MAIMON WORKING PAPER No. 29 DECEMBER 2020

ICER ANALYTICS: FAIR PRICES FOR FUN AND PROFIT

Paul C Langley Ph.D., Adjunct Professor, College of Pharmacy, University of Minnesota

Abstract

The release of the ICER Analytics platform presents an unanticipated opportunity to finally discredit the ICER modeling component of the regular product specific evidence reports. It has been argued on a number of occasions that the ICER modeling fails the standards of normal science, not only in rejecting the need for credible, evaluable and replicable claims, but in ailing to understand the limitations imposed by the axioms of fundamental measurement. Utility scales, whether generic or disease specific, are ordinal. They cannot be used to create quality adjusted life years (QALYs). This is usually ignored; but what cannot be denied is that the utility score essentially determines the ICER claim for a 'fair price'. The purpose of this note is to demolish the notion of a reference ICER 'fair price' in the US market. To do this we suspend belief in the impossible QALY. This allows QALY scenarios to be developed to demonstrate how multiattribute utility scores will almost certainly generate price discounting recommendations..

INTRODUCTION

It is not often that the opportunity is presented to discredit completely a widely accepted analytical framework for social pricing of goods and services. This opportunity is the release of the ICER Analytics cloud platform which allows the user to repopulate ICER product and disease specific lifetime simulation models with alternative assumptions to generate competing fair prices to those prosed by the ICER version of this analytical framework [1]. From a practical perspective this is an unsurprising outcome; changing assumptions will change model outcomes. From a technology assessment perspective the implications are more serious: who are we to believe? Does the ICER 'base' model which is an integral part of their product evidence report to be the 'gold' standard for fair price claims or are the multitude of possible alternative assumption driven models to be considered as equal contenders? Surprisingly, there is nothing in the choice of ICER 'base' model assumptions which necessarily put the ICER 'fair price' on a pedestal. Indeed, for those who believe in the approximate information generated by the ICER 'base' model, any modification of assumptions in presumably are an unnecessary distraction as it points to the possibility of a multitude of 'fair prices'. Indeed some of these may provide the basis for recommending price increases as well as the typically expected price discounts. The ability to manipulate assumptions to change outcomes, which may be manipulated to achieve a desired outcome, should be sufficient to convince health care decision makers to abandon introducing ICER 'fair prices' into contract negotiations.

CHOOSING THE UTILITY SCORE

The principal objective in lifetime cost-per QALY simulation modeling is to make the case for a 'fair price'. If there are additional concerns that, in the US market, drug and device prices are 'too high', then the simulation modelling framework is a potential tool to argue for a 'social' fair price and possibly associated price discounts from the wholesale acquisition cost (WAC) set by the manufacturer. The modeled 'social fair price' becomes a bargaining chip. The key element in the simulation are the modeled QALYs and the choice of utility score to create the impossible or I-QALY [2]. The ICER Analytics platform allows the choice of utility scores to be explored with the impact on incremental cost-per-QALY claims exhibited. In an important sense, however, the choice of utility instrument is a waste of time as the distances between the scores are unknown; this is after all an ordinal score [3]. However, we can proceed on the false assumption that the various utility scores are ratio scores which support the construction of QALYs.

Two options are open in the choice of utility score: (i) a generic multiattribute or similar score or (ii) a disease specific utility score. The distinction is critical because the generic score, with a limited number of symptoms and response levels, may not capture attributes of disease state experience that are relevant to patients (and caregivers). This will inevitably bias the utility score towards 'perfect health' with attempts to capture utilities for the various stages of a disease clustered towards this end of the scale. This is, in fact, the case with the EQ-5D-3L with the added issue of a floor effect around 0.40. The implication is that the utility gains from switching to a new from a new product are likely to be small, with the cost-per-QALY gain dominated by the assumed lifetime cost of the drug.

Choice of a disease specific measure of utilities will likely yield a more reasonable distribution of between perfect health and death. QALY gain are likely to be greater and, in consequence, less dominance by the cost of the drug. The social 'fair price' will be higher and less discounts demanded of manufacturers. Unfortunately there are few disease specific utility scores. If we are concerned to develop at least a first step then we could consider for each disease stage a 'lay person' description of this stage and then ask, with a visual analogue scale (VAS) a sample of patients (and caregivers) to value the 'utility' they would attach to that disease stage. The average of the scores (we are assuming the VAS is a ratio scale) would then yield 'scores' for our modelling. This may suffice for illustrative purposes with an audience of health decision makers, but would fail if the audience included measurement specialists. In this respect we should note the impossibility of actually creating a ratio utility score. At best, with Rasch measurement theory (RMT), we could create an interval scale to assess response to therapy [4]. This illustrates how far from reality the ICER modelling framework is in even practical terms.

REJECTING ICER 'FAIR PRICES"

As will be illustrated here, the ICER concept of a 'fair price' is a non-starter. To the economist a 'fair' price is a market clearing price. If markets are well organized and reasonably transparent (the efficient market hypothesis) a market price will approximate the fair value: a rational and unbiased estimate of

potential market price. Depending upon the degree to which a market diverges from the reference point of a perfect market, a number of pricing criteria can be considered given market conditions, e.g., degree of monopoly market power reflecting the monopolist's demand curve and cost structure. This has nothing to do with imaginary or I-QALYs embedded in models that fail the standards of normal science.

Even so, to illustrate the absurd nature of ICER 'fair price' claims, let's pit the I-QALY to one side and consider the impact of utility choice consider in the following example. The standard of care is drug B; the new competitor is drug A. Simulated lifetime costs are assumed at $250,000 and $1 million respectively. The benefits conferred by Drug A over Drug B are measured by time spent in each of four successive stages of disease. There is no increase in life expectancy (it can be assumed) but the time spent (in years) shows longer time in disease stage 1 for drug A (41 vs. 25 years).

Two utility scales are assumed: a generic scale and a disease specific scale. The former clusters utilities at the perfect health end of the scale; the disease specific is more reflective of disease experience with utilities substantially less by disease stage and more 'spread out'. The standard modeling approach is to propose a series of stages which a hypothetical patient is proposed to experience over the natural course of a disease. Each 'stage' is assumed to have a utility weight so that time spent can be re-imagined as the equivalent time in perfect health (QALYs). Thus, as the disease becomes worse, time spent in each stage yields fewer QALYs (it is best to avoid a discussion of what perfect health actually means).

TABLE 1 IMAGINARY QALY SIMULATION FOR SAME LIFE EXPECTANCY

	Number of Stages: A (= years)	Numb er of Stages: B (= years)-	Utility Score (0-1) Generic Multiattribute	Utility Score (0-1) (Disease specific)	QALYS: A: generic multiatt ribute	QALYS: B generic multiatt ribute	QALYS: A disease specific utility	QALYS: B disease specific utility
Stage 1	41	25	0.98	0.95	40.18	24.50	38.95	23.75
Stage 2	10	15	0.92	0.65	9.20	13.80	6.50	9.75
Stage 3	3	10	0.88	0.30	2.64	8.8	0.90	3.00
Stage 4	1	5	0.75	0.05	0.75	3.75	0.05	0.25
Total years	55	55						
Total QALYS					52.77	50.85	46.40	36.75
QALY Gain					1.92		9.65	

Although this is a trivial example is shows how sensitive the results are to the utility scores. In the case of the generic scores which show little difference in utility weights by stage of disease the number of QALYS from 55 years is 52.77 for the new drug A and 50.85 for the comparator standard of

care. The QALY difference is 1.92. With the disease specific utilities the results are markedly different with QALYs under drug A at 46.40 and the standard of care 36.75, a difference of 9.65 QALYs.

Assume that the cost of drug A over 55 years is $1.5 million dollars while the standard of care is $250,000. Then the cost per QALY for drug A is $28,463 and $4,916 for drug B. Incremental costs over the 55 years are $1,250,000 compared to 1.92 QALY gains or $657,895 per QALY. If we assume a mythical cost-per-incremental QALY threshold of $150,000 then for a 'social fair price' drug A would have to be discounted by $507,895 or 77%. ICER would claim victory.

If we consider the disease specific utilities then the picture changes completely: with an increment of 9.75 QALYs the cost per additional QALY drops to $1,250,000/9.75 or $128,205. This is below the $150,000 threshold and the company could justify a price increase of $21,795 over the 55 years or 17%.

At product entry, drug A could be reasonably put forward as one that extends life (gene therapies would fall into this category). Consider the results presented in Table 2 where for drug A life expectancy increases from 55 to 67 years. In this case the QALY gain with the generic utilities is 12.95 QALYs and 17.6 QALYs with the disease specific utilities. Again, with even an increased cost of drug A over an additional 12 years (increasing to $1.68 million) drug A, against a threshold of $150,000 is seriously underpriced.

TABLE 2 IMAGINARY QALY SIMULATION WITH DRUG A INCREASED LIFE EXPECTANCY

	Number of Stages: A (= years)	Number of Stages: B (= years)-	Utility Score (0-1) Generic Multiattribute	Utility Score (0-1) (Disease specific)	QALYS: A: generic multiattribute	QALYS: B generic multiattribute	QALYS: A disease specific utility	QALYS: B disease specific utility
Stage 1	45	25	0.98	0.95	44.10	24.50	42.75	23.75
Stage 2	15	15	0.92	0.65	13.80	13.80	9.75	9.75
Stage 3	5	10	0.88	0.30	4.40	8.8	1.50	3.00
Stage 4	2	5	0.75	0.05	1.50	3.75	0.10	0.25
Total years	67	55						
Total QALYS					63.80	50.85	54.10	36.75
QALY Gain					12.95		17.65	

Note: ignores axioms of fundamental measurement and assume utility a ratio scale

ICER SCENARIOS

Arguing that these 'differences' can be subsumed by scenario analysis is a non-starter. The standard ICER model assumes only one utility (generic) scale targeted to a model structure of disease stages which cannot be modified within ICER Analytics. While this can be seen as providing a common 'backbone' for the model assumptions 'variation' the fixed disease progression stages merely makes it easier to illustrate the manifest deficiencies in the ICER model as the common 'backbone' can support a variety of 'fair prices'. This goes beyond the ICER scenario presentations as it involves assessing the impact of changing the ordinal utility scores from 'generic' to 'disease specific'. Whether this was anticipated by ICER in its 'open season' policy, the impact illustrates the absurdity of modeled simulations driven by malleable assumptions to establish 'fair prices'. ICER would have been better advised to disallow utility modification.

CONCLUSIONS

ICER's 'status' as the US arbiter of drug and device 'fair prices' rests on the choice of utility score. Whether this was conscious or not, the choice of multiattribute instruments virtually guarantees a minimum QALY gain and, by definition, claims for WAC discounting. This is apparent from the ICER Analytics platform. The various product models do not allow any change in the disease stages; all that can be done is to vary the utility scores (and justify these) for each stage. This is the break point as it makes quite clear the importance of disease specific utilities. The further weakness is gene therapy. Even with multiattribute scores this creates significant utility gains as the model must acknowledge a possibly natural lifespan free of the disease and, in most cases, a much shorter time spent with substantial utility decrements.

REFERENCES

[1] Institute for Clinical and Economic Review. ICER Analytics Launched to Accelerate the Real World Application of ICER Reports, Helping Payers, the Lifde Science Industry and Health Care Policy Makers move from Insight to Action. Media Release. 1 December 2020. https://icer-review.org/announcements/icer-analytics-launch

[2] Langley P. The Great I-QALY Disaster. *InovPharm*. 2020;11(3): No. 7 https://pubs.lib.umn.edu/index.php/innovations/article/view/3359/2517

[3] Langley PC, McKenna SP. Measurement, modeling and QALYs [version 1; peer reviewed] *F1000Research* 2020, 9:1048 https://doi.org/10.12688/f1000research.25039.1

[4] Bond T, Fox C. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 3rd Ed. New York: Routledge, 2015