DAMAGED GOODS: IMAGINARY ICER VALUE ASSESSMENTS AND THE DEPARTMENT OF VETERANS AFFAIRS

According to a June 15, 2020 *Health Affairs* blog, authored by representatives of the Department of Veterans Affairs (VA) and the Institute for Clinical and Economic Review (ICER), concerns initially raised by critics have been successfully addressed. There is now a mutually reinforcing collaboration, supported by monthly meetings, that the partners believe has added significantly to the data available to the VA for formulary decisions. Without wishing to unduly unsettle what appears to be a cozy and mutually satisfying relationship, it should be pointed out that the ICER value assessment framework not only fails the standards of normal science, but that the pièce de résistance – the Quality Adjusted Life Year (QALY), is a mathematically impossible construct. The implications are serious, casting doubt on whether there are any benefits resulting from this mutually attractive engagement.

Introduction

In common with other formulary assessments, the VA Pharmacy Benefits Management Services (PBMS) presumably undertakes a systematic review of the available clinical evidence for new products, in addition to ongoing disease area and therapeutic class reviews. Systematic reviews recognize the primacy of normal science in the testing of competing hypotheses as to treatment effects and the discovery of new facts. A provisional consensus can be achieved for the potential place of products in therapy – subject to quality assessments for the merits of individual studies, and in particular the validity of study protocols for the VA target patient populations within disease areas.

Unfortunately, the VA PBMS has introduced ICER value assessments into the evaluation process. As the June 15 *Health Affairs* blog informs us: "….missing from these information resources (pharmaceuticals) was a reliable source of non-biased cost-effectiveness analysis …. The opportunity to contract with the Institute for Clinical and Economic Review (ICER) was seen … as a means to better contextualize value for decision making on behalf of veteran patients and US taxpayers."

Three years into the partnership, no one has yet apparently seen the need to consider the weaknesses, indeed, the fatal flaws, in the ICER value assessment paradigm.[1] The VA has taken ICER recommendations at face value without a more extensive assessment of the merits of this paradigm.

Abandoning Normal Science

The VA, in accepting ICER, has accepted that the role of health technology assessment is not to test hypotheses and develop protocols and propose evidence platforms for the discovery of

new facts in treatment options for the patient population, but to rely on imaginary simulations to generate imaginary evidence.[2] In previous reviews of the ICER paradigm, going back now 5 years, the key criticism of the ICER approach is that it fails to meet the demarcation test between science and pseudoscience. There is a complete absence of credible, evaluable, and replicable cost-effectiveness claims. The hallmark of the scientific method is measurement and hypothesis testing. The VA – in contracting with ICER for their 'pivotal' imaginary claims for cost-effectiveness – has cast this hallmark aside.

Rather, in embracing ICER, the VA has devoted the last three years to factoring in value assessments based on the ICER cost-per-incremental QALY reference case. This model generates lifetime modeled assumption-driven scenarios for quality adjusted time spent in disease stages, with further assumptions regarding lifetime direct medical costs to create evidence for incremental QALY costs. These are matched to cost-per-QALY thresholds with consequent recommendations for pricing and access.

The VA has utilized the evidence created by ICER's imaginary worlds in a variety of ways, hastening to point out that ICER's input is only one element within the formulary decision and price determination process. The VA has utilized this created evidence in price negotiations and in the choice of products for the VA patient population. At the same time, the VA is at pains to point out that it has not used negative ICER assessments of value for new products to limit their use. Instead, the VA points to the use of prior authorization for high priced products.

Assumptions

While this apparently cautious use of ICER modeled recommendations is laudable, the fact is that the VA incorporates into its decision making an analytical framework that fails the standards of normal science. Driven by assumptions plucked from trial data and the wider literature, the VA and ICER have fallen into the trap of what is known as Hume's problem of induction (David Hume 1704 – 1776): basing general statements (e.g., assumptions) on accumulated observations. The problem with ICER's approach is that you cannot construct future scenarios on past observations. The ICER value assessment framework is illogical since it cannot be established from the fact that all past futures have resembled past pasts it does not follow that all future futures will resemble future pasts.[3] Assumption-driven futures with no credible claims for evaluation are logically nonsensical. The VA has, apparently, opted to introduce pseudoscientific claims into its formulary evaluation process: this additional information resource is seen to provide an objective transparent standard to guide its price negotiations. While apparently the results have not undermined in any way the clinical focus of the VA drug coverage process, it is an open question as to whether the factoring in of ICER recommendations has been in the best interests of VA patients.

The Impossible QALY

Clearly, the VA will not respond well to being told that the VA-ICER collaboration is questionable. But there is a more fundamental and fatal objection to this collaboration: the failure of both parties to recognize the limitations placed on the ICER value assessment by the axioms of fundamental measurement. As detailed in previous Current Issues reports the ICER use of QALYS has no basis in reality. The favored utility value, the multi-attribute EQ-5D-3L, is an ordinal score. That is, the utility values have magnitude, an ordered relationship to each other but an unknown distance between the ranks. The ordinal scale can only generate medians and models, supporting nonparametric statistics (e.g., correlations). What the utility values lack are interval properties, where there are equal intervals, supporting mathematical operations of addition and subtraction and, at a higher order,

ratio properties. A scale with ratio properties has interval properties but, with a true zero, can also support multiplication and division. If you want to create a QALY then you have to have a ratio scale. The EQ-5D-3L demonstrably fails on this criterion as the utility values can range from -0.59 to 1.0. The former utility (-0.59) indicates, apparently, a state worse than death (death is arbitrarily assigned zero) with 1.0 perfect health. Thus, we can have negative QALYs. This is a common feature of all multi-attribute utility instruments.

More to the point, the incontrovertible fact is that that the EQ-5D-3L has only ordinal properties. This means that creating a QALY is mathematically impossible. It is based on symptom levels ranked by degree of impact, ordinal scores; aggregating over these with preference weights also creates an ordinal score. The frustrating aspect of this failure to recognize the axioms of fundamental measurement is that there were ample warnings in the literature. Following the seminal work of Stevens in formalizing measurement theory in the 1940s, it has become even more obvious that if you want to measure an attribute, such as needs-based quality of life, then the instrument has to be developed to meet the required measurement standards; specifically Rasch measurement theory.

This is commonplace in the physical sciences: measurement precedes hypothesis testing. A fact that ICER the VA and others appear to have either ignored or failed to recognize in creating multi-attribute utilities.

Reliable Data

If the VA thought it could rely on ICER to provide "reliable data" then it was misinformed. What the VA was buying were modeled imaginary scenarios that were intended to provide, given the position taken by groups such as the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), **approximate information**. No thought was given to the lack of credible and evaluable claims for competing products in the ICER value assessment, the failure to conform to the standards of normal science and, most egregiously, a failure to recognize that the QALY was an impossible construct. This was not even approximate information; it was impossible information. It is ridiculous even to attempt to apply some measure of "reliability" to the ICER value assessment. Take away the mathematically impossible QALY and the entire value assessment edifice collapses. To describe these as damaged goods would be an understatement.

More Believers

The extent to which health system decision makers include ICER recommendations in formulary decisions is debatable. A few recent surveys of formulary committees would suggest an uptick, although it is far from clear that these decision makers, like the VA, have any idea of the basis on which the ICER recommendations are created. Accepting an ICER recommendation does not mean, or even imply, that the recipient has any idea of the flawed methodology.

According to a link from the health affairs blog there are other who have swallowed the ICER value assessment line.[4] Apart from the lack of any awareness of the fatal flaws in the ICER value assessment framework, a survey of 31 payers (out of 400 stakeholders) in July 2019 which, given the response rate should not be taken seriously, reports that some 48.4% believed there was a place for an independent health technology assessment body in the US. Overall, 58.1% were familiar with ICER reports, processes, and assessments. A total of 54.8 % of the 31 respondents were likely and 9.7% extremely likely to follow an ICER cost-effectiveness threshold. ICER reports were mostly used to provide references for literature reviews (61%) with 58% using a report to inform choices for preferred products and the same percentage

using the reports to inform prior authorization and step edit criteria.

While this report is hardly a ringing endorsement of the awareness of ICER's presence in the world of technology assessment, care has to be taken to ensure, unlike the VA, that they understand what they are buying into. The ICER value assessment models have a plausible if superficial attraction, yet few appear to go beyond its recommendations to understand that it is actually a fatally flawed product.

Unfortunately, the VA through the PBMS is not alone in subscribing to the ICER story. A recent paper pointed to the lack of appreciation of ICER's limitations by the National Pharmaceutical Council.[5]

Conclusions

Beliefs in the mysteries of imaginary worlds, a phenomenon Richard Dawkins describes as a mind virus,[6] have been addressed in previous reports. A key feature is what has been described as the ISPOR meme: a coherent, strongly held belief with high transmission fidelity that supports the ongoing commitment in technology assessment to creating approximate information and rejecting the axioms of fundamental measurement; rather than the hard slog of developing measurement tools and creating evidence platforms for hypothesis testing.

Certainly, the VA can continue its collaboration with ICER, although it is difficult to see what role in formulary decisions the production of impossible information can play. If the VA is concerned with the cost-effectiveness of competing products, its resources might be better employed to generate real world and not imaginary evidence to support comparative assessments. Building imaginary worlds is a trivial and ultimately worthless exercise.

Paul C. Langley, Ph.D
Adjunct Professor, College of Pharmacy
University of Minnesota
paul@accessandaffordability.org.

REFERENCES

[1] Langley P. Nonsense on Stilts – Part 1: The ICER 2020-2024 value assessment framework for constructing imaginary worlds, *InovPharm*. 2020;11(1): No. 12
[2] Neumann P. Willke R, Garrison L. A health economics approach to US value assessment frameworks – Introduction: An ISPOR Special Task Force Report (1). *Value Health*. 2018;21:119-23
[3] Magee B. Popper. London: Fontana, 1973
[4] ICON. Payer reliance on ICER assessments and perceptions on value-based pricing. September 2019.
[5] Langley P. The National Pharmaceutical Council: Endorsing the construction of imaginary worlds in health technology assessment. *Pharmacy*. 2020;8(119);doi:10.3390/pharmacy8030119
[6]Dawkins R. A Devil's Chaplain. New York: Houghton Miflin, 2004