

2-13-2017

Imaginary Worlds and Efficiency Frontiers: Should We Abandon the IQWiG Health Technology Assessment Model?

Paul C. Langley

University of Minnesota, langley@maimonresearch.com

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

Recommended Citation

Langley PC. Imaginary Worlds and Efficiency Frontiers: Should We Abandon the IQWiG Health Technology Assessment Model?. *Inov Pharm.* 2017;8(1): Article 13. <http://pubs.lib.umn.edu/innovations/vol8/iss1/13>

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

Imaginary Worlds and Efficiency Frontiers: Should We Abandon the IQWiG Health Technology Assessment Model?

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

Abstract

The contribution of cost-effectiveness analysis to the pricing of pharmaceuticals in Germany is at best marginal and in many, if not most cases, absent. While this may reflect a reasonable belief that cost-effectiveness analysis adds little if anything to pricing and formulary placement decisions, its marginalization reflects considerable dissatisfaction, if not frustration, with modeling efforts by the Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). In part, this reflects the rejection of quality adjusted life years (QALYs) as the common outcome standard, together with the adoption of the efficiency frontier as the default framework for modeled claims. The purpose of this commentary is to consider the merits, in the German context, of an efficiency frontier framework for cost-effectiveness and pricing decisions. The commentary concludes that the efficiency frontier framework for health technology assessment, in supporting the creation of non-evaluable claims from models or simulations, fails of to meet the standards of normal science: it fails to support claims that are credible, evaluable and replicable. It should be abandoned. If cost-effectiveness modeling is to play a constructive role in pricing negotiations in Germany then manufacturers should be required to submit evaluable claims. The most effective way of ensuring this is to require manufacturers to accompany any submission for a new product with a protocol detailing how their claims, to include those for clinical outcomes, cost-effectiveness and budget impact, are to be evaluated and reported to decision makers in a meaningful time frame.

Keywords: IQWiG, QALYs, imaginary worlds, efficiency frontier, credible standards

Introduction

Over the past 25 years recommendations and standards for assessing the merits of competing health care interventions have focused models or simulations that are intended to 'inform' decision makers through the construction of imaginary worlds ¹. In the past nine months, a number of commentaries have been published in *INNOVATIONS in Pharmacy* pointing to the lack of scientific merit in this approach to the economic evaluation of claims for pharmaceutical products and devices. These commentaries include reviews of technology assessment standards in the US, UK, Ireland, New Zealand, Australia, the Netherlands and France, together with the proposed European Union standards under the EUnetHTA umbrella ^{2 3 4 5 6 7 8 9}. A common feature of these guidelines is the adoption of the National Institute for Health and Care Excellence (NICE) reference case which focuses on the 'gold standard' of constructing non-evaluable claims for competing pharmaceutical products in terms of cost-per-quality adjusted life years (QALYs). The typical reference case framework for

chronic disease is a modeled or simulated lifetime (or long-term) tracking of the natural course of the disease for a hypothetical patient cohort. Adoption of a reference case or similar framework that focuses on lifetime models is in contrast to the standards of normal science. As stated in the draft for the latest Canadian guidelines: *Economic evaluations are designed to inform decisions. As such, they are distinct from conventional research activities, which are designed to test hypotheses* ¹⁰. Unfortunately, as pointed out on a number of occasions in the previous commentaries, building simulations to 'inform decisions' is hardly an acceptable basis for formulary decision making as we have no idea whether the constructed claims are right or even if they are wrong ^{11 12}. In short, the commentaries have pointed out that outcomes for competing therapies expressed over the long-term or a lifetime means that the claims made are not credible, evaluable or replicable ^{13 14 15 16 17 18 19 20}. They are best seen as pseudoscience sharing the platform with intelligent design rather than natural selection ²¹.

Although they are the preferred 'end-point' in many, if not the majority of technology assessment guidelines, QALYs have come in for increasing criticism over the last few years. Apart from ethical issues, there have been concerns expressed over the methodological foundations for QALY claims, the number of competing QALY instruments, differences in the process of utility evaluation and their lack of sensitivity, and an outright rejection of their use in claims for certain formulary decisions in the US under the Patient Protection and Affordable Care Act

Corresponding author: Paul C Langley, PhD

Adjunct Professor, College of Pharmacy

University of Minnesota

Director, Maimon Research LLC

5061 North Apache Hills Trail, Tucson, AZ 85750

Tel: 520-577-0436; Email: langley@maimonresearch.com

Web: www.maimonresearch.net

²² ²³. In Germany the *Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG))*, as detailed below, has rejected QALYs as a ‘gold standard’ of comparative effectiveness in its ‘General Methods’ for health technology assessment ²⁴. Rather, the focus is on disease specific therapy measures. There is no single preferred measure in specific disease states; rather, there is the flexibility, within the efficiency frontier methodology to generate maximum reimbursable prices for individual effectiveness measures. Under the efficiency frontier methodology total benefits are compared to total costs rather than applying incremental cost-effectiveness ratios (ICERs). The principal reason for this rejection, although a case can be made that if QALYs are seen as an international standard then Germany should follow suit, is that there are competing constitutional rights and legal restrictions. These prevent the application of cost-per-QALY ratios together with thresholds for pricing and out-of-pocket payment assessments. There may also be cultural reasons for an efficiency frontier with a ‘reluctance to frame healthcare decisions around cost-based valuations of human health’ and their potential implications for health care rationing ²⁵. Unfortunately, as noted below, even if QALYs are rejected, the same objections apply if alternative outcomes are modeled in a long-term or lifetime framework. They still fail to meet the standards for normal science.

The purpose of this commentary is to consider the merits, in the German context, of an efficiency frontier framework for cost-effectiveness and pricing decisions. The principal focus is on the need to create a health technology assessment framework that supports credible and evaluable claims for anticipated impact of a new product in German target patient populations. Technology assessment standards should be consistent with the standards for hypothesis testing that characterize normal science. If technology assessment standards, such as those proposed by IQWiG fail this test then they should be rejected.

Price Determination in Germany

There are three agencies involved in price negotiations with manufacturers in Germany. These are: (i) The Federal Joint Committee (Gemeinsamer Bundesausschuss, GBA); (ii) IQWiG; and (iii) The National Association of Statutory Health Insurance Funds (SFI). Under the 2011 Pharmaceutical Market Reorganization Act (AMNOG) cost-effectiveness analysis has been put to one side from its previous central role as providing a framework for determining maximum reimbursable prices under the IQWiG umbrella. The current regulations require the negotiation of new drug prices to be informed by a comparative clinical effectiveness benefit assessment by IQWiG. The remit for IQWiG is to determine whether a

manufacturer’s dossier has made the case that a new drug provides added net patient-relevant benefit, including health status improvement, shortening disease duration, life expectancy extension, alleviation of side effects or improved quality of life. The product is then categorized by IQWiG as providing a major, considerable or minor benefit.

Cost-effectiveness considerations are only introduced if there is a failure to achieve a negotiated price. These take place under two scenarios: (i) an assessment of submitted dossiers and (ii) a commissioned health economics evaluation. Both scenarios are initiated by the GBA and undertaken by IQWiG as two of its product-specific procedures. The only scenario under which IQWiG considers price is where the GBA and the manufacturer co-ordinate and agree to IQWiG’s undertaking a cost-effectiveness assessment. Otherwise, IQWiG’s role is to assess the drug’s benefit.

It should be noted that apart from the structured assessment process for new drugs the GBA has the authority under AMNOG to commission IQWiG to undertake assessments of existing medications and devices. This covers drugs that entered the German market before the 2011 reform and were not evaluated for net benefit as well as follow-up studies for drugs previously assessed. Manufacturers may also request a follow-up assessment.

Comparative Effectiveness Assessment

As noted, IQWiG’s role is to assess the added benefits a drug brings to the target population. This is achieved through evaluating a manufacturer’s submission to the GBA. This assessment may involve IQWiG undertaking its own literature search as well as soliciting external expert advice on the clinical benefits of a new product. The assessment should, hopefully, come to a clear conclusion as to whether or not the product contributes added benefits and harms for the individual patient. The key patient relevant outcomes are mortality, morbidity and health-related quality of life. A hierarchy of outcomes may be created with analyses focused on the higher weighted ones.

The term ‘benefit assessment’ refers to the ‘whole process of the assessment of medical interventions with regard to their positive and negative causal effects compared to a clearly defined alternative treatment, a placebo or no treatment’. Surrogate endpoints are considered unreliable and will only be considered if they have been appropriately validated.

The term ‘harm’ refers to adverse events and which are an ‘essential and equal component in the benefit assessment of an intervention’. Relevant adverse events include: (i) those that completely or almost completely offset the benefit of an

intervention'; (ii) vary substantially between two or more otherwise equivalent treatment options; (iii) occur predominantly in particularly effective treatment options; (iv) have a dose-effect relationship; (v) are regarded as particularly important by patients; and (vi) are accompanied by serious morbidity, increased mortality or substantial impairment in quality of life.

Manufacturer Submission

The benefit assessment of a drug starts from a dossier submitted by a manufacturer. The information provided should include: (i) approved therapeutic indication; (ii) medical benefit; (iii) likelihood of added medical benefit compared with appropriate comparator therapy; (iv) number of targeted patients and patient groups; (v) cost of treatment for the SRI; and (vi) requirements for quality-assured use of drug. The added benefit claimed in the dossier should be categorized as major, considerable, minor, non-quantifiable, no added benefit and benefit less than for comparator.

Evidence Evaluation

The IQWiG evidence evaluation considers what is described as both qualitative and quantitative certainty in results. The former is determined by an assessment of the study design and outcome related measures to avoid or minimize bias. The latter by effect size. Studies are classified as having high, moderate or low qualitative certainty of results. Depending upon the number of studies assessed together with an evaluation of study heterogeneity, the proof of evidence is determined for the benefit inferred for each patient-relevant outcome. If possible, a summary is presented in the form of a weighting of benefits and harms. There is the option of aggregating the various outcomes to a single measure. Supplementary information may be requested from the manufacturer, supported if necessary by a systematic literature review.

IQWiG Modeled Efficiency Frontier

The efficiency frontier was developed at the request of IQWiG as a method to inform decision makers of the net costs and additional benefits of an intervention given other interventions in the disease or therapy area²⁶. Its purpose is to generate evidence to support negotiations for an acceptable reimbursable price. It is not IQWiG's role to make recommendations as there is no evidence for the financial capacity of the SHI insurers or their willingness-to-pay. The efficiency frontier exercise is merely to provide information. There is no evidence of any intent to utilize the efficiency frontier as the basis for generating evaluable comparative claims for either clinical effect or cost-effectiveness.

The modeled efficiency frontier is constructed for all available treatments for a given indication within a disease or therapy

area. The frontier is constructed as a plane with health effects on the vertical axis and net costs on the horizontal. The effects, for example modeled lifetime outcomes for a hypothetical patient cohort, should meet cardinal measurement standards over the appropriate benefit range. Treatment strategies that are not considered dominated are linked in ascending order of effectiveness. A maximum reimbursable price is determined as an input to total costs by the linearly extrapolated last segment of the frontier. This extrapolation establishes the same trade-off for benefits and costs as the trade-off between the most efficient or effective intervention relative to the second most efficient. This establishes the 'authorities' implicit willingness-to-pay for additional benefits in that indication. The reimbursable price is considered in the context of total costs that include proposed drug costs (which comprise the number of units by price proposed by the manufacturer) and other direct medical costs. The model can be extended to include non-medical costs such as productivity losses.

There are no restrictions on the number of outcomes that could be identified within an indication. Where there are multiple end-points, clinical or patient reported, it is possible to create an aggregate efficiency frontier in relation to the net costs of the interventions. This translates to possible techniques to aggregate weighted outcomes to create this 'overall' frontier.

Each non-dominated intervention or efficient point on the frontier is constructed from a standard model for that indication. The modeled claims for benefits and net costs are, given the timeframe (see below) typically non-evaluable. Uncertainty in the models for each point on the frontier (including the new product) will reflect variations in its position on the cost-benefit plane.

Discussion

The comments and recommendations below should be seen alongside the current review of the IQWiG guidelines for a proposed Version 5.0²⁷. Unfortunately, it was not possible to review the draft proposals as these are only available in German. The focus, therefore, is on Version 4.2.

The IQWiG Model

In common with other health technology assessment groups IQWiG requires a modeled claim to support product benefits. According to IQWiG in populating the efficiency frontier modeled estimates of each of the potential comparators are to be developed with their cost-effectiveness ratios. Presumably this means that the frontier should capture cost-effectiveness estimates of the new product, utilizing the identical model structure, together with those for the 'less efficient' ranked products. The new product estimate should

include the proposed unit price, with existing market prices included in the modeled claims for the other non-dominated frontier product coordinates. If there are a number of efficiency frontiers constructed, then these should be weighted to generate an overall measure of overall benefit for these various products. Each of these, presumably, utilizing the common model structure.

The development of a modeled claim is seen by IQWiG as the key component of a health technology assessment in extrapolating benefits and costs beyond those generated by the studies used to support model assumptions. In practical terms, this can be interpreted as a preference for a long-term if not a lifetime modeling of costs and benefits. The principal objection to this hierarchy of modeled claims, where they are used to create an efficiency frontier of lifetime outcomes is that the claims for benefits and costs for each point on the frontier are imaginary. There is no basis for evaluating their merits. They fail to meet the standards for hypothesis testing. Modeled claims and simulations, of course, can always fail²⁸. Unfortunately, in IQWiG's case the simulations driving the efficiency frontier are immune to failure. The only exceptions would be in end-stage disease states where the expected survival profile for the target population is short and potentially open to evaluation. There is no requirement that modeled claims for products already on the German market should be evaluated as validation for the efficiency frontier claim.

In common with other health technology assessment guidelines, there appears to be no attempt by IQWiG to ask manufacturers to stand behind their claims for the new product. Irrespective of the dossier requirement for manufacturers to categorize the extent of intended benefit. Together with IQWiG's possible reassessment of the appropriate benefit category, there is no apparent interest in the potential for claims evaluation, and indeed replication, in the target German populations. The possibility of feedback to decision makers is not considered as part of a possible program of ongoing disease area and therapeutic class reviews. Indeed, given the effort IQWiG apparently puts into determining the quality of the evidence base and the estimate and categorization of the effect size for each outcome as inputs to an overall assessment of an overall categorization of the added benefit of the drug, it is not clear why the issue of evaluation and replication in the target populations is not considered. As it stands these are constructed claims of added benefit, but we have no idea if these are right or even if they are wrong. The notion of progress through hypothesis testing or systematic observation of claims through comparative effectiveness assessments is absent.

Comparators and Simplification

In deriving the efficiency frontier, IQWiG requires 'all healthcare-relevant interventions in the therapeutic area to be considered'. In this context, 'considered' means that each of the relevant interventions, not necessarily those that the new product is likely to supplant, should be modeled. Each of these models is seen as a 'simplified depiction of reality' with analytic clarity achieved by reducing complexity. This 'simplification', IQWiG acknowledges, cannot be specified *a priori*, instead there should be agreement with experts to check on the models external validity. Presumably, as noted above, the model structure should be the same for each of the relevant interventions, where the modeled claims for benefits and net costs will determine whether or not comparators yield the greatest benefit in relation to costs (i.e., fall on the efficiency frontier).

There is no requirement, as also noted above, even if the intervention has been on formulary for some years, for the modeled claim to have ever been evaluated in the target population. Indeed, the present modeled claim for an existing product may bear no resemblance in either its structure or parameter values to previous claims utilizing an alternative model by a different manufacturer. In effect, all the points on the efficiency frontier could represent interventions for which the claims are non-evaluable. The frontier, therefore, is an entirely imaginary construct. This situation is compounded where there might be a family of frontiers and an overall efficiency frontier is developed as a weighted average of these imaginary benefit and cost claims.

Validation

IQWiG points out that a simulation model that is valid for one research question might not be valid for another. Validation for IQWiG encompasses three elements: whether the model (i) adequately reflects the course of the disease and treatment; (ii) is technically correct; and (iii) generates predictive results that reflect the 'real world'. While IQWiG recognizes that prediction is the 'most desirable form of validity', it points out that this is difficult (if not impossible) and falls back on proposing a comparison with previous modeled results. Given the apparent importance that IQWiG attaches to a model's ability to generate credible predictions, it is surprising that IQWiG does not emphasize the importance of only recognizing models that generate credible and evaluable claims. As it stands, IQWiG would appear to accredit equal status to models that are capable of generating evaluable claims and those that generate claims that are immune to failure. This means that IQWiG is prepared to accept an efficiency frontier framework where the 'efficient' comparators claims are non-evaluable for the target German patient population.

Adherence and Persistence

In common with other health technology assessment agencies IQWiG puts to one side any attempt to assess the potential impact on the modeled benefits of poor adherence and discontinuation of therapy. There is now ample evidence that for many if not most pharmaceuticals the majority of patients have discontinued therapy within two to three years of an index prescription. Adherence, for those remaining on therapy is also poor with low medication possession or days covered ratios. IQWiG's standards for modeling benefits and costs take no account of anticipated compliance behavior or, indeed, of intervention strategies that may have accompanied comparator therapies to support compliance.

Failure to take compliance behavior into account means that it is likely that the benefits attributed to the comparator products together with the modeled benefits from the new product will be overstated. Given the fact that the comparator therapies will have been on the market in Germany for some years, it would surely be possible to provide compliance data to modify modeled claims for chronic disease interventions. An assessment of product and comparator benefits, particularly in chronic disease, may be (at best) only short term once the implications of poor compliance behavior are taken into account.

A further issue which is not considered by IQWiG and which is clearly relevant for adherence and persistence is the presence of comorbidities in the target population. The majority of older patients with chronic conditions will have one or more comorbidities. Failure to accommodate these in modeled claims casts further doubt on the model's relevance.

Uncertainty

Three types of uncertainty are typically considered in decision models: stochastic, parameter and structural. IQWiG recognizes these and requires modeled claims to address these issues both for individual outcome-specific models as well as for overall benefit assessments. In attempts to defend the viability of modeled claims, considerable effort has gone into developing techniques to capture the impact of uncertainty in the claims that accompany, for example, lifetime cost-per-QALY claims under the NICE reference case model. While these have a possible theoretical interest, their practical benefit is difficult to see given that the underlying construct is still an imaginary world. Of course, where competing models may come to quite different conclusions for the impact of new therapies in disease areas there seems little scope for assessing the various models outside of evaluable claims.

These arguments also apply to attempts to model uncertainty for efficiency frontiers. A recent attempt to model the effect of parameter uncertainty utilizing a Monte Carlo simulation to create scatter plots, confidence intervals and contour plots together with the construction of price acceptability curves (PACs), illustrates the limitation of this for decision-making²⁹. Apart from the imaginary nature of the final PAC construct, it is far from clear how decision makers would interpret this and which interquartile range would be selected as input to pricing negotiations.

Replication of Results

Even though IQWiG may undertake a comprehensive systematic review of the relevant literature to assess the likely effect size for the competing interventions, there are recognized limitations to accepting claims for comparative effects from indirect comparisons. More to the point, there have been increasing concerns over the ability to replicate phase 2 and phase 3 randomized clinical trial (RCT) results³⁰. This puts increased emphasis on asking manufacturers to underwrite protocols to evaluate claims made for treatment effect and adverse effect profiles. These, as noted below should be an integral part of a protocol to establish claims for cost-effectiveness in the target patient population. The issue of replication of clinical effect claims in German treatment populations does not appear in the IQWiG guidelines.

IQWiG's categorization of a drug product as of major benefit across the range of outcomes considered just simply sets the stage for providing credible claims, evaluating the claims and replicating those claims in target patient populations. Evaluating those claims should be a necessary next step.

Efficiency Frontier Redundancy

Irrespective of the contribution of efficiency frontiers to portfolio management models, the application of the notion of an efficiency frontier adds little if anything to the framework for the competing assessment of drug therapies. While there is no doubting the visual appeal of a frontier representation, it needs to be kept in mind that the frontier is simply an imaginary construct. There may in fact be a family of imaginary efficiency frontiers, each relying upon constructed evidence for the potential outcomes identified from the reviews of RCTs and supplementary literature searches. There is also, apparently, a 'master' imaginary frontier which is a weighted aggregate of the individual efficiency frontiers. The weighting criteria are unknown as are the measurement properties of this putative health benefit index.

This does not mean that it is not impossible to construct imaginary lifetime models utilizing the efficiency frontier concept. A recent paper by Gissel et al applied the efficiency

frontier method to the impact of direct-acting retrovirals in Hepatitis C in Germany³¹. A lifetime Markov framework was developed to simulate immediate treatment effect (up to 72 weeks) in sustained virologic response (SVR) as well as the impact on possible long-term complication costs associated with disease progression across a range of single therapy and combination therapy options. While the long-term claims are obviously not evaluable and should be put to one side, there is the potential within the model to evaluate short term SVR and associated costs and to generate evaluable claims that can be assessed and reported to decision makers in a meaningful time frame. The critical question is whether or not it is possible to defend empirically a reference marginal efficiency frontier cost (e.g., cost per additional SVP percentage point increase) to establish the threshold values for evaluating the cost consequences of new products and recommendations for pricing and possible rebates. Clearly, if a reference point willingness-to-pay as determined by previously negotiated prices for products in the disease area yields a commonly agreed marginal willingness to pay for unit or percentage point improvements in health outcomes, then these can be applied to claims that have been empirically assessed. It becomes a simple exercise to apply the willingness to pay amounts to the short term claims assessments. Unfortunately, we may have doubts as to the empirical base on which the previous prices were negotiated; a situation which is further confounded by a possible lack of agreement on what is considered the 'primary' outcome standards within specific disease areas.

Systematic Price Increases

A further issue, which has been touched upon in previous commentaries, is the currency of a simulated or modeled claim. Presenting competing claims in the framework of a lifetime reference case cost-per-QALY model raises the question of how to accommodate potential price increases for drug products and devices if the exercise is to generate claims that 'meet', for example, lifetime willingness-to-pay cost-per-QALY thresholds. In the US, for example, models developed by the Institute for Clinical and Economic Review (ICER) exclude any attempt to factor in possible long-term price increases³². Even if the NICE lifetime reference case paradigm is accepted by decision makers, the absence of modeling potential price increases restricts the currency of any argument that the product meets willingness-to-pay thresholds. It flies in the face of abundant evidence that, even with claims for moderation in price increases and, in the US in particular, policies to offer discounts or free access to offset co-payments, manufacturers are committed to a policy of systematic price increases over the patent life of a product³³. These price increases, semi-annual and annual, are all too often in double digits which means that within five or six years the product price, putting to one side possible discounting arrangements, is doubled. To

the extent that, in Germany, there are anticipated price increases then these need to be factored into any assessment of future costs.

Minnesota Proposed Guidelines

Rather than focusing on the construction of imaginary worlds as an input to pricing negotiations, the solution proposed here is for IQWiG to mandate claims based on short term models; models that can generate claims that are credible, evaluable and replicable, and as a result, provide the opportunity for feedback to decision makers in a meaningful time frame. In order to illustrate how this could be implemented, the Program in Social and Administrative Pharmacy at the University of Minnesota, has published a set of proposed guidelines for formulary committees^{34 35}. These set standards for modeled claims, either as extrapolations from clinical trials or as stand-alone models, which can be evaluated within a two-year time frame. The key requirement is that submissions for new products should be accompanied by a protocol detailing how the claims are to be evaluated and reported to a formulary committee. It is responsibility of the manufacturer to underwrite the study design or to report on the results of a parallel study that may have been undertaken as part of another submission. There is no restriction on the type of claim as long as it is evaluable and is acceptable to the formulary committee. The claim can be expressed in utility as well as clinical effectiveness terms. Unless the timeframe for disease survival is short (e.g., in metastatic cancer interventions) claims expressed in lifetime or long-term cost-per-QALY or cost-per-life year saved would not be considered credible.

There are, in fact, instances of short-term models that have been published over the last few years^{36 37}. These short term models which in these instances, consider the effectiveness of biologic treatments in rheumatoid arthritis in a Spanish target population and percutaneous coronary interventions in a French target population demonstrate the ease of constructing short-term models with evaluable claims.

Options for GBA, SRI and IQWiG

If the proposition is accepted that competing claims in Germany for the added benefits and potential harms of drug products should be credible, evaluable and replicable, then there are a number of options that could be considered for Version 5 of the IQWiG 'General Methods'. These include requirements that:

- The focus should be on short-term evaluable claims that can be assessed and reported back to health care decision makers in a meaningful (e.g., under two year) time frame. Long-term or lifetime modeled claims should be rejected. The claim submission,

following the proposed Minnesota guidelines for formulary evaluation, should be accompanied by a protocol detailing how claims are to be evaluated and reported to IQWiG, the GBA and the SFI. The protocol would be required to be agreed between the parties. The protocol should be comparative, identifying the product (or possibly products) most like to be replaced in clinical practice in the target treating population.

- Rather than accommodating any number of potential outcome measures within a disease or therapy area, IQWiG should agree with the GBA, SFI and professional groups on a primary outcome measure. This could be related to the stage of the disease and linked to care pathways (e.g., median survival for metastatic cancers). This would not exclude the identification of what may be described as secondary outcome measures. Manufacturers would then be informed that in developing modeled claims for the product and comparators, the model should be developed to generate evaluable claims for the primary outcome. This would serve two purposes; (i) it would provide a focus for the development of a claims assessment protocol that would be underwritten by the manufacturer and (ii) it would avoid having to generate weighted aggregate claims across the various outcome measures that may be considered appropriate to the disease area. The claims protocol could, of course, accommodate secondary outcome measures.
- In the absence of a report on the results of the protocol assessment for the modeled primary claim, any interim pricing negotiation should take as the reference point the existing unit prices for the respective comparator (e.g., the product most likely to be replaced in clinical practice). Multiple comparisons, particularly for those products deemed less 'efficient', are an unnecessary burden for both manufacturers and assessment agencies. A final price determination would then occur once the protocol results were available for the target population in Germany. This would reflect the assessed comparative benefits and costs, as well as the potential impact on these of compliance behavior.
- The concept of an efficiency frontier should be abandoned. IQWiG should adopt a simpler and more robust approach to evaluating the benefits and costs from products entering the market. Attempting to

populate one or more efficiency frontiers and aggregating across these with multiple criteria decision analysis (MCDA) or some variant of conjoint analysis is unnecessary and, from a price negotiation perspective, an irrelevant exercise. Attempting to agree, for example, on the elements of a MCDA and then the preference weights to be attached when so many parties are involved is unrealistic. Impractical and irrelevant.

- Any attempt to re-introduce long-term or lifetime cost-per-QALY models and thresholds should be avoided. This does not mean that primary claims expressed as QALYs should be put to one side. While there may be substantive objections in Germany to QALYs and thresholds as rationing criteria, there is no reason why evaluable claims expressed in terms of a specific QALY instrument, or other patient reported outcome (PRO) measure specific to a disease or therapy area should not be accepted in prospective assessments of short-term claims.
- A short-term model should accommodate uncertainty, but only as it is predicted to impact the claims made. Proposed study designs to evaluate claims, whether they are for a phase 4 trial design or a systematic observational study would be expected to meet the standards for evidence based medicine. This will avoid, in the focus on evaluable claims, attempts to assess the impact of uncertainty in models which are not designed to generate evaluable claims and further attempts, in the case for example of parameter uncertainty, imaginary price reimbursement acceptability curves.

Conclusions

Whether the GBA, SFI and IQWiG have any interest in a formulary evaluation program that meets the standards of normal science is a moot point. The options considered here provide a simpler and more direct approach to assessing the value of new products rather than trying to continue to support the application of efficiency frontiers. While the efficiency frontier has some interest as a theoretical concept, it is not a viable practical tool for pricing and formulary decisions, particularly when the focus is on constructing evaluable claims. Abandoning the efficiency frontier will bring to the fore a commitment to disease and therapeutic area strategies that support a commitment to progress and feedback in health care interventions. Rather than relying on imaginary constructs, the approach proposed here will create an evidence base that is of benefit to target patient populations.

References

- ¹ Drummond M, Sculpher M, Torrance G et al. *Methods for the Economic Evaluation of Health Care Programmes* (3rd Ed). Oxford: University Press, 2005
- ² Academy of Managed Care Pharmacy. *The AMCP Format for Formulary Submissions (Version 4.0)*. AMCP: 2016
- ³ United Kingdom. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal*. London: Nice, April 2013.
- ⁴ Ireland. Health Information and Quality Authority. *Guidelines for the Economic Evaluation of Health Technologies in Ireland*. Dublin: HIQA; 2014
- ⁵ Australia. Department of Health. *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (PBAC)*. Version 5.0. Canberra, 2016
- ⁶ New Zealand. Pharmaceutical Management Agency. *Prescription for Pharmacoeconomic Analysis: Methods for cost-utility analysis (v 2.1)*. PHARMAC, 2012
- ⁷ Zorginstituut Nederland. *Guideline for Economic Evaluation in Healthcare*. Diemen: 2016
- ⁸ Haute Autorité de Sante. *Choices for Methods for Economic Evaluation*. Paris, 2012
- ⁹ European Network for Health Technology Assessment (EUnetHTA). *Methods for health economic evaluations*. May 2015
- ¹⁰ Canada. Canadian Agency for Drugs and Technologies in Health. *Guidelines for the Economic Evaluation of Health Technologies: Canada. Methods and Guidelines. Draft Report*. Ottawa: CADTH, October 2016
- ¹¹ Langley PC. Supporting formulary decisions: The discovery of new facts or constructed evidence? *Inov Pharm*. 2016; 7(2): No.15
- ¹² Langley P. Validation of modeled pharmacoeconomic claims in formulary submissions. *J Med Econ*. 2015;18(12):993-999
- ¹³ Langley PC. Modeling Imaginary Worlds: Version 4 of the AMCP Format for Formulary Submissions. *Inov Pharm*. 2016;7(2): No. 11
- ¹⁴ Langley PC. Sunlit Uplands: The Genius of the NICE Reference Case. *Inov Pharm*. 2016;7(2): No. 12
- ¹⁵ 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
- ¹⁶ 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
- ¹⁷ Langley PC. Dreamtime: Version 5.0 of the Australian Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee (PBAC). *Inov Pharm*. 2017;8(1): No. 5
- ¹⁸ Langley PC. Denkbeeldige Wereld: The New Dutch Guidelines for Economic Evaluations in Healthcare. *Inov Pharm*. 2017;8(1): No. 4
- ¹⁹ Langley PC. HAS Should Not Be NICE: Rejecting Imaginary Worlds in the French Technology Assessment Guidelines. *Inov Pharm*. 2017;8(1): No. 8

- ²⁰ Langley PC. Imaginary Worlds: The European Network for Health Technology Assessment (EUnetHTA) recommendations for health economic evaluations. *Inov Pharm*. 2016;(7) No. 3
- ²¹ Pigliucci M. Nonsense on Stilts: How to tell science from bunk. Chicago: University of Chicago Press, 2010
- ²² Beresniak A, Dupont D. Is there an alternative to quality adjusted life years for supporting healthcare decision making? *Expert Rev Pharmacoecon Outcomes Res*. 2016;16(3):353-57
- ²³ Beresniak A, Medina-Lara A, Auray JP et al. Validation of the underlying assumptions of the quality-adjusted life-years outcome: results from the ECHOOUTCOME European project. *Pharmacoeconomics*. 2015;33(1):61-9
- ²⁴ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care). General Methods. Version 4.2. IQWiG: April 2015,
- ²⁵ Economist Intelligence Unit. Value-based Healthcare in Germany. London, 2015
- ²⁶ Caro JJ, Nord E, Siebert U et al. The efficiency frontier approach to economic evaluation of health-care interventions. *Health Econ*. 2010;19:1117-1127
- ²⁷ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden. Entwurf für Version 5.0 vom 07.12.2016
- ²⁸ Grim P, Rosenberger R, Eason RE et al. How simulations fail. *Synthese*. 2013;190(12):2367-2390
- ²⁹ Stollenwerk B, Lhachimi SK, Briggs A et al. Communicating the parameter uncertainty in the IQWiG efficiency frontier to decision-makers. *Health Econ*. 2015;24:481-90
- ³⁰ Baker M. Is there a reproducibility crisis? *Nature*. 2016;553:452
- ³¹ Gissel C, Götz G, Mahlich J et al. Cost-effectiveness of interferon-free therapy for Hepatitis C in Germany – an application of the efficiency frontier approach. *BMC Infect Dis*. 2015;15:297
- ³² Institute for Clinical and Economic Review (ICER). Targeted Immunomodulators for the Treatment of Moderate-to-Severe Plaque Psoriasis: Effectiveness and Value - Final Evidence Report December 2, 2016
- ³³ Cheng J, Feldman SR. The cost of biologics for psoriasis is increasing. *Drugs in Context*. 2014;3:212266
- ³⁴ 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
- ³⁵ University of Minnesota, College of Pharmacy. Program in Social and administrative Pharmacy. Guidelines for Formulary Evaluations [proposed]. Version 2.0. December 2016.
https://www.pharmacy.umn.edu/sites/pharmacy.umn.edu/files/minnesota_guidelines_december_2016.pdf
- ³⁶ Beresniak A, Ariza-Ariza R, Garcia-Llorente JF et al. Modelling cost-effectiveness of biologic treatments based on disease activity scores for the management of rheumatoid arthritis in Spain. *Int J Inflammation*. 2011; ID 277634.
- ³⁷ Beresniak A, Caruba T, Sabatier B et al. Cost-effectiveness modelling of percutaneous coronary interventions in stable coronary artery disease. *World J Cardiology*. 2015;7(10):594-602