Determination of the Ideal Number of Response Levels for the EQ-5D as a US Population Health Measure

Jonathan L. Nazari, PharmD, Maja Kuharic, MPharm, MSc, A. Simon Pickard, PhD
University of Illinois at Chicago, Department of Pharmacy Systems, Outcomes, and Policy

UNIVERSITY OF ILLINOIS CHICAGO
PHARMACY SYSTEMS OUTCOMES AND POLICY
COLLEGE OF PHARMACY

Objectives

- The expansion in descriptive system from the EQ-5D-3L (3L) to the EQ-5D-5L (5L) was largely motivated by an interest in greater discriminative ability as a measure of health[1]. However, the adequacy of the 5L descriptive system in the context of a population health measure has received limited attention.
- The primary aim of this study is **to examine the usefulness of 5 - levels as a measure of population health and evaluate other variations in the number of response levels in terms of scale efficiency.**

Methods

- We analyzed data from the 2017 US EQ-5D-5L valuation study[2].
- In addition to valuation tasks, respondents self-reported their health using the 5L, 3L, and dimension specific VAS rating scales with 101 response levels (101L), where 0 represents extreme problems and 100 represents no problems).
- For each level of each EQ-5D dimension, mean VAS ratings were calculated to gain insight into how respondents self-calibrated their dimension-specific health on a more granular scale. Mean differences in VAS ratings were calculated between adjacent levels, including a level comprised of combined level 4 and 5 responses (4L).
- Descriptive richness was characterized by Shannon's index (H'; a measure of informativity) and Shannon's evenness index (J'; a measure of descriptive efficiency, controlling for number of levels)[3-5] across variations on the number of EQ-5D response levels, including the 101L, 5L, 4L, and 3L.

* Combined level 4 and 5 responses of the EQ-5D-5L

Results

Figure 1: Example Dimension-Specific VAS Rating Task

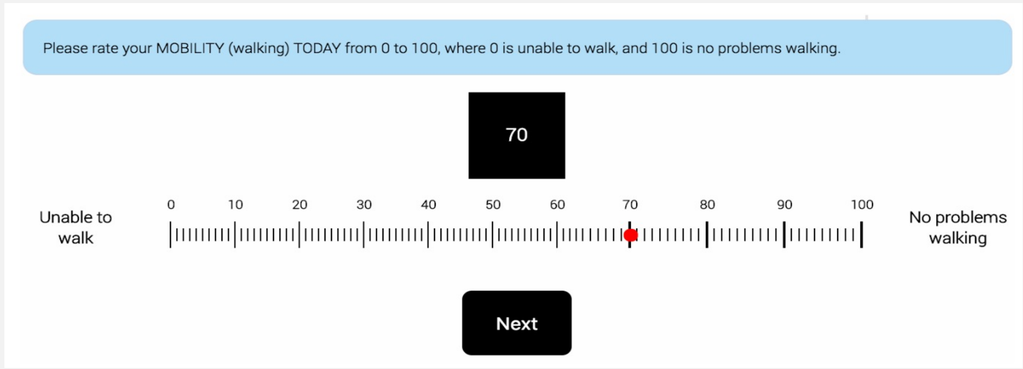
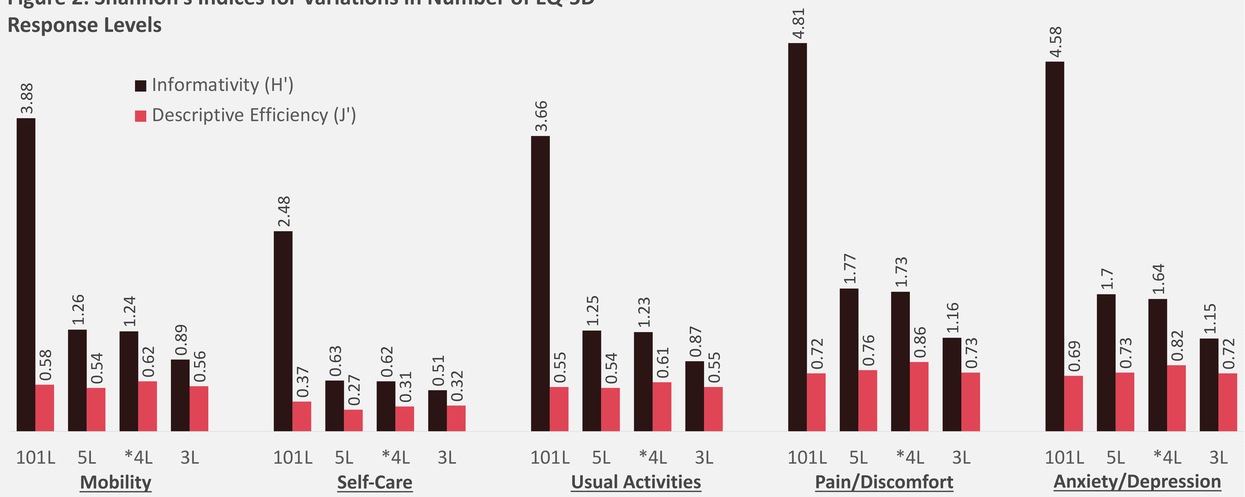


Figure 2: Shannon's Indices for Variations in Number of EQ-5D Response Levels



*The dataset included a total of 3151 participants, with data collected in-person (n=1133) or online (n=2018) using quota sampling based on age, sex, and race/ethnicity. There was poor differentiation between EQ-5D-5L level 4 and 5 responses when calibrated to the VAS. Only anxiety/depression demonstrated a substantial difference in mean VAS scores between levels 4 and 5 (**Table 1**). Averaged across all EQ-5D-5L dimensions, mean (SD) VAS ratings by levels 1/2/3/4/5 were 95.5 (11.1), 79.4 (19.0), 63.4 (23.4), 47.6 (27.6), 48.1 (33.6), respectively; a consistent mean difference of approximately 16 between levels 1/2, 2/3, 3/4, but minimal difference between 4/5. Across dimensions, descriptive efficiency (J') increased when combining EQ-5D-5L levels 4 and 5 into a 4-level (4L) system, with only marginal reductions in informativity (H') (**Figure 2**).*

Table 1: Dimension Specific VAS Ratings by EQ-5D Level

EQ-5D-5L Response Level	VAS Rating [Mean (SD)]	Number of Respondents
Mobility		
1	96.2 (9.0)	2236
2	79.6 (16.8)	569
3	63.7 (21.7)	251
4	43.8 (25.4)	73
5	44.1 (42.4)	22
4+5	43.9 (29.9)	95
Self-Care		
1	97.7 (7.6)	2815
2	75.8 (20.6)	215
3	69.6 (21.0)	89
4	58.4 (33.3)	20
5	68.0 (33.4)	12
4+5	62.0 (33.2)	32
Usual Activities		
1	96.1 (8.9)	2242
2	81.9 (17.7)	563
3	62.5 (24.6)	267
4	45.8 (30.4)	62
5	62.2 (27.6)	17
4+5	49.3 (30.4)	79
Pain/Discomfort		
1	94.3 (13.7)	1304
2	82.4 (18.5)	1134
3	61.5 (23.1)	534
4	43.1 (26.3)	143
5	39.9 (33.0)	36
4+5	42.5 (27.7)	179
Anxiety/Depression		
1	93.0 (17.4)	1685
2	77.5 (21.5)	796
3	59.8 (24.2)	480
4	47.0 (27.8)	124
5	26.4 (32.1)	66
4+5	39.9 (30.9)	190

Conclusions

- US general population **respondents poorly differentiate between levels 4 and 5 when anchoring on an independent VAS scale** for all dimensions except anxiety/depression.
- Compared to the 5L or 3L, **a system with 4 levels of response may be more optimal when considering the equipoise of descriptive richness and efficiency.** These findings may be related to the specific choice of labels for the 5L system (severe vs extreme).
- The generalizability **of these results could be considered in other settings and populations to inform future development of measures**.

No funding was received for this study. The EuroQol Research Foundation & Bristol Myers Squibb provided funding for the US Valuation study[2]. Dr. Nazari is supported by a UIC/Pfizer fellowship in HEOR. Dr. Kuharic is a past UIC/Takeda fellow in HEOR.

1. Herdman M, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736.
2. Pickard, A.S., et al., *United States Valuation of EQ-5D-5L Health States Using an International Protocol.* Value Health, 2019;22(8):931-941.
3. Bas Janssen MF, et al. Evaluating the discriminatory power of EQ-5D, HUI2 and HUI3 in a US general population survey using Shannon's indices. *Qual Life Res*. 2007;16(5):895-904.
4. Pickard AS, et al. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. *Med Care*. 2007;45(3):259-263.
5. Janssen MF, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717-1727.