This is a repository copy of *Psychometric performance of the EQ Health and Wellbeing short (EQ-HWB-S) in a UK population sample*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/212547/

Version: Published Version

ScienceDirect

ELSEVIER

Patient-Reported Outcomes

Psychometric Performance of the EQ Health and Wellbeing Short in a United Kingdom Population Sample

Emily McDool, PhD, Clara Mukuria, PhD, Tessa Peasgood, PhD

A B S T R A C T

Objectives: This study assessed the psychometric performance and construct validity of the EQ Health and Wellbeing Short (EQ-HWB-S), using a validated measure, the EQ-5D-5L, as a comparator.

Methods: The experimental version of the EQ-HWB-S was compared with the EQ-5D-5L to assess the psychometric performance of the measures. Data were drawn from the valuation stages of the Extending the Quality-Adjusted Life-Year project (UK general population, n = 429) and the EQ-5D-5L UK valuation pilot study (UK general population, n = 248). Construct validity was assessed based on convergent validity, using Spearman correlations and Pearson correlations. Known-group validity was assessed by estimating effect sizes to assess the ability of the EQ-HWB-S and EQ-5D-5L to discriminate between known groups based on "healthy" status, presence of a long-term condition, health and life satisfaction, age, and employment status. The degree of agreement in utility values across instruments was also evaluated using Bland-Altman plots.

Results: Strong associations ($r_s \geq 0.5$, $P < .001$) were found between conceptually overlapping dimensions and the utility scores of the EQ-HWB-S and EQ-5D-5L. The instruments performed comparably in discriminating between known groups including healthy versus unhealthy groups (based on the visual analog scale ≥ 80), long-term condition (vs no long-term condition), and above versus below average health and life satisfaction and employed (vs unemployed and long-term sick).

Conclusions: The EQ-HWB-S performs favorably with utility values successfully discriminating between groups in which differences are expected. Convergence between the EQ-HWB-S and EQ-5D-5L is evident, especially between conceptually overlapping dimensions.

Keywords: EQ Health and Wellbeing Short, preferences, quality-adjusted life-years, utilities.

VALUE HEALTH. 2024; ■(■):■–■

Highlights

- There is evidence of convergence between the EQ Health and Wellbeing Short (EQ-HWB-S) and EQ-5D-5L instruments, especially for constructs that are similar across both instruments, including mobility (getting around or walking), activities (daily or usual), anxiety and depression (separate or combined items), and pain (physical pain or pain and discomfort).

- The EQ-HWB-S successfully distinguished between groups with known differences in health based on self-reported health, experience of long-term conditions, and health satisfaction in addition to life satisfaction and employment status.

- The EQ-HWB-S extends beyond the EQ-5D-5L by capturing additional aspects of health-related quality of life, namely, loneliness, cognition, and control.

Introduction

The quality-adjusted life-year (QALY) is a widely adopted outcome used in the economic evaluation of health interventions. A QALY is a single index that combines the length of life and the value of health-related quality of life (HRQoL). HRQoL can be measured using preference-based measures that are scored using utility values.[1] Utility values are typically obtained by eliciting the preferences relating to health states from members of the general population.[2] The most widely used preference-based measure for measuring HRQoL in economic evaluation is the EQ-5D[3] with 3 (EQ-5D-3L)[3] or 5 (EQ-5D-5L)[4] levels that each describe health across 5 dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression.

The EQ Health and Wellbeing (EQ-HWB) is a new generic measure that covers health and wellbeing with 7 dimensions: activity, relationships, cognition, self-identify, autonomy, feelings, and physical sensations.[5,6] The measure was developed as part of the "Extending the QALY project," jointly funded by the UK

Medical Research Council and the EuroQol Research Foundation. It was designed to enable economic evaluation using QALYs when assessing interventions in health, public health, and social care, for care recipients and other beneficiaries, such as informal carers.[6] Two versions of the measure have been developed, the EQ-HWB (25-items) and the short version, the EQ-HWB-S (9-items). The EQ-HWB was developed to include items identified as being important to patients, carers, and social care users for their quality of life with input from various stakeholders, whereas the EQ-HWB-S includes a subset of core generic items that were ranked by stakeholders as being important for decision making and included items that would be amenable to preference elicitation. The 9 items included in the EQ-HWB-S are mobility, activities, exhaustion, loneliness, cognition, anxiety, sadness/depression, control, and physical pain. A recent feasibility valuation study has been undertaken in the United Kingdom to generate a scoring for the EQ-HWB-S based on general public preferences.[7]

Because the instrument was recently developed, limited evidence is available on its validity and psychometric performance in

the assessment of health and wellbeing. The EQ-5D-5L acts as a sound initial comparator in a first step toward understanding the descriptive and psychometric properties of the EQ-HWB-S relative to alternative measures because the EQ-5D-5L is a broadly utilized measure for assessing health. Additionally, the same international protocol was adopted in the valuation of the EQ-5D-5L and the UK EQ-HWB-S feasibility valuation study, enhancing the comparability.[8] The EQ-HWB-S and EQ-5D-5L measure some overlapping HRQoL concepts, including mobility, pain, activities, and anxiety/depression, each with 5 response levels, although they vary in the recall period because the EQ-HWB asks about the past 7 days, and the EQ-5D-5L asks about today. There are also differences in response options and item wording, 1 example of this is pain where the EQ-HWB-S asks respondents about physical pain, with responses ranging from no physical pain to very severe physical pain, whereas the EQ-5D-5L includes pain or discomfort, with responses ranging from no pain or discomfort to extreme pain or discomfort. The differences all across items that measure overlapping concepts are highlighted in the Appendix in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.05.007.

The available evidence on the psychometric performance of the preliminary version of the EQ-HWB-S relative to the EQ-5D-5L, based on data from the United States, indicates that the EQ-HWB-S performed similarly or better than the EQ-5D-5L among patient groups and is able to distinguish between groups, including caregivers and patient groups. The evidence also identified convergence between measures, especially dimensions of similar construct.[9] However, the preliminary version of the EQ-HWB-S was changed before valuation to reverse and negatively frame the control question, reword the sadness question to include "depressed," and combine mobility inside and outside and concentrating and thinking clearly questions. Further evidence is therefore required to support the use of the current version of the EQ-HWB-S.[7] Additionally, few studies have examined the properties of the EQ-HWB-S utility values, and it is important to build upon this existing knowledge base to understand the potential implications of its use in economic evaluation to inform Health Technology Assessment (HTA). This study therefore aimed to expand upon the current knowledge base surrounding the psychometric performance of the EQ-HWB-S while highlighting the similarities and differences between the EQ-HWB-S and EQ-5D-5L.

Methods

Data

Data were drawn from 2 existing data sets: (1) the valuation stage of the Extending the QALY (E-QALY) project and (2) data from the EQ-5D-5L valuation pilot study. Both data sets were collected as part of valuation studies conducted in the United Kingdom with members of the general population. The data were combined to produce a large (n = 677) data set that provided EQ-HWB-S and EQ-5D-5L responses, in addition to a wealth of demographic information from participants.

Data from the valuation stage of the E-QALY project

The EQ-HWB-S feasibility valuation study data were collected between May and November 2021 from 521 members of the UK general public recruited through advertising online and through targeted postal invites.[7] Individuals were sampled via quota sampling based on age, sex, and ethnicity. Before undertaking the online valuation interview, respondents were asked to complete the EQ-HWB-S followed by the EQ-5D-5L

and a demographic background survey, which also included questions around health, life satisfaction, and health satisfaction. EQ-HWB-S and EQ-5D-5L responses were collected from 521 and 429 individuals, respectively. Further details are available in the valuation study.[7] We utilized data on the 429 individuals who provided responses both to the EQ-5D-5L and EQ-HWB-S.

Data from the UK EQ-5D-5L pilot valuation study

As part of the UK EQ-5D-5L valuation study, interviewers initially completed a pilot phase of interviewer training with members of the public. Participants were members of the public who were recruited through advertising online to be interviewed online or in-person with no indication that they would be interviewed as part of a pilot. Participants were asked to complete the EQ-5D-5L and some brief background questions, before completing valuation exercises. Participants then completed further sociodemographic questions and the EQ-HWB-S. The pilot data will not be used in the main EQ-5D-5L valuation study but were collected in the same manner as the main study data. The data collected by the pilot study and utilized in this study were collected in interviews conducted between October 2022 and February 2023. The majority of interviews were conducted online (81%), with the remainder being conducted face-to-face in universities across the United Kingdom. The usable sample with both EQ-HWB-S and EQ-5D-5L included 248 individuals from across the United Kingdom.

These data were combined to generate a single data set on which the analyses were conducted because the same versions of the instruments were administered in a comparable format and setting, that is, as part of valuation interviews conducted with members of the UK population and predominantly administered online.

Measures and Additional Questions

Respondents completed the EQ-HWB-S (experimental version 1.0) and the EQ-5D-5L in both surveys. The EQ-HWB-S has 9 questions, and utility values were calculated using values from the value set from the UK feasibility study.[7] The utility values range from −0.384 to 1. The EQ-5D-5L has 5 questions and a visual analog scale ranging from 0 (worst imaginable health) to 100 (best imaginable health). As recommended by the National Institute of Health and Care Excellence (NICE),[10-12] the utility values of the EQ-5D-5L were calculated by mapping the EQ-5D-5L descriptive system onto the EQ-5D-3L UK value set. Predicted EQ-5D-5L utility values obtained by mapping do not cover the full range of scores; they can range from −0.255 to 0.960.[13]

Participants also completed additional questions surrounding their health and wellbeing, including their experience of serious illness (in themselves, family, and someone they cared for). Participants were also asked whether they had a long-term condition, but the question varied slightly across the 2 studies, with the EQ-5D-5L study asking, "Do any of your conditions or illnesses reduce your ability to carry out day-to-day activities?" and the EQ-HWB-S study asking, "Do you have any long-standing physical or mental impairment, illness, or disability which has been diagnosed by a doctor?" Life and health satisfaction scores, which were rated on a 0 to 10 scale, were collected in the EQ-HWB valuation study only. A score of 7 and below was coded to indicate below average satisfaction, in line with the average UK life satisfaction score of 7.54 and in line with mean values in the data (mean life satisfaction = 7.18; mean health satisfaction = 7.01).[14]

Carer status data were collected in the EQ-HWB valuation study only. Further sociodemographic information was collected

by both studies, including age, gender, ethnicity, employment status, education, marital status, and housing tenure.

Ethical approval to utilize the secondary data sources was obtained from the University of Sheffield School of Health and Related Research Ethics Committee (052211).

Statistical Analysis

Ceiling effects were examined for both the EQ-HWB-S and EQ-5D-5L and defined as the proportion reporting "no problems" (or equivalent) at the item level and the proportion reporting "no problems" (or equivalent) in all dimensions at the instrument level. In line with previous analyses of the EQ-HWB-S, distributions were deemed problematic if they were skewed such that >70% responded in the top category.[5] It may be expected that the EQ-HWB-S is less likely to suffer instrument ceiling effects because of the broader aspects of health and wellbeing that are not covered by the EQ-5D-5L. The proportion reporting the worst possible level of health was examined at the item level only and deemed problematic if the distribution was skewed such that <5% responded in the bottom category.[5] The distribution of EQ-HWB-S responses for respondents who reported no problems in the EQ-5D-5L (ie, health state 11111) were evaluated and presented in the Appendix in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.05.007. It is expected that similar responses across instruments would be given in overlapping domains. We may expect to see differences in domains in which the recall period (ie, today or 7 days) may affect responses, for example, anxiety or depression,[15,16] whereas physical functioning domains, that is, mobility, may be more stable and less affected by recall period.[15,17]

The degree of association and convergent validity between the EQ-HWB-S and the EQ-5D-5L dimensions and utility scores were assessed. Spearman rank order and Pearson correlations were estimated to assess the relationship between EQ-HWB-S and EQ-5D-5L dimensions and utility values, respectively. Correlations were assessed as strong ($\geq$0.5), moderate (<0.5 to $\geq$0.3), and weak (<0.3).[18] It was hypothesized that dimensions that measured similar aspects of health would have moderate to strong correlations, including mobility items across instruments, activities (EQ-HWB-S), and usual activities (EQ-5D-5L); anxiety (EQ-HWB-S) and anxiety and depression (EQ-5D-5L); sadness/depression (EQ-HWB-S) and anxiety and depression (EQ-5D-5L); and pain (EQ-HWB-S) and pain/discomfort (EQ-5D-5L). Differences in responses to related dimensions across instruments were descriptively evaluated. Correlations were also examined in those who would be classified as having lower health based on the visual analog scale (VAS) (<80), given the general population focus of valuation studies, and above or below the sample average health satisfaction score.

A Bland-Altman plot, which plots the mean EQ-HWB-S and EQ-5D-5L utility scores against the difference between EQ-HWB-S and EQ-5D-5L values, was generated to explore the agreement between values and to identify whether potential differences in the relationship exists across the distribution of utility values or in a particular part of the distribution.[19] Lin's concordance correlation coefficient (CCC) was also reported for the EQ-HWB-S and EQ-5D-5L scores to assess agreement between utility scores. The CCC value close to unity implies good concordance between measures. The CCC was assessed as poor (CCC < 0.40), fair (CCC = 0.40 to 0.59), good (CCC = 0.60 to 0.74), and excellent (CCC > 0.75).[20,21]

The ability of the EQ-HWB-S utility values to discriminate between groups with known differences, that is, known-group validity, was also assessed. For health and wellbeing, known groups were based on being healthy (VAS $\geq$ 80) or unhealthy (VAS

< 80), whether individuals reported a long-term health condition or not, life satisfaction above or below the UK population average and the sample average (mean life satisfaction = 7), and health satisfaction that was above or sample average (mean health satisfaction = 7).[14] It was hypothesized that the EQ-HWB-S scores would be lower among individuals with lower health (lower EQ-VAS scores and with long-term health conditions) and those with below average life and health satisfaction.

Known-group validity was also assessed based on whether or not individuals reported being a carer, which may affect health and wellbeing. The EQ-HWB-S questions relate to the aspects of quality of life that may be negatively affected because of caring activity; therefore, it was hypothesized that carers would have lower EQ-HWB-S utility scores. Finally, known-group validity was explored based on age and employment (vs unemployed or report "long-term sick").

Mean differences and effect sizes (Cohen's d) were estimated to compare groups. The effect size is estimated by dividing the difference in mean scores between 2 adjacent subgroups by the pooled SD of scores. Effect sizes of $\geq$0.2 to <0.5, $\geq$0.5 to <0.8, and $\geq$0.8 indicate small, medium, and large effect sizes, respectively.[18] The statistical significance of differences between groups were tested using a t-test to compare 2 groups or an overall F-test from an analysis of variance to compare across >2 groups. EQ-5D-5L mean differences and effect sizes were also estimated for comparison purposes.

Results

Sample Description

Table 1 provides the characteristics of the sample that included 677 individuals with 59% females and a mean age of 47 (SD 15.85) years. The majority of the sample were White (83%) and in employment (59%), with postcompulsory education (90%), including a degree (69%).

Responses to the EQ-HWB-S and EQ-5D-5L

At the instrument level, 9.45% reported "no problems" across all dimensions of the EQ-HWB-S (36% in the EQ-5D-5L), and no participants reported the worst health and wellbeing on the EQ-HWB-S (555555555) or EQ-5D-5L (55555) (Table 1). Potentially problematic distributions at the ceiling (ie, >70% reporting no problems) were identified in the EQ-HWB-S mobility dimension only (83%) (Table 2). For each of the EQ-HWB-S and EQ-5D-5L dimensions, few respondents (<5%) reported the lowest level (floor), and there were no responses at the worst level for EQ-HWB-S mobility or for EQ-5D-5L self-care. The exhaustion dimension received the highest proportion of responses at the worst level (4%). A lower proportion of participants reported "no problems" in the anxiety (43%), sadness/depression (55%) or in both (36%) dimensions of the EQ-HWB compared with in the overlapping EQ-5D-5L dimensions anxiety/depression (58%).

The EQ-VAS responses ranged from 0 to 100, with a mean score of 76 (SD 17.54). The EQ-HWB-S utilities ranged from −0.212 to 1 (SD 0.20), and the EQ-5D-5L utility scores ranged from −0.25 to 0.988 (SD 0.19). Because of the mapping algorithm used,[12] the predicted EQ-5D-5L utility score does not cover the full range of values (ie, with an upper limit at 1). The distribution of EQ-HWB-S and EQ-5D-5L utility scores is presented graphically in Figure 1 for the full sample and for those deemed "healthy" (VAS $\geq$ 80) and "unhealthy" (VAS < 80).

Table 1. Sample characteristics.

Characteristics	Mean/% (SD)	n
Age	47.02 (15.85)	677
Age group, years		
18-30	21.86	148
31-50	34.12	231
51-65	28.66	194
66+	15.36	104
Gender		
Male	41.06	278
Female	58.64	397
Other	0.30	2
Ethnicity		
White	82.87	561
Asian	7.98	54
Black	4.87	33
Mixed	1.92	13
Other	1.77	12
Employment/activity status		
In employment/self-employment	59.38	402
Retired	19.79	134
Student	5.76	39
Unemployed	4.28	29
Long-term sick	3.69	25
Carer or volunteer	3.10	21
Not seeking work	0.59	4
Other occupation	2.95	20
Post-16 education	89.51	606
Degree	69.28	469
Caring responsibilities (total n = 429)	14.45	62
Long-term health condition	36.69	248
Experience of serious illness		
Serious illness in self	32.79	222
Serious illness in family	52.29	354
Serious illness in others cared	23.78	161
Health and wellbeing		
Life satisfaction	7.19 (1.80)	429
Health satisfaction	7.01 (1.78)	429
Healthy (VAS≥80)	55.69	377
EQ-5D-5L 11111	36.04	244
EQ-HWB-S 111111111	9.45	64
EQ-5D-5L 55555	0.00	0
EQ-HWB-S 555555555	0.00	0
VAS	76.03 (17.54)	677
EQ-5D index score	0.81 (0.19)	677
EQ-5D Devlin scores	0.88 (0.17)	677
EQ-5D index scores (crosswalk 3L) BvH)	0.82 (0.20)	677
EQ-HWB index score	0.83 (0.20)	677

Note. (1) Mean or percentage reported and SD for continuous variables only. (2) Employment categories in valuation data include volunteer as "other," but in EQ-5D-5L pilot data, volunteer is grouped with carer. Volunteers coded as "other" in EQ-HWB valuation study data but carer or volunteer in EQ-5D-5L pilot data. Looking after home (EQ-HWB valuation study data) is also grouped with other. (3) Long-term health condition questions varied across the data sets, with the EQ-HWB-S valuation data questionnaire asking whether the individual has a long-term health condition lasting for a period of at least 12 months and the EQ-5D-5L pilot data asking if any conditions or illnesses reduce the ability of the individual to carry out day-to day activities. Responses are combined in the EQ-5D-5L pilot study so that "limited a little" and "limited a lot" are coded as having a condition. (4) Life and health satisfaction were collected in the EQ-HWB valuation study only.
EQ-HWB indicates EQ Health and Wellbeing; EQ-HWB-S, EQ Health and Wellbeing Short; VAS, visual analog scale.

Convergent Validity

All correlations between overlapping dimensions were strong (≥0.5) (Table 3). In the full sample, strong correlations were also observed between the EQ-HWB-S utility score and EQ-5D-5L dimensions (0.507-0.624), with the exception of self-care in which there was moderate correlation (−0.440). Moderate to weak correlations (0.151-0.495) were identified in items assessing concepts of wellbeing, including exhaustion, loneliness, and cognition, with all EQ-5D-5L items, with the exception of anxiety and depression, which were strongly correlated with the EQ-5D-5L combined item (0.685 and 0.660, respectively). Convergent validity was also apparent in the subsamples including healthy and unhealthy individuals and those with both above and below average health satisfaction scores; strong correlations (0.508-0.694) were identified in these groups in the overlapping dimensions and utility scores across instruments. In the "healthy" sample and those with above average health satisfaction, weak to moderate correlations (<0.5) were observed across the majority of nonoverlapping domains. In these samples, weak correlations (<0.3) were especially apparent in the EQ-HWB-S wellbeing–related dimensions, including loneliness, cognition, anxiety, depression, and control, and EQ-5D-5L items, with the exception of anxiety/depression.

The Bland-Altman plot (Fig. 2) indicated some evidence of agreement between the utility values of the instruments, particularly at the upper end of the scales around 0.7 to 1. The difference in the utility values ranged from −0.61 to 0.31, with a mean absolute difference of 0.17. There is evidence of a lack of agreement in outliers. Outliers with utility score differences >0.2 are found across the EQ-HWB-S scale. Further assessment of agreement performed by calculating the CCC indicated a coefficient equal to 0.848, which signals excellent correlation and concordance.

Known-Group Validity

The results of known-group validity analyses are presented in Table 4. Large effect sizes (≥0.8) were identified across all groups in which comparisons were based on health and wellbeing characteristics, including the healthy versus unhealthy groups (based on the VAS), long-term condition (vs no long-term condition), and when based on above versus below average health and life satisfaction. The EQ-HWB-S was able to successfully discriminate between each of these groups. When grouping individuals based on observed sociodemographic characteristics, a large effect size was identified when comparing the employed with unemployed (including those who are long-term sick). Small effect sizes were identified when distinguishing between carers and noncarers and across age groups. There were only minor differences between the effect sizes calculated for the EQ-HWB-S and EQ-5D-5L across all groups. The EQ-5D-5L results did not seem to be driven by the choice of value set utilized in the analyses because the same conclusions were reached when adopting an alternative value set[22] (see Appendix in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.05.007).

Discussion

This study aimed to assess the psychometric performance, including the construct validity of the EQ-HWB-S, when applying the UK value set. The results indicate evidence of convergent validity of the EQ-HWB-S and the EQ-5D-5L with significant strong correlations (rs ≥ 0.5) between overlapping dimensions (mobility, activities, depression, anxiety, and pain), and between the EQ-HWB-S scores and the majority of EQ-5D-5L dimensions. Weak correlations were generally observed around EQ-HWB-S wellbeing domains, for example, loneliness, and all EQ-5D-5L domains except anxiety and depression. These findings mirror those identified in the comparison of the preliminary version of the EQ-HWB-S and EQ-5D-5L.[9] The high degree of convergence in

Table 2. Dimension responses EQ-HWB-S and EQ-5D-5L (N = 677).

Measure	Dimensions	Ceiling				Floor
		1	2	3	4	5
EQ-HWB-S	Mobility*	83.01	9.60	4.87	2.51	0
	Activities*	68.24	18.02	8.71	4.28	0.74
	Exhaustion[†]	31.46	31.46	19.35	13.59	4.14
	Loneliness[†]	61.30	20.83	11.23	4.87	1.77
	Cognition[†]	42.54	26.88	20.38	7.83	2.36
	Anxiety[†]	42.54	26.14	19.79	8.27	3.25
	Sadness/depression[†]	55.10	24.67	12.11	5.61	2.51
	Control[†]	63.66	20.38	9.01	4.43	2.51
	Pain[‡]	46.97	35.45	13.15	3.69	**0.74**
EQ-5D-5L	Mobility	**78.58**	13.29	5.76	2.22	**0.15**
	Self-care	**90.99**	5.47	2.95	0.59	**0**
	Usual activity	69.28	20.24	7.24	2.36	**0.89**
	Pain/discomfort	54.80	31.91	9.75	3.10	**0.44**
	Anxiety/depression	58.20	26.44	11.67	2.66	**1.03**

Note. (1) Percentages presented. (2) Values above the ceiling threshold (70%) in column 1, or below the floor threshold (5%) in column 5 are bolded.
EQ-HWB-S indicates EQ Health and Wellbeing Short.
*Difficulty scale: 1 = none of the time, 2 = only occasionally, 3 = sometimes, 4 = often, 5 = most or all of the time.
[†]EQ-HWB-S response scales include frequency scale: 1 = no difficulty, 2 = slight difficulty, 3 = some difficulty, 4 = a lot of difficulty, 5 = unable.
[‡]EQ-HWB-S response scales include severity scale: 1 = no physical pain, 2 = mild physical pain, 3 = moderate physical pain, 4 = severe physical pain, 5 = very severe physical pain. EQ-5D-5L response scales: 1 = no problems, 2 = slight problems; 3 = moderate problems; 4 = severe problems; 5 = extreme problems (pain/discomfort and anxiety/depression) or unable to (mobility, self-care, and usual activities).

overlapping items supports the existing evidence of the construct validity of the EQ-HWB-S.[9] The EQ-HWB-S also performed favorably with respect to known-group validity and was able to successfully discriminate between known groups in which differences in utility values were expected a priori, for example, healthy versus unhealthy groups and long-term conditions versus

Figure 1. Distribution of EQ-HWB-S and EQ-5D-5L utility values full sample (n = 677).
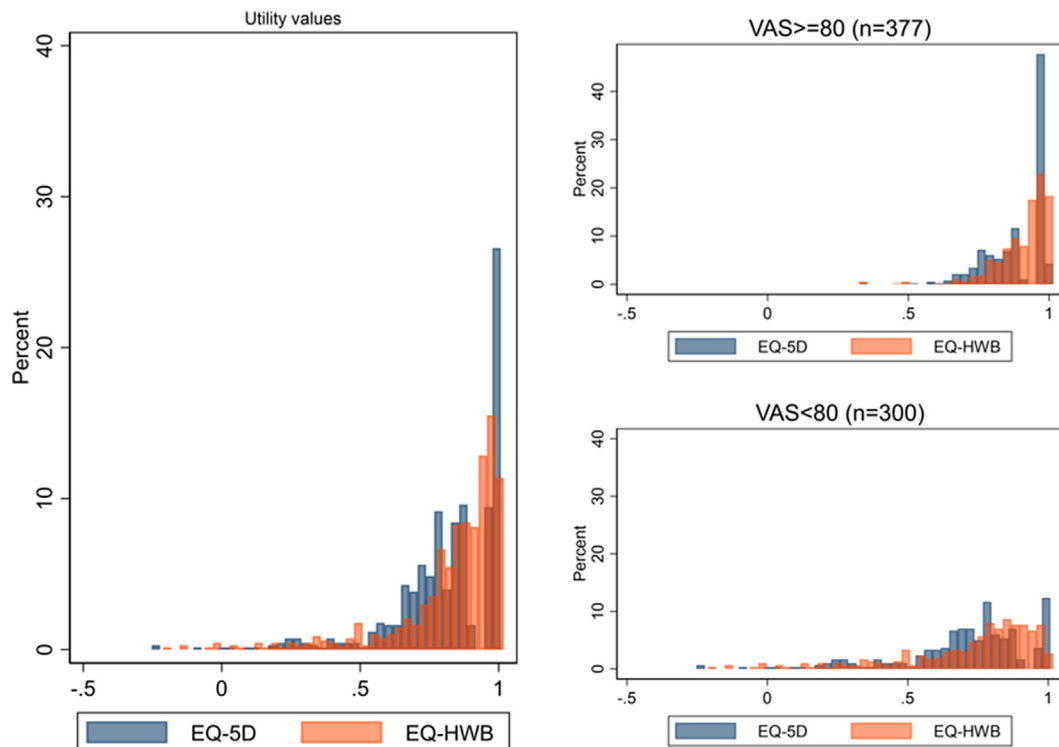
EQ-HWB-S indicates EQ Health and Wellbeing Short.

Table 3. Convergent validity of EQ-HWB-S and EQ-5D-5L.

Utilities/dimensions	EQ-HWB-S score	EQ-HWB-S Mobility	EQ-HWB-S Activities	EQ-HWB-S Exhaustion	EQ-HWB-S Lonely	EQ-HWB-S Cognition	EQ-HWB-S Anxious	EQ-HWB-S Sad/depression	EQ-HWB-S Control	EQ-HWB-S Pain
Full sample (n=677)										
EQ-5D-5L score	0.852*	−0.558*	−0.616*	−0.457†	−0.276‡	−0.424†	−0.468†	−0.461†	−0.407†	−0.682*
EQ-5D-5L Mobility	−0.507*	**0.764***	0.597*	0.276‡	0.165‡	0.234‡	0.172‡	0.220‡	0.234‡	0.544*
EQ-5D-5L Self-care	−0.440†	0.580*	0.517*	0.370†	0.262‡	0.295‡	0.256‡	0.288‡	0.336†	0.404†
EQ-5D-5L Usual activity	−0.624*	0.615*	**0.760***	0.437†	0.240‡	0.378†	0.267‡	0.334†	0.342†	0.538*
EQ-5D-5L Pain/discomfort	−0.586*	0.514*	0.524*	0.364†	0.151‡	0.299‡	0.266‡	0.255‡	0.277‡	**0.775***
EQ-5D-5L Anxious/depression	−0.605*	0.231†	0.334†	0.443‡	0.424†	0.495†	**0.685***	**0.660***	0.470	0.263‡
Unhealthy VAS < 80 (n = 300)										
EQ-5D-5L score	0.865*	−0.633*	−0.691*	−0.548*	−0.283‡	−0.447†	−0.475†	−0.489†	−0.451†	−0.704*
EQ-5D-5L Mobility	−0.559*	**0.773***	0.636*	0.311†	0.140‡	0.193‡	0.136‡	0.191‡	0.233‡	0.583*
EQ-5D-5L Self-care	−0.536*	0.616*	0.567*	0.425†	0.258‡	0.305†	0.256‡	0.272‡	0.377†	0.453†
EQ-5D-5L Usual activity	−0.662*	0.596*	**0.776***	0.490†	0.210‡	0.352†	0.265‡	0.306†	0.363†	0.533*
EQ-5D-5L Pain/discomfort	−0.601*	0.528*	0.523*	0.407†	0.127‡	0.269‡	0.241‡	0.246‡	0.262‡	**0.791***
EQ-5D-5L Anxious/depression	−0.587*	0.189‡	0.302†	0.417†	0.454†	0.514*	**0.689***	**0.699***	0.505*	0.197‡
Healthy (VAS ≥ 80) (n = 377)										
EQ-5D-5L score	0.593*	−0.319†	−0.379†	−0.207‡	−0.090‡	−0.238‡	−0.311†	−0.225‡	−0.222‡	−0.556*
EQ-5D-5L Mobility	−0.285‡	**0.631***	0.366†	0.051‡	0.014‡	0.052‡	−0.004‡	−0.020‡	0.001‡	0.386†
EQ-5D-5L Self-care	−0.200‡	0.266‡	0.292‡	0.178‡	0.136‡	0.098‡	0.098‡	0.128‡	0.069‡	0.210‡
EQ-5D-5L Usual activity	−0.402†	0.482†	**0.599***	0.188‡	0.072‡	0.211‡	0.049‡	0.087‡	0.129‡	0.398†
EQ-5D-5L Pain/discomfort	−0.431†	0.359†	0.372†	0.174‡	−0.002‡	0.165‡	0.116‡	0.032‡	0.125‡	**0.686***
EQ-5D-5L Anxious/depression	−0.465†	−0.002‡	0.145‡	0.305†	0.250‡	0.338†	**0.616***	**0.508***	0.311†	0.097‡
Above average health satisfaction (N = 201)										
EQ-5D-5L score	0.599*	−0.367†	−0.316†	−0.158‡	−0.029‡	−0.171‡	−0.331†	−0.272‡	−0.312†	−0.481‡
EQ-5D-5L Mobility	−0.276‡	**0.694***	0.373†	0.032‡	−0.036‡	0.150†	0.069‡	0.063‡	0.149‡	0.262‡
EQ-5D-5L Self-care	−0.109‡	0.305†	0.227‡	0.040‡	−0.047‡	−0.054‡	−0.060‡	−0.049‡	−0.039‡	0.143‡
EQ-5D-5L Usual activity	−0.324†	0.830*	**0.534***	0.087‡	0.028‡	0.171‡	0.089‡	0.098‡	0.141‡	0.282‡
EQ-5D-5L Pain/discomfort	−0.416†	0.352†	0.299‡	0.189‡	0.023‡	0.130‡	0.201‡	0.138‡	0.256‡	**0.604***

Table 3. Continued

Utilities/ dimensions	EQ-HWB-S score	EQ-HWB-S Mobility	EQ-HWB-S Activities	EQ-HWB-S Exhaustion	EQ-HWB-S Lonely	EQ-HWB-S Cognition	EQ-HWB-S Anxious	EQ-HWB-S Sad/ depression	EQ-HWB-S Control	EQ-HWB-S Pain
EQ-5D-5L Anxious/ depression	−0.542*	0.117‡	0.198‡	0.302†	0.300†	0.354†	**0.649***	**0.597***	0.439†	0.210‡
Below average health satisfaction (n = 228)										
EQ-5D-5L score	0.857*	−0.636*	−0.629*	−0.460†	−0.285‡	−0.415†	−0.463†	−0.462†	−0.415†	−0.685*
EQ-5D-5L Mobility	−0.527*	**0.787***	0.588*	0.259‡	0.170‡	0.196‡	0.137‡	0.138‡	0.189‡	0.587*
EQ-5D-5L Self-care	−0.458†	0.578*	0.475†	0.314†	0.284‡	0.307†	0.281‡	0.298‡	0.348†	0.401‡
EQ-5D-5L Usual activity	−0.631*	0.625*	**0.769***	0.441†	0.246‡	0.327†	0.197‡	0.295‡	0.296‡	0.503*
EQ-5D-5L Pain/ discomfort	−0.583*	0.564*	0.481†	0.346†	0.116‡	0.254‡	0.256‡	0.198‡	0.223‡	**0.816***
EQ-5D-5L Anxious/ depression	−0.594*	0.233‡	0.336†	0.470†	0.450†	0.546*	**0.697***	**0.736***	0.541*	0.158‡

Note. (1) Pearson correlation presented for continuous data (EQ-HWB-S and EQ-5D-5L utility scores) and Spearman correlations presented for correlations including dimensions. (2) All correlations significant at 1% level. (3) The correlations of overlapping dimensions are bolded.
EQ-HWB-S indicates EQ Health and Wellbeing Short; VAS, visual analog scale.
*Correlations ≥0.5 are assessed as strong.
†Correlations <0.3 are assessed as weak.
‡Correlations <0.5 to ≥0.3 are assessed as moderate.

no long-term condition, in which large effect sizes (≥0.8) were estimated.

The evidence is encouraging, suggesting that the EQ-HWB-S successfully captures dimensions that overlap with those measured by the EQ-5D-5L. It also illustrates that the EQ-HWB-S extends beyond the EQ-5D-5L by capturing additional aspects of HRQoL specifically relating to wellbeing. Individuals deemed as being in "full health" (EQ-5D-5L) report problems with other domains, and weak correlations (<0.3) are identified between the EQ-5D-5L score and EQ-HWB-S wellbeing domains. In this study, a general population sample was utilized, rather than a patient group; thus, we may expect fewer reported problems on both instruments. The value of using the EQ-HWB-S is likely to be gained in conditions or populations in which problems

Figure 2. Bland-Altman plot of differences in EQ-HWB-S and EQ-5D-5L utility scores.
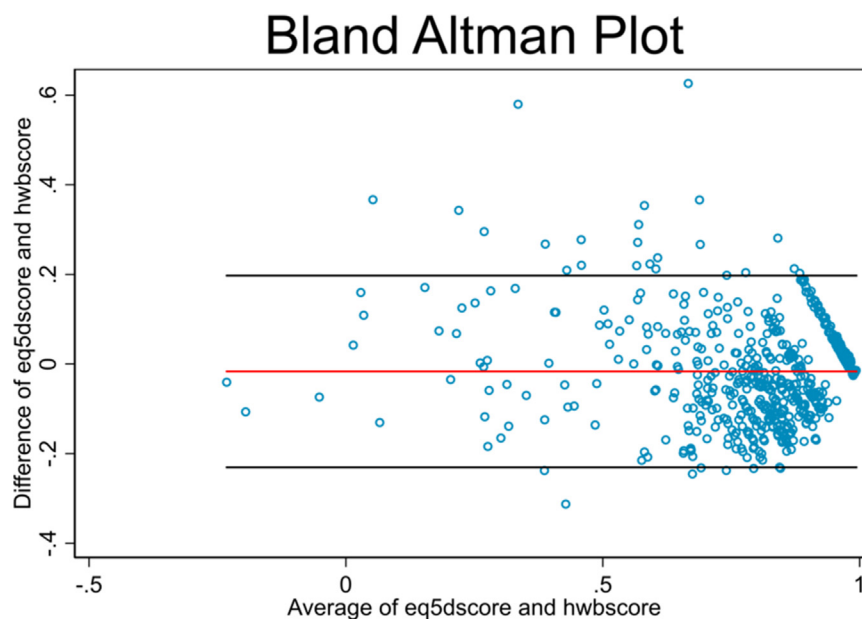
EQ-HWB-S indicates EQ Health and Wellbeing Short.

Table 4. Known-group validity assessment of the EQ-HWB-S and EQ-5D-5L.

Groups	n	EQ-HWB-S					EQ-5D-5L				
		utility value	SD	Differences in means	Test statistics (*P* value)	Effect size, Cohen's *d*	Mean utility value	SD	Differences in means	Test statistics (*P* value)	Effect size, Cohen's *d*
Health and wellbeing											
Unhealthy (VAS < 80)	300	0.725	0.246				0.712	0.224			
Healthy (VAS ≥ 80)	377	0.912	0.096	−0.187	−13.53 (.000)	−1.047	0.893	0.106	−0.181	−13.816 (.000)	−1.069
No long-term condition	428	0.905	0.104				0.892	0.104			
Long-term condition	248	0.700	0.255	0.205	14.648 (.000)	1.169	0.677	0.228	0.214	16.669 (.000)	1.330
Below average life satisfaction*	208	0.761	0.215				0.770	0.202			
Above average life satisfaction	221	0.924	0.079	−0.163	−10.527 (.000)	−1.017	0.895	0.116	−0.125	-7.903 (.000)	−0.763
Below average health satisfaction*	228	0.777	0.211				0.765	0.198			
Above average health satisfaction	201	0.922	0.086	−0.145	−9.123 (.000)	−0.883	0.914	0.0953	−0.149	-9.732 (.000)	−0.942
Sociodemographic characteristics											
Carer*	62	0.779	0.209				0.771	0.193			
Noncarer	367	0.856	0.172	0.076	3.120 (.002)	0.428	0.845	0.170	0.074	3.102 (.002)	0.426
Age 18-30	148	0.793	0.210				0.820	0.178			
Age 31-50	231	0.807	0.216	−0.014		−0.069	0.803	0.205	0.016		0.086
Age 51-65	194	0.864	0.181	−0.057		−0.289	0.822	0.187	−0.019		−0.101
Age 66+	104	0.866	0.175	−0.002	5.669 (.000)	−0.012	0.805	0.185	0.017	0.469 (.704)	0.091
Employed	402	0.862	0.158				0.852	0.144			
Unemployed or long-term sick	54	0.571	0.300	0.291	11.123 (.000)	1.612	0.561	0.294	0.290	11.881 (.000)	1.722

Note. (1) The test statistic provided is a *t* value for the comparison of 2 groups and the F-statistic from analysis of variance when comparing across groups (by age and gender). (2) Effect sizes of ≥0.2 to <0.5, ≥0.5 to <0.8, and ≥0.8 denote small, medium, and large effect sizes, respectively.
EQ-HWB-S indicates EQ Health and Wellbeing Short; KGV, Known Group Validity; VAS, visual analog scale.
*Carer and life/health satisfaction KGV is only estimated on the EQ-HWB-S valuation data because comparable carer or life/health satisfaction data are not collected in the EQ-5D-5L pilot study.

relating to exhaustion, loneliness, cognition, or control are common but are not reflected adequately by the EQ-5D-5L. Additionally, the measure was developed for use in social care users and carers to overcome the current challenges relating to the use of alternative measures and limited opportunities for comparison across sectors.[6] Although we cannot say that the EQ-HWB-S will be adopted by NICE, in these populations, there is a potential for the EQ-HWB-S to be a viable alternative for cases in which the EQ-5D-5L may not be appropriate, for example, if performing poorly in terms of construct validity and responsiveness. In such a case, NICE guidance stipulates that an alternative HRQoL measure may be used.[23] However, the implications of using the EQ-HWB-S on QALY estimation in HTA are not straightforward. The shift from one instrument to another may not affect equally across the distribution of health and wellbeing and may therefore depend on the patient or population group. This was clear in the move from the EQ-5D-3L to the EQ-5D-5L.[24] Furthermore, the analysis was undertaken using the experimental version of a measure and a definitive EQ-HWB-S value set is yet to be produced. Finally, the current UK EQ-5D-5L value set is likely to be replaced imminently; thus, HTA implications may need to be drawn from comparisons with the updated value set if adopted.

This study utilized 2 data sets collected as part of the EQ-HWB-S development (E-QALY project) and the EQ-5D-5L valuation study. Given the similarities in the questions and data collection of the valuation studies, these data were combined to provide a larger data set on which the value set was applied, and the main

psychometric analysis was conducted. It should be acknowledged that the EQ-HWB-S was collected during the COVID-19 pandemic, which may have affected subjective HRQoL and mental well-being.[25] Additionally, the order in which the instruments were presented varied across studies. Thus, some differences may be attributable to either or both ordering effects of instruments and the mode of administration.[26,27] However, the same version of each self-completion questionnaire was administered to respondents. The combination of data aids larger sample sizes, allowing for more accurate estimation involved in the assessment on construct validity.

The EQ-HWB-S value set applied was obtained directly from the preferences of members of the public (using time-trade-off and discrete choice experiments) and data modeling using a hybrid Tobit heteroscedastic model. The EQ-5D-5L value set observed was the result of mapping the EQ-5D-5L to the EQ-5D-3L value set as currently recommended by NICE with a mapping algorithm that is age and gender specific and features loss of dispersion, which leads to an upper limit of the value set below 1.[13] However, the results of the known-group validity sensitivity analysis conducted using different value sets indicate that the adoption of an alternative EQ-5D-5L UK value set would provide the same results and conclusions.

Upon the publication of the new UK value set, future work may wish to revisit this analysis and reconsider the assessment of known-group validity and the agreement of the utility values. Additionally, to further expand the evidence on the psychometric performance of the EQ-HWB-S, future work may wish to analyze the construct validity of the EQ-HWB-S in specific disease and condition groups when applying the UK value set.

Conclusions

The EQ-HWB-S performs favorably with regard to known-group validity, with the utility values successfully discriminating between healthy individuals and those with long-term conditions. There is evidence of agreement between the utility scores from the EQ-HWB-S and EQ-5D-5L instruments and of convergent validity, especially in the overlapping dimensions.

Author Disclosures

Author disclosure forms can be accessed below in the Supplemental Material section.

The views expressed in this article are those of the authors only.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2024.05.007.

Article and Author Information

Author Affiliations: Sheffield Centre for Health and Related Research (SCHARR), University of Sheffield, Sheffield, England, UK (McDool, Mukuria, Peasgood).

Correspondence: Clara Mukuria, PhD, Sheffield Centre for Health and Related Research, Regent Court, Regent Street, Sheffield, England S1 4DA, United Kingdom. Email: c.mukuria@sheffield.ac.uk

REFERENCES

1. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford, United Kingdom: Oxford university press; 2017.
2. Rowen D. Preference-based measures of health-related quality of life. In: Michalos AC, ed. *Encyclopedia of Quality of Life and Well-Being Research*. Dordrecht, The Netherlands: Springer Netherlands; 2014:5026–5028.
3. Brooks R, Group E. EuroQol: the current state of play. *Health Policy*. 1996;37(1):53–72.
4. Herdman M, Gudex C, Lloyd A, 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.
5. Peasgood T, Mukuria C, Brazier J, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ-HWB. *Value Health*. 2022;25(4):525–533.
6. Brazier J, Peasgood T, Mukuria C, et al. The EQ-HWB: overview of the development of a measure of health and wellbeing and key results. *Value Health*. 2022;25(4):482–491.
7. Mukuria C, Peasgood T, McDool E, Norman R, Rowen D, Brazier J. Valuing the EQ Health and Wellbeing Short using time trade-off and a discrete choice experiment: a feasibility study. *Value Health*. 2023;26(7):1073–1084.
8. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goñi JM. Overview, update, and lessons learned from the international EQ-5D-5L valuation work: version 2 of the EQ-5D-5L valuation protocol. *Value Health*. 2019;22(1):23–30.
9. Monteiro AL, Kuharic M, Pickard AS. A comparison of a preliminary version of the EQ-HWB short and the 5-level version EQ-5D. *Value Health*. 2022;25(4):534–543.
10. Hernández Alava M, Pudney S, Wailoo A. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from a UK population study. *Pharmacoeconomics*. 2023;41(2):199–207.
11. NICE health technology evaluations: the manual. National Institute of Health and Care Excellence. https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation. Accessed January 2023.
12. Hernández-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: the switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. *J Health Econ*. 2017;55:139–152.
13. Hernández-Alava M, Pudney S. eq5dmap: a command for mapping between EQ-5D-3L and EQ-5D-5L. *Stata J*. 2018;18(2):395–415.
14. Personal well-being in the UK: April 2021 to March 2022. *Stat Bull*. 2022.
15. Peasgood T, Caruana JM, Mukuria C. Systematic review of the effect of a one-day versus seven-day recall duration on patient reported outcome measures (PROMs). *Patient*. 2023;16(3):201–221.
16. Mendoza TR, Dueck AC, Bennett AV, et al. Evaluation of different recall periods for the US National Cancer Institute's PRO-CTCAE. *Clin Trials*. 2017;14(3):255–263.
17. Armstrong TS, Vera-Bolanos E, Acquaye A, Gilbert MR, Mendoza TR. Impact of recall period on primary brain tumor patient's self-report of symptoms. *Neurooncol Pract*. 2014;1(2):55–63.
18. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155–159.
19. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8(2):135–160.

20. Bulsei J, Pfister J, Leroy S, et al. Do we need to assess quality-of-life with both the St George's respiratory questionnaire and the EuroQol 5-dimension questionnaire in a clinical study with an economic component: insights from the REVOLENS study in severe emphysema. *Int J Chronic Obstruct Pulm Dis.* 2020;15:135–142.

21. Kocks JW, Kerstjens HA, Snijders SL, et al. Health status in routine clinical practice: validity of the clinical COPD questionnaire at the individual patient level. *Health Qual Life Outcomes.* 2010;8:135.

22. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ.* 2018;27(1):7–22.

23. NICE Health Technology Evaluations: The Manual 2022. National Institute for Health and Care Excellence (NICE). https://www.nice.org.uk/process/pmg36. Accessed June 10, 2024.

24. Hernandez Alava M, Wailoo A, Grimm S, et al. EQ-5D-5L versus EQ-5D-3L: the impact on cost effectiveness in the United Kingdom. *Value Health.* 2018;21(1):49–56.

25. Serrano-Alarcón M, Kentikelenis A, McKee M, Stuckler D. Impact of COVID-19 lockdowns on mental health: evidence from a quasi-natural experiment in England and Scotland. *Health Econ.* 2022;31(2):284–296.

26. Childs AL, Submacular Surgery Trials Research Group. Effect of order of administration of health-related quality of life interview instruments on responses. *Qual Life Res.* 2005;14(2):493–500.

27. Hays RD, Kim S, Spritzer KL, et al. Effects of mode and order of administration on generic health-related quality of life scores. *Value Health J Int Soc Pharmacoecon Outcomes Res.* 2009;12(6):1035–1039.