INNOVATIONS in pharmacy

Volume 8 | Number 2 Article 10

5-4-2017

Imaginary Worlds and the Institute for Clinical and Economic Review (ICER) Evidence Report: Targeted Immune Modulators for Rheumatoid Arthritis

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

Imaginary Worlds and the Institute for Clinical and Economic Review (ICER) Evidence Report: Targeted Immune Modulators for Rheumatoid Arthritis

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

Abstract

In April 2017, the Institute for Clinical and Economic Review (ICER) issued its evidence report on the value of targeted immune modulators (TIMs) in rheumatoid arthritis. The report made the case that for the TIMs to be accepted for formulary placement in the US, where notional willingness-to-pay thresholds are the ICER gateway criteria, manufacturers should be prepared to offer substantial unit price discounts. The purpose of this commentary is to make the case that the methodology underpinning the ICER claims for value assessment does not meet the required standards of normal science. None of the claims made for clinical and comparative cost-effectiveness are credible, evaluable and replicable. As such, formulary committees have no idea whether ICER recommendations are right or even if they are wrong. They are, in fact, immune to failure and should be rejected. Utilizing ICER claims generated by simulated projections, this review points out that it is entirely possible to justify the current WAC or net pricing structure of TIMS. The review concludes that if ICER is to contribute to the successful formulary placement of drugs and devices the methodology for pricing recommendation should be re-assessed. As it stands, questions must be raised regarding recommendations for, possibly unnecessary, price discounts. ICER needs to develop an assessment framework that focuses on developing claims for competing therapies that are robust, evaluable and replicable together with recommendations on how these claims are to be evaluated in a timeframe meaningful to health care decision makers.

Keywords: ICER, rheumatoid arthritis, economic evaluations, imaginary worlds, pseudoscience

Introduction

In April 2017, the Institute for Clinical and Economic Review (ICER) issued an evidence report for the comparative clinical effectiveness, cost-effectiveness and value of targeted immune modulators (TIMs) for patients with moderately-to-severely active rheumatoid arthritis (RA) despite prior treatment with conventional disease-modifying anti-rheumatic drugs (cDMARDs) [1]. The purpose of this commentary is to consider the merits or otherwise of the claims made by ICER for the lifetime cost effectiveness of TIMS and the consequent recommendations, based upon lifetime cost-per quality (QALY) projections, for value-based benchmark prices for the TIMs in treatment practice.

This commentary follows upon two previous commentaries of ICER evidence reports in heart failure and multiple sclerosis [2] [3]. Together these commentaries are part of an ongoing series of commentaries published in the University of Minnesota peer-reviewed journal *INNOVATIONS in Pharmacy* that have provided an ongoing critique of current standards in health technology assessment [4]. 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 and potentially misleading evidence base for effective formulary decisions. 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 a perceived 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 are apparently irrelevant.

The argument put forward in these commentaries is that advocates of models that are intended to 'inform' decision makers in health care systems (whatever that means) rather than establish a practical research program, put to one side a commitment to standards that have been in place since the seventeenth century in favor of what has been described as pseudoscience [5]. 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 reject 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 [6] [7]. Formulary submissions are to be supported by protocols to detail how the claims are to be evaluated and reported. This requirement is not new. It was put forward as a standard over ten years ago in formulary submission guidelines developed for the Wellpoint (now Anthem) health system in the US [8] [9].

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
Email: langley@maimonresearch.com
Web: www.maimonresearch.net

The recent commentary on the ICER evidence review in heart failure drugs pointed out, perhaps not surprisingly, that different models can generate quite different perspectives on the presumption of 'cost-effectiveness' [2]. If modeled claims are to be useful for formulary decision making, then we need to eschew 'black box' lifetime cost-per-QALY models with non-evaluable claims in favor of transparent and public domain accessible models that yield credible, evaluable and replicable claims that can support defensible product placement and pricing decisions. A similar conclusion was reached, again not surprisingly, in the more recent review of the ICER evidence assessment for multiple sclerosis interventions [3]. The claims made for comparative effectiveness and value were not acceptable because they were immune to failure. The review concluded that if ICER models are to contribute to improving our understanding of the effectiveness and costs of DMTs then they should be evaluable in the short-term to allow feedback to formulary committees in a meaningful timeframe.

If ICER is to contribute to formulary decision making then the primary objective must be to support the assessment of robust, evaluable and replicable claims for therapy interventions. The purpose of the present commentary is to point out the limitations of the present ICER methodology and to suggest how these objectives might be met through eschewing lifetime modeled claims and focusing instead of modeled claims, in this case for TIM initiated treatment pathways, that are meaningful.

The ICER Evidence Review

For present purposes the focus is on: (i) outcomes of the ICER assessment of the comparative clinical efficacy of targeted immune modulator treatment pathways (TIMs); (ii) the ICER modeled claims for long term cost-effectiveness of the TIMs; and (iii) the claims for TIM value-based benchmark prices. The therapies comprise: five subcutaneous TIMs (adilumumab, certolizumab, etanercept, golimumab, sarilumab); two subcutaneous or intravenous TIMs (abatacept and tocilizumab), two intravenous TIMs (infliximab, rituximab) and two oral TIMs (baricitrinib, tofacitinib). At the time of the evidence review publication two had yet to receive FDA approval: sarilumab and baricitrinib.

Comparative Efficacy Claims

As the ICER evidence review points out, the claims for comparative product performance should be seen in the context of a number of recent reviews of TIMs, together with network meta-analyses of the biologic disease modifying anti-rheumatic drugs (bDMARDS). ICER provides a brief summary of five evaluations (Appendix E) of TIMS in patients aged 18 years and over with moderate to severe rheumatoid arthritis and an inadequate response to or intolerance of cDMARDs. These are the NICE technology assessment report [10]; the AHRQ comparative effectiveness review [11]; two Cochrane reviews [12] [13]; and a Canadian Agency for Drugs and Technologies in Health (which is presently being updated) [14] [15]. The NICE evidence review group report is of particular note as it includes not only critiques of manufacturer's submissions to support NICE assessment and recommendations, but a comprehensive network meta-analysis and scenario evaluations to support a lifetime cost effectiveness model which, as might be expected., anticipates and mirrors the model presented in the ICER evidence review. At the same time, the focus on the UK is of interest because it offers insight into the process by which the assessment presented the review group was evaluated by NICE and factored into the technology appraisal guidance issued by NICE in January 2016 [16].

ICER estimates of comparative clinical effectiveness were based on published papers as well as abstracts/presentations. These comprised 132 reports, 67 RCTs and 17 observational studies. Of the RCTs evaluated, 60 focused on TIM combination therapy with methotrexate or other cDMARDs, five on TIM monotherapy and two with both combination and monotherapy. Outcomes in the network meta-analysis included the ACR criteria for 20%, 50% and 70% response and, in a more limited data set, radiographic progression. Individual product analysis found that all TIMS in combination with a cDMARD produced statistically and clinically improved symptom response improvements and associated outcomes compared to cDMARDs as monotherapy. All were between two and three times likely to achieve a ≥ACR20 response. The overall incidence of serious infections, deaths and serious adverse events were comparable between treatments.

With only eight head-to-head comparisons of the TIMs of interest the network meta-analysis was driven largely by indirect comparisons. Even so, findings were consistent with those for head-to-head studies as well as assessments of relative differences in ACR response in comparison to cDMARD therapy. The network meta-analysis results were also consistent with findings from other recent assessments. Importantly, however, in the network meta-analysis none of the comparisons between the individual TIMS differed at conventional decision levels.

In pointing to controversies and uncertainties in drawing inferences for therapy options and therapy sequencing, the ICER review points to the absence of data on the long-term effects of initial DMARD treatment, in particular the timing of treatment switch decisions. The ICER review notes that comparisons of TIM combination therapy or monotherapy with cDMARDS do not provide an evidence base for comparing treatment sequencing options.

The ICER Model

The ICER sequential treatment or pathway cohort model is intended to provide a framework for assessing the long-term cost-effectiveness of each of the TIMs as combination therapy detailed in the clinical review relative to cDMARDs as monotherapy. The model parameters are taken from the

network meta-analysis and the published literature. The simulated outcomes are: (i) discounted lifetime total payer costs; (ii) life years lived; (iii) quality adjusted life years (QALYs) and incremental cost-effectiveness ratios (ICERs).

The simulation is for a hypothetical cohort of patients with a tracking simulation from a selected TIM initial combination therapy to death (in the model approximately 16-17 years). Treatment was considered effective if the patient achieved an ACR≥ 20. A model treatment cycle of six months was assumed for all therapy initiations and sequences. Beyond the first six months patients were assumed to discontinue therapy only due to adverse events. If the patient failed to respond to initial TIM therapy the stimulation required switching to up to three times: first, within the same TIM class; (ii) to a different TIM class; and then (iii) to palliative care with cDMARDs. The level of ACR response determined quality of life improvements; quality of life was degraded if patients were on cDMARDs. Each TIM was assumed to be used in combination with methotrexate.

The base-case simulation (which is the focus here) assumed a hypothetical cohort with a mean age of 55 years, 79% female, 84% Caucasian, and a mean weight of 170 pounds. Base HAQ prior to cDMARD benefit was 1.7 (range 1.37 to 2.03) and a baseline mTSS of 54 (SD:64). Age and gender were used to estimate mortality risk, with mean weight to calculate dosing. Costs were categorized as: (i) drug acquisition costs (WAC discounted to give a net price); (ii) administration and monitoring costs; (iii) health care utilization costs (e.g., hospitalizations, office visits); (iv) severe event costs; and, in an extended analysis, (v) productivity costs. Clinical events were captured by (i) response to treatment by ACR score and adverse event discontinuations; (ii) mortality (i.e., HAQ as a predictor or mortality); and (iii) utilities. Utilities were captured by mapping from the HAQ score. The model was subject to one way sensitivity analysis and probabilistic sensitivity analyses together with multiple scenario analyses.

Model Outcomes

The base case model results for (i) total cost; (ii) proportion of drug costs to total costs; (iii) life years; (iv) QALYs and (v) QALY days gained vs. average QALY are presented in Table 1 for each assumed TIM initiated pathway. Simulated (i.e., imaginary) lifetime total pathway costs ranged from $424,674 (tocilizumab iv) to $583,449 (etanercept). The TIM average ratio of drug to total costs was 79.50% (range 77.55 to 80.77%). Simulated average life years lived were 16.82 years (range 16.78 – 16.94 years [or 58 days]) versus 16.16 years for cDMARDs. Simulated average TIM pathway QALYs were 12.79 (range 12.69 – 13.12 QALYs [or 157 quality adjusted days]) versus 10.69 QALYs for cDMARDs.

Willingness-to-Pay Threshold Analysis

This similarity in QALY outcomes for the various TIM initiated treatment pathways is to be expected from the model structure and, in the second and third therapy sequences, the bundling together of TIMs by class. Even so, there are substantial differences between the TIMs in both their WAC price and their assumed net discounted prices. This leads to different estimates of the unit prices required to achieve $50,000, $100,000 and $150,00o QALY thresholds together with estimates of the range of discounts from WAC to achieve these thresholds. The base case results for this analysis are achieved through adjusting the respective ICER pairwise comparisons between the results for the individual TIM therapy pathway with the cDMARD reference treatment pathway. As an example, in the case of adalimumab, the market leader, the current WAC price per unit is $2,220.62 and the net price $1,554.43. To achieve the willingness-to-pay thresholds ICER would recommend a unit price of $373.06 for the $50,000 threshold, $699.49 for the $100,000 threshold and $1,010.38 for the $150,000 threshold. These translate to discounts from WAC in the range 55% to 83%.

Discussion

The principal objective of the ICER modeling is to generate recommendations to support discounting of drug costs by manufacturers to ensure that they fall within arbitrary willingness-to-pay cost per QALY thresholds. To justify these discounts, the case rests upon a modeled lifetime cost-per-QALY simulation that projects from a short term evidence base with efficacy claims for the competing TIMs created by a network meta-analysis. A network analysis that rests principally on indirect comparisons, the outcomes of which are best seen as hypotheses themselves for comparative treatment effects across the TIM therapy options.

From the perspective of the standards of normal science, the principal objection to claims for the impact and costs of competing therapy interventions from the ICER and similar models is that the claims made for treatment pathway outcomes are immune to failure [17]. This is the inevitable result of employing a modeled lifetime perspective with an assumed treatment pathway sequence of fixed interval therapy switching. The reader has no idea whether the claims for the costs of treatment, years lived or QALYs are right or even if they are wrong; indeed, the reader will never know. As such, despite claims that they may, in some sense, be sufficiently representative of the model builders' perception a present and future reality to be taken seriously, they are entirely imaginary constructs. They fail to meet standards for generating evaluable hypotheses to establish product-specific value claims that can be assessed and replicated in target treating populations.

As argued in previous commentaries in this series, non-evaluable claims generated by lifetime cost-utility or reference case models, rejects experimentation through hypothesis testing or systematic observation in favor of relativism and the acceptance of constructed evidence [18]. The 'gold standard' of experimentation has been accepted for the last 350 years,

exemplified in the motto of the Royal Society (founded 1660; Royal Charter 1662): *nullius in verba* (take no man's word for it). As stated on the Royal Society website, this motto 'is an expression of the determination of Fellows to withstand the domination of authority and to verify all statements by an appeal to the facts determined by experiment' [19].

A Multiverse of Models and Modeled Scenarios

If we were to suspend disbelief in the ICER methodology and subscribe to the information value of evidence from constructing imaginary worlds, the question then becomes: from a potential multiverse of models and modeled scenarios which one do we chose? Are all models equally valid? Do we only choose models that meet arbitrary reference case standards? Do we only choose models that mandate a specific generic HRQoL measure? Do we only choose models that mandate a specific disease-specific HRQoL measure? Can we map HRQoL utilities from clinical markers? Do we only choose models that assume a minimum sequence of options in therapy sequencing over a patient's lifetime? How do we justify future claims for clinical status when they are predicated on multiple sequential therapy switches for which there are no data? How do we accommodate long-term response to biologics in rheumatoid arthritis when the majority of patients have discontinued bDMARD therapy within two years? Can we justify claims for future costs of treatment when the possibility of manufacturers following a strategy of substantial annualized price increases is ignored? Does everyone subscribe to willingness-to-pay thresholds? Should our criterion be that if a modeled is judged 'sufficiently' realistic then the claims will necessarily follow? Why should formulary committees subscribe to claims generated by imaginary simulations?

We could, of course, brush these questions to one side and take the relativist position on constructed evidence. As long as the community of scholars in health technology assessment, including ICER, agree that reference case modeling is acceptable then why should we disagree? The latest CADTH guidelines, for example, are quite explicit in their rejection of the standards of normal science in favor of lifetime-cost-per-QALY claims: *Economic evaluations are designed to inform decisions. As such, they are distinct from conventional research activities, which are designed to test hypotheses* [20]. The latest version of these guidelines, released March 2017, has been reviewed as part of this commentary series [21]. The recommendation, which is unlikely to be accepted given the investment in imaginary worlds by professional groups and agencies such as CADTH is to abandon a relativist position and focus on evaluable and replicable clinical and cost-outcomes claims.

In any event, whether we agree or disagree that technology assessments should focus on hypothesis testing, it is worth considering a number of the assumptions that underpin the ICER model structure and claims for lifetime cost-effectiveness of the TIMs. This does not endorse the construction of imaginary worlds. Rather, it points to the flexibility we always have in building imaginary worlds and for ongoing disputes between competing model builders [22].

The Lifetime Horizon

As noted, the adoption of a model framework that tracks patients through a series of switching algorithms over the lifetime of a hypothetical cohort ensures that the claims made are immune to failure. At the same time, there is no account of the potential impact of products, which may now be in phase 2 or phase 3 of product development, on prospective therapy switching and persistence with current therapies. As this would, presumably, admit of a range of market entry scenarios, pricing and switching options, it would be foolhardy to attempt such as exercise.

The model builder is faced with two options, allow simplified market entry of competitor scenario(s) or admit that this possibility has been discounted. Neither are likely to appeal to decision makers. A more reasonable approach would be to admit the potential for new entrants and focus on short term credible and evaluable claims. Support for these claims would potentially put existing products in a stronger evidentiary position for continued formulary placement and patient/physician acceptance. New entrants would have to demonstrate, again with credible and evaluable claims, their comparative benefits and harms to those already accepted as integral to treatment pathway decisions.

The fact that agencies such as NICE in the UK and the PBAC in Australia, among others, have adopted a reference case lifetime cost-per-QAY model that excludes consideration of from, in this case the pipeline of new compounds for target rheumatoid arthritis patients, is no reason to accept uncritically the same reference framework for modeling claims in the US market - let alone the underlying acceptance of a lifetime modeling framework [23]. It would be of interest to ask ICER how they justify their position on the prospective entry of new compounds given their commitment to lifetime cost-utility models.

Health Related Quality of Life

Claims for therapy options couched in terms of quality of life as a single metric that combines morbidity and mortality have an obvious appeal. The downsides are the questions of the choice and acceptability of the particular instrument together with the ability to map or crosswalk from clinical scores or patient reported outcomes (PROs) to an agreed standard score [24]. Unlike NICE in the UK there is no standard for QALY measurement in the US to support product claims nor, more to the point, is there any interest is collecting QALY measures to validate modeled QALY claims by the overwhelming majority of health care systems.

ICER does not admit a standard for QALY measurement, either in terms of a measure's ability to cover generically a range of disease states or its psychometric properties such as Rasch interval measurement. In multiple sclerosis the ICER lifetime model relies upon an algorithm linking ACR status to utilities; in the NICE model the link is to EULAR status.

It is of interest to note that the latest (March 2017) version of the Canadian guidelines for health technology assessment cautions against accepting uncritically mapping algorithms as their 'predictive value can vary dramatically depending on instruments being mapped, the algorithm being used and the severity of the health states included'. At the same time, it is also worth noting the recent experience with attempting to move from the EQ-5D-3L to the purportedly more sensitive EQ-5D-5L version. Unfortunately, the application of the EQ-5D is now in doubt as the attempt to move to a more sensitive version of the instrument with five levels within each health domain (EQ-5D-5L) as opposed to the three levels (EQ-5D-3L) has led to a situation where these two versions produce substantially different estimates of cost-effectiveness [25]. This is because of the combined effect of differences in the way individuals respond to the changed descriptive system and the changed valuation system in the 5L compared to the 3L. The two versions are not consistent with each other. This problem is likely to plague NICE for years given the number of accepted modeled evaluations and claims for cost-per-QALY outcomes based on the EQ-5D-3L. If ICER, for example, used the EQ-5D-3L as the multi-attribute utility instrument in accepting the ACR mapping algorithm, revising the algorithm to generate EQ-5D-5L utilities may result in quite different utility scores and cost-per-QALY projections and price discounting recommendations.

The issue is further compounded by range of available possible measures (time-trade off, standard gamble) as well as generic multi-attribute instruments (EQ-5D-3L, EQ-5D-5L, HUI, ST-6D) and algorithms that have been proposed to map or crosswalk items from disease specific patient reported outcomes (PRO) instruments or clinical markers to generic instrument standard. Aside, of course, from the possibility that directly measuring EQ-5D-5L utilities from a pragmatic TIM sequencing trial may generate a different set of utility profiles that may result from crosswalking within an EQ-5D-3L framework. Again, there are no evaluable claims from the ICER model that would enable health system decision makers to assess the contribution of alternative QALY constructs or manufacturers to challenge the claims. Claims for QALYs for the various TIM pathways are non-starters for comparative validation. In challenging ICER, one can presumably choose one's own instrument.

Comorbidities and Symptoms

Lifetime modeled claims for competing therapy interventions typically ignore the presence of comorbidities in modeling outcomes in target treating populations. This is a significant oversight as the presence of comorbidities, typically present in older populations such as the ICER modeled cohort in multiple sclerosis, can have a potentially significant impact not only on the patient's quality of life but also on outcomes achieved in the target disease state. The importance of managing comorbidities in rheumatoid arthritis is seen in the Canadian Dermatology-Rheumatology Comorbidity Initiative where the point is made that comorbidities such as cardiovascular disease can 'contribute to increased early mortality, affect disease activity and response to treatments, and generate costs in these patients' [26].

This situation is further compounded by the presence of symptoms such as fatigue and pain, together with possible depression and anxiety, which may characterize older populations. In this context it is worth noting a the systematic review and network meta-analysis reported by Jansen et al of patient reported outcomes (PROs) in patients with an inadequate response to cDMARDs. This review captured a range of outcomes including pain, self-reported disease activity, functional ability, physical and mental health and fatigue within the different classes of bDMARD therapies [27]. The downside for those who subscribe to the construction of imaginary lifetime modeled claims is that attempts to introduce comorbidities into a model yield even more scenarios to be considered. At the same time, to put comorbidities to one side makes the construction of imaginary scenarios even less attractive to the prospective health system audience. It would, of course, be possible to factor out only those costs directly attributable to rheumatoid arthritis. However, this raises the further question of the choice of technique to partition these costs.

Adherence and Persistence

Making assumptions regarding lifetime adherence and persistence with therapy and attempting to draw conclusions regarding possible differences with competing therapies seems a particularly pointless exercise if the assumptions generate compliance profiles that are at variance with observed behavior. In the ICER model case the assumptions relate not to compliance with the initial therapy, but to compliance over the potential three sequential TMI therapies in the specific pathways. In the ICER model patients discontinued therapy if they received an ACR score < 20 in the first six month cycle. Beyond six months, discontinuation was due to the occurrence of adverse events. As the cycle length was six months this presumably occurred at the end of a further six months of therapy following an earlier discontinuation. After three TIMs the patient is assumed to revert to cDMARD therapy for the rest of their life. Patients are apparently not allowed to switch to a cDMARD or other palliative care before experiencing at least 18 months of TIMs.

Unfortunately, and this is acknowledged in the ICER model, data for long term persistence with therapy, evidence for different patterns of adherence and persistence between the

various bDMARDS and evidence for the causes of discontinuation and therapy switching are limited, if not non-existent. In a recent paper, Souto et al go some way towards a more complete picture of the prevalence, causes and predictors of discontinuation with biologic therapies from a systematic review and meta-analysis of drug registries and health care data bases [28]. In total studies with > 200,000 patients were included in the analysis, 81 of which included TNF inhibitors, 14 all biologics and 3 abatacept. There were no data on rituximab. Overall, 63 European and 35 non-European registries were included. The patterns observed pointed to (i) substantial differences in persistence behavior between the individual TNK inhibitors (etanercept, infliximab, adalimumab, monoclonal antibodies) and (ii) substantial variation within each of the therapies. At six months, for example, discontinuation of all TNF inhibitors was 21% (range 14 to 28%), 27% at 1 year (range 23 – 32%0 and at 2 years 37% (range 35 – 40%). In the case of adalimumab, to give one example again, at six months discontinuation was 33% (range 32-35%) and at 2 years 42% (37 to 47%).

Also noteworthy, is the observation that discontinuation of individual TNF inhibitors was significantly higher in patients treated after 2005. There was considerable heterogeneity in the studies in the contribution of lack of efficacy versus adverse events. Predictors of time to discontinuation, for example, showed lower discontinuation for any cause in patients treated with etanercept versus infliximab or adalimumab. Lower discontinuation was associated with concomitant use of cDMARDs. Disease duration predicted higher discontinuation for adverse events but not for lack of efficacy or any cause. Female sex predicted higher discontinuation for any cause.

Given the contribution of drug costs to overall assessments of cost-effectiveness among competing TIMs, it would be far more useful from a formulary decision perspective to avoid modeled claims that attempt to capture persistence and switching behavior over the lifetimes of a hypothetical cohort of patients, focusing instead on short-term models where claims for persistence and the relative contributions of adverse events and efficacy can be evaluated for competing therapies. Presumably, manufacturers would want to put forward claims for fewer adverse events and improved efficacy to support value claims for persistence with their products. With the observed heterogeneity observed in persistence patterns, it would seem more useful to focus on shorter term impacts that can be evaluated and replicated in target populations, to include the contribution of polytherapy due to comorbid conditions, rather than make unsupported assumptions regarding long term persistence in a range of treatment pathways when the majority of patients may have abandoned bDMARD therapies within 2 to 3 years.. Manufacturers could also be encouraged to propose intervention and patient management strategies to support their product and standards for the choice of successor therapies.

Drug Prices

The ICER evidence report recognizes that the prices of TIMS have risen substantially. The evidence report cites two TIMS with the leading market share, adalimumab and etanercept, having risen in price by 70-80% in last 3 years which, even allowing for discounts, rebates or patient assistant programs is still substantial. Given this, it seems odd that ICER has not apparently attempted to build into the model assumptions regarding expected WAC price increases, together with projected price increases for the other cost components, for the individual TIMs over the anticipated treatment lifetime. After all, if prices increase for selected TIMs at 15% per annum, compounded over five years the WAC price will have almost doubled. A situation that is further compounded by the possibility that annualized price increases may vary between the respective TIMs.

Model Validation

The ICER evidence review also reports on steps taken to validate the model. This is achieved, apart from internally evaluating mathematical functions, by comparing the model to other models which were similar in hypothetical populations, setting, perspective and treatments. In effect, comparing one imaginary construct to other imaginary constructs; one set of facts with alternative facts. At best, in the absence of any commitment to generating and testing evaluable claims, this serves to demonstrate the inadvisability of actually committing to the construction of imaginary worlds: models generate different outcomes because the models are, well, different. Presumably this sets the stage for even more imaginary lifetime rheumatoid arthritis constructs; a plethora of imaginary worlds with non-evaluable claims. Of course, this is in the pharmacoeconomic tradition with literally thousands of modeled non-evaluable therapy claims published in the leading health care technology journals over the past 30 years [29] [30] [31].

An Alternative Perspective

Although this is not an endorsement of the ICER model, it is possible with estimates provided in the evidence review for treatment pathway costs and outcomes to put a quite different perspective on drug pricing and recommendations for discounting. As detailed above, Table 1 presents the base-case ICER modeled estimates of TIM pathway lifetime treatment costs by initial second line therapy for those who have failed prior cDMARD therapy. However, rather than considering the implications for drug pricing by constructing ICERs referenced by the cDMARD therapy pathway, we can consider a comparison, not of incremental costs and QALY gain over cDMARD but on the projected lifetime costs of each of the TIM treatment pathways compared to lifetime QALYs. This yields a cost-per-QALY projection for each pathway. The results are presented in Table 1 as (i) costs per life year; (ii) costs per QALY and (iii) costs-per-QALY differences from the average cost-per-QALY for the 12 TIM pathways.

From this perspective, a formulary committee has to judge whether it is prepared to pay, on average, $39,682 (at current prices) for each QALY. Rather than the somewhat artificial and self-serving comparison (from a price discounting perspective) of the additional cost-per-QALY gained over traditional cDMARDs, where the result is a foregone conclusion. This more pragmatic perspective considers what it will cost in life years and quality of life to introduce those failing or non-responsive to cDMARDs to any one of the TIM pathways. In this case, the differences in cost-per-QALY compared to the average (or to any other reference TIM such as adalimumab) range from a 'gain' of $4,788 for etanercept to a 'loss' with tocilizumab (sc) of $6,350. For those who subscribe to willingness-to-pay thresholds, all of the TIMs come in under $50,000 per QALY at current prices. If we accept the results of the ICER model then the choice in cost-per-QALY terms is the tocilizumab (sc) treatment pathway; total projected cost (subject to the caveats considered above) is the lowest, while the life-years and QALYs are essentially no different from the competing TIM pathways.

Conclusions

It is not the intent here to suggest how the ICER model can be 'improved'. The position taken is that in emulating NICE, in focusing on lifetime cost per QALY models as the standard for evaluating competing therapies and making recommendations for price discounting, the ICER approach is simply misplaced; it is a pointless exercise. The construction of lifetime cost-per-QALY models is unacceptable, if we subscribe to the standards of normal science, because the claims are unevaluable. More to the point: they were never intended to be evaluated. However much we may tinker with the core model through the construction of alternative scenarios, assumptions about long term therapy choice and response to therapy, the methodology is insupportable.

Rather than attempting to inform decision makes through the construction of imaginary worlds, price negotiations should be predicated on evidence that meets the standards of normal science. Unless evaluable and replicable claims are presented by ICER to support recommendations for price discounting, the recommendations should be rejected. This applies not only to the current recommendations for price discounting of TIMs, but to other ICER evidence reviews that have generated non-evaluable claims. If we reject relativism in favor of the standards of normal science, the only acceptable claims are those that are credible, evaluable and replicable.

Table 1
ICER Multiple Sclerosis Model: Costs per QALY and Incremental QALYs for Base Case Targeted Immune Modulator Treatment Pathway Simulation

TMI: Treatment Pathway	Simulation: Total Cost ($)	Drug Costs (%)	Simulation: Life Years	Simulation: QALYs	QALY Days Difference vs. Average QALY Days	Cost per Life Year ($)	Cost per QALY ($)	Cost per QALY: vs Average Cost per QALY ($)
Rituximab	464,864	78.90	16.79	12.70	-32.54	27,687	36,603	-3,078
Abatacept (iv)	466,733	78.79	16.82	12.78	-3.34	27,749	36,520	-3,162
Abatacept (sc)	566,053	79.90	16.87	12.90	0.45	33,554	43,880	4,198
Tocilizumab (iv)	470,205	78.66	16.85	12.88	33.15	27,905	36,506	3,176
Tocilizumab (sc)*	424,674	77.55	16.83	12.81	7.06	25,233	33,152	6,530
Tofacitinib	579,140	80.77	16.78	12.57	-80.00	34,493	46,073	-6,391
Adalimumab	530,720	80.20	16.78	12.68	-39.85	31,628	41,855	-2,172
Certolizumab pegol	522,473	79.95	16.84	12.86	25.85	31,025	40,628	-946
Etanercept	583,449	80.56	16.94	13.12	120.75	34,442	44,470	-4,788
Golimumab (sc)	512,875	79.63	16.79	12.69	-36.20	30,547	36,632	3,050
Golimumab (iv)	488,380	79.24	16.81	12.75	-14.30	29,053	38,304	1,378
Infliximab	480,448	79.54	16.79	12.73	-21.60	28,615	37,741	1,941
Average	507,501	79.50	16.82	12.79		30,164	39,682	
cDMARD	67,819	26.85	16.16	10.69	2.10*	4,197	6,344	-33,338**

Note: Treatment pathway as defined in ICER model following previous failure on cDMARDs and initiated with combination therapy (MTX + bDMARD/TIM). The simulated life years/QALYs together with simulated total costs and drug costs are from initiation of combination therapy for the treatment pathway over the projected lifetime of the hypothetical cohort. * versus average QALYs for TIMs. ** versus average cost per QALY for TIMs.
Source: ICER Evidence Review Table ES4

References

1 Institute for Clinical and Economic Review (ICER). Targeted Immune Modulators for rheumatoid Arthritis: Effectiveness and Value.: Evidence Report. April 2017

2 Langley PC. Cost-Effectiveness and Formulary Evaluation: Imaginary Worlds and Entresto Claims in Heart Failure. *Inov Pharm.* 2016;7(3): No. 6

3 Langley PC. Multiple Sclerosis and the Comparative Value Disease Modifying Therapy Report of the Institute for Clinical and Economic Review (ICER). *Inov Pharm.* 2017;8(1): No. 12

4 Langley PC. Intelligent Design and Imaginary Worlds in Cost-Effectiveness Claims: An Overview of Commentaries in INNOVATIONS in Pharmacy from July 2016 to February 2017. *Inov Pharm.* 2017;8(1): No. 22

5 Pigliucci M. Nonsense on stilts: How to tell science from bunk. Chicago: University of Chicago Press, 2010

6 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

7 University of Minnesota. 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

8 Sweet B, Tadlock CG, Waugh W, et al. The WellPoint outcomes based formulary: enhancing the health technology assessment process. *J Med Econ* 2005;8:13-25

9 Langley PC. Recent developments in the health technology assessment process in Fulda TR, Wertheimer AI, eds. Handbook of Pharmaceutical Public Policy. New York: Pharmaceutical Products Press, 2007

10 Stevenson M, Archer R, Tosh J et al. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. *Health Tech Assess.* 2016;20(35):1-610

11 Donahue K, Jonas D, Hansen R et al. Drug therapy for rheumatoid arthritis in adults: An Update. Comparative Effectiveness Review No. 55Rockville MD: Agency for Healthcare Research and Quality. April 2012

12 Singh J, Hossain A, Tanjong E et al. Biologics or tofacitinib for rheumatoid arthritis in incomplete responders to methotrexate or other traditional disease-modifying anti-rheumatic drugs: a systematic review and network meta-analysis. *Cochrane Database Syst Rev*. 2016(5)

13 Singh J, Hossain A, Tanjung E.. Biologic or tofacitinib monotherapy for rheumatoid arthritis in people with traditional disease-modifying anti-rheumatic drug (DMARD) failure: a Cochrane Systematic Review and network meta-analysis (NMA). *Cochrane Database Syst Rev*. 2016(11)

14 Canadian Agency for Drugs and Technologies in Health. Clinical and Economic Overview: Biological Response Modifier Agents for Adults with Rheumatoid Arthritis. Ottawa: CADTH, 2010

16 National Institute for Health and Care Excellence (NICE). Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDS or after conventional DMARDs only have failed. Technology Appraisal Guidance (TA375). London: NICE, 2017

17 Langley PC. Supporting Formulary Decisions: The Discovery of New Facts or Constructed Evidence?. *Inov Pharm.* 2016;7(2): No. 15

18 Wootton D. The Discovery of Science. New York: Harper Collins, 2015

[19] https://royalsociety.org/about-us/history/

[20] Canadian Agency for Drugs and Technologies in Health (CADTH). Guidelines for the Economic Evaluation of Health Technologies: Canada (4th Ed.). Ottawa:March 2017

[21] Langley PC. True North: Building imaginary worlds with the revised Canadian (CADTH) Guidelines for Health Technology Assessment. *Inov Pharm*. 2017;8(2): No. 9

[22] Langley PC, Rhee TG. Imaginary Worlds: The Status of Simulation Modeling in Claims for Cost-Effectiveness in Diabetes Mellitus. *Inov Pharm.* 2016;7(2): No. 17.

[23] Langley PC. Sunlit Uplands: The Genius of the NICE Reference Case. *Inov Pharm.* 2016;7(2): No. 12

[24] Langley PC. Great Expectations: Cost-Utility Models as Decision Criteria. *Inov Pharm.* 2016;7(2): No. 14

[25] National Institute for Health and Care Excellence. Decision Support unit. Comparing the EQ-5D-3L and 5L versions. What are the implications for cost-effectiveness estimates? NICE: March, 2017
http://www.nicedsu.org.uk/USERIMAGES/DSU_3L%20to%205L%20FINAL.pdf

[26] Roubille C, Richer V, Starnino T et al. Evidence-based recommendations for the management of comorbidities in rheumatoid arthritis, psoriasis, and psoriatic arthritis: Expert opinion of the Canadian Dermatology-Rheumatology Comorbidity Initiative. *J Rheumatol*. 2015;42:1767-80

[27] Jansen J, Buckley F, Dejonckheere f et al. Comparative efficacy of biologics as monotherapy in combination with methotrexate on patient reported outcomes (PROs) in rheumatoid arthritis patients with an inadequate response to conventional DMARDS – a systematic review and meta-analysis. *Health Qual Life Outcomes*. 2014;12:102

[28] Souto A, Maneiro J, Gómez-Reino J. Rate of discontinuation and drug survival of biologic therapies in rheumatoid arthritis: a systematic review and meta-analysis of drug registries and health care databases. *Rheumatology*. 2016;55:523-534

[29] Langley PC. Imaginary Worlds: Modeled claims for cost-effectiveness published in PharmacoEconomics January 2015 to December 2015. *Inov Pharm.* 2016;7(2): No. 9

[30] Langley PC, Rhee T. Imaginary Worlds: A Systematic Review of the Status of Modeled Cost-Effectiveness Claims Published in the Journal of Medical Economics from January 2015 to December 2015. *Inov Pharm.* 2016;7(2): No. 16

[31] Langley PC, Rhee T. Imaginary Worlds: The Status of Modeled Economic Evaluation Claims Published in Value in Health January 2015 to December 2015. *Inov Pharm.* 2016;7(2): No. 18