January 31, 2023

The Honorable Brett Guthrie
Chairman
Health Subcommittee
Washington, DC 20515

The Honorable Anna Eshoo
Ranking Member
Health Subcommittee
Washington, DC 20515

The Honorable Cathy McMorris Rodgers
Chairwoman
Energy and Commerce Committee
Washington, DC 20515

The Honorable Frank Pallone
Ranking Member
Energy and Commerce Committee
Washington, DC 20515

Re: Statement Submission for Congressional Record on the Ban of QALYs in all government programs

Dear Chairman Guthrie, Ranking Member Eshoo, Chairwoman Rodgers, and Ranking Member Pallone,

As a health economist with links to pharmacy colleges in the US, I am impressed with the almost universal lack of understanding of measurement theory when it comes to arguments for or against quality adjusted life year (QALY) scores in health technology assessment (HTA). The QALY is a mathematical impossibility, yet it has a strange fascination for those with a limited understanding (or awareness) of modern measurement theory. The QALY debates are, clearly, a waste of time; the QALY must be discarded, not for issues such as disability, but for the incontrovertible fact that it is an impossible measure [1].

We have accepted for some 60 or more years that if a measure of response is required it must be a single attribute, unidimensional, interval measure [2] [3]. Nothing else will do. It must meet the standards of Rasch measurement, and the issue is straightforward: observations produce ordinal scales. To assess therapy response, we need an interval (or ratio) measure. This can only be achieved by applying Rasch rules to the ordinal counts of observations to transform them to interval measures. The point is made in a paper by Wight and Linacre in 1989 published before the QALY was developed to become a mainstay of HTA in the mid 1990s: *Quantitative observations are based on counting observed events or levels of performance. Meaningful measurement is based on the arithmetical properties of interval scales. The Rasch measurement model provides the necessary and sufficient means to transform ordinal counts into linear measures* [4].

Rasch measurement has been accepted for over 60 years as the basis for creating single attribute, unidimensional, linear, interval (as well as ratio) measures in the scores in the social sciences; it has just been ignored for over 30 years. Those advocating QALYs such as the Institute for Clinical

and Economic Review (ICER) with their assumption driven modelled simulations to support imaginary non-empirically evaluable claims just perpetuate the QALY myth. The issue is the failure of those creating multiattribute generic preference scores to understand the limitations of fundamental measurement. The preference scores are ordinal scales as sums over questionnaire items. There are, in fact, several software packages (RUMM2030, WINSTEPS) that undertake a Rasch assessment of patient reported outcomes and support interval scores. They were first introduced in the 1980s and are used extensively in the social sciences.

The QALY multiplies time spent in a disease state with a preference scale (range 0 to 1; but the algorithms create negative scores). This fails because: (i) no one thought they needed, not just an interval but a ratio scale (a measure with a true zero) to support multiplication; and (ii) that the scale had to represent a single unidimensional attribute not a bundle of symptoms and response levels. Typically applied are ordinal preference scores from a multiattribute instrument such as the EQ-5D-5L. This is where it all falls apart. The EQ-5D-5L as a multiattribute score fails the requirements of Rasch modelling for subjective responses. The entire exercise is a waste of time and resources.

The result, unfortunately, is that HTA is locked into a belief system that is unique among the physical and social science disciplines: putting to one side value claims that meet the standards of normal science for credibility, empirical evaluation, and replication together with a failure to recognize the importance of Rasch measurement [5]. Instead, HTA rests on the simulated creation of non-evaluable claims for pricing and product access which would be rejected out of hand in other disciplines.

Although the criticisms presented here have been voiced over the past 30 years, together with instruments developed in many disease areas to capture response in terms of Rasch requirements, they are largely ignored. The reason is obvious: a dominant belief system or meme that excludes criticism. Truth is consensus. I have no doubt the good ship QALY will sail on; too many people have too much to lose to make clear they were wrong. It is perhaps asking too much for the government to step in and make clear why it will no longer support the QALY in any program.

Sincerely,

Paul C. Langley, Ph.D., Adjunct Professor, College of Pharmacy University of Minnesota, Minneapolis, MN; Director, Maimon Research LLC, Tucson AZ

Contact: Email langley@maimonresearch.com
Tel: (520) 577-0436

References

[1] Langley PC and McKenna SP. Measurement, modeling and QALYs [version 1; peer review: 2 approved]. *F1000Research* 2020, **9**:1048 (https://doi.org/10.12688/f1000research.25039.1)

[2] Bond T, Yan Z, Heene M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences (4th Ed.). New York: Routledge, 2021

[3] Andrich D, Marais I. A Course in Rasch Measurement Theory: Measuring the Educational, social and Health Sciences. Singapore, Springer: 2019

[4] Wright B, Linacre J. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989; 70(12):857-60

[5] Langley P. Nothing to Cheer About: Endorsing Imaginary Economic Evaluations and Value Claims with CHEERS 22 [version 1; peer review: 2 approved]. *F1000Research* 2022, **11**:248 (https://doi.org/10.12688/f1000research.109389.1)