouwer W. From good to better: New Dutch guidelines for economic evaluations in healthcare. *Pharmacoeconomics.* 2016;34:1071-1074