eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Patient-Reported Outcomes

The Measurement Properties of the EQ-HWB and the EQ-HWB-S in Italian Population: A Comparative Study With EQ-5D-5L

Maja Kuharić, MPharm, MSc, PhD, A. Simon Pickard, PhD, Clara Mukuria, PhD, Aureliano Paolo Finch, PhD

A B S T R A C T

Objectives: The EQ Health and Well-being, EQ-HWB (25-item) and the EQ-HWB-S (9-item), are new generic measures of health and well-being. The purpose of this study was to examine the measurement properties of the EQ-HWB and EQ-HWB-S measures in relation to the EQ-5D-5L among the Italian general population.

Methods: A cross-sectional survey was conducted from October 2020 to February 2021, followed by secondary analysis of collected data from Italian adults. This analysis included response pattern distributions, correlation strength, and known-group comparison. Known-group comparison was assessed using effect sizes (ES) across health conditions, caregiver status, and social care usage. The EQ-HWB-S index-based score was based on the UK pilot value set, whereas the Italian value set was used for scoring the EQ-5D-5L index.

Results: Out of the 1182 participants, 461 reported having a chronic condition, 185 identified as caregivers, and 42 were social care users. EQ-HWB items (7.5%) showed fewer ceiling effects than EQ-5D-5L items (34.7%). Strong correlations ($r_s > 0.5$) were found between overlapping EQ-HWB and EQ-5D-5L items. EQ-HWB-S and EQ-5D-5L index scores demonstrated similar discrimination based on symptomatic chronic conditions (ES d = 0.68 vs d = 0.71), but EQ-HWB-S had slightly higher ES for social care users (ES d = 0.84 vs d = 0.74).

Conclusions: Initial evidence supports the validity of EQ-HWB/EQ-HWB-S because outcome measures in the Italian population. EQ-HWB-S performed comparably to EQ-5D-5L among patients and was better in differentiating social care users. A slight decrease in discriminative properties for caregivers was observed when transitioning from EQ-HWB to EQ-HWB-S.

Keywords: EQ Health and well-being, EQ-HWB, EQ-5D-5L, Italy, psychometrics.

Highlights

- The EQ Health and Well-being, EQ-HWB (25-item) and EQ-HWB-S (9-item) were developed to meet the need for a single instrument applicable across different beneficiaries of health and social care services. The EQ-HWB is intended for use in clinical practice and population health surveys, whereas the EQ-HWB-S is designed for use in economic evaluations across health, social care, and public health.

- This study aimed to assess the psychometric performance of the EQ-HWB/EQ-HWB-S in the general Italian population, including patients, caregivers, and social care users. It evaluated ceiling/floor effects, convergent validity, and discriminative ability against EQ-5D-5L.

- Results support the validity of EQ-HWB/EQ-HWB-S as outcome measures in the Italian population. EQ-HWB-S performed comparably to EQ-5D-5L among patients and was better in differentiating social care users. There was a slight decrease in discriminative properties for caregivers was observed when transitioning from EQ-HWB to EQ-HWB-S. This study contributes to the growing body of literature on the EQ-HWB/EQ-HWB-S, highlighting its potential as a prominent measure in health and social care.

Introduction

Measures used in economic evaluations vary significantly in their descriptive systems' content and size.[1] EQ-5D, the most widely used generic preference-based measure,[2] is applicable to many medical conditions.[3,4] However, it may lack validity for certain conditions, such as hearing-related conditions,[5] visual impairments,[6] severe mental health disorders,[7-9] dementia,[10] and multiple sclerosis.[11] There is growing recognition that medical conditions affect not only health outcomes but also broader aspects of quality of life (QoL), which contribute to a sense of well-being.[12] These broader aspects may not be effectively captured by traditional health-focused measures, which becomes especially relevant for informal caregivers or users of social care.[13] Furthermore, in elderly care, decisions around the selection of healthcare services may prioritize well-being improvement alongside health gains.[14] This has led to several proposed strategies for decision making when evaluating new interventions. One approach is to use sector-specific measures, which can be problematic when the intervention's impact spans different sectors.[13] Another strategy is to extend a reference health-related quality of life (HRQoL) measure's descriptive system with "bolt-ons" (ie, single-item dimensions) relevant for a specific disease or capturing broader QoL elements.[15] However, bolt-ons have their limitations, including the need for consistency with the parent measure's wording and additional costs for bolt-on valuation.[16,17] An alternative approach is to develop a new generic measure relevant to health and social care.[13] This led to the development of the EQ-HWB.

The EQ Health and Well-being (EQ-HWB) is a new measure for health and well-being.[18-20] It consists of a 25-item profile measure (long version, EQ-HWB) and a 9-item classifier (short version, EQ-HWB-S). The EQ-HWB is intended for use in clinical practice and population health surveys, whereas the EQ-HWB-S is designed for use in economic evaluations across health, social care, and public health. Both measures were developed to meet the need for a single instrument applicable across different beneficiaries of health and social care services. This development was based on the views of health and social care users, as well as informal caregivers. The EQ-HWB was developed in an international study conducted across 6 countries (Argentina, Australia, China, Germany, the United Kingdom, and the United States) using 4 language versions: Argentinean Spanish, English, German, and Simplified Chinese.[18-20]

Recent research has shown that the EQ-HWB has face and content validity in the Italian population.[21] However, evidence regarding its feasibility, ceiling effects, convergent validity, and known group validity in this demographic is lacking. A study by Monteiro et al.[22] demonstrated significant convergence between the EQ-HWB-S and EQ-5D-5L for overlapping dimensions in a US-based study. This study also reported limited ceiling effects for the EQ-HWB-S and demonstrated the measure's ability to distinguish between caregivers and non-caregivers and those with low and high caregiver burden. After the US-based study that used the experimental version of the EQ-HWB, the instrument has undergone revisions, which included merging items, necessitating further evaluation of EQ-HWB refinements. This study aimed to assess the psychometric performance of the EQ-HWB/EQ-HWB-S in the general Italian population, including social care users, caregivers, and patients with health conditions by examining response distributions, ceiling and floor effects, convergent validity, and the ability to distinguish between chronic health conditions, social care users, and informal caregivers and compare it with the widely used generic measure of health, the EQ-5D-5L.

Methods

Data Collection

This study, conducted from October 2020 to February 2021, was a part of the Italian EQ-5D-5L valuation study approved by Bocconi University's Ethics Committee (2020-SA000136.4).[23] Utilizing Zoom and Lifesize for online interviews because of COVID-19, a team of 11 Bocconi University researchers and graduate students collected data. Managed by Pepe Research srl, the study sought a demographically representative sample of Italy's adult population using quota sampling by age, gender, and geographic location. Participants were categorized into 3 groups—people with chronic health conditions, social care users, and caregivers—by their self-reported responses during the interviews. They were asked to identify their chronic conditions based on classification from the International Classification of Diseases 11th revision. Additionally, participants indicated their roles as informal caregivers or social care user, the latter defined as recipients of social assistance in the last 2 years. The interview included first completing the EQ-5D-5L and EQ Visual Analog Scale (VAS), then a valuation exercise, followed by socio-economic, and health-related questions. Finally, they self-completed the EQ-HWB. For this study, the EQ-HWB was adapted into Italian from the English version, which included translation and content validation.[21] For a more detailed account of the overall survey methodology, readers are referred to Finch et al.[23]

Measures

While introducing the instruments in our study, it is important to highlight the conceptual differences between the EQ-HWB and the EQ-5D-5L. The EQ-HWB is intended to cover a wider range of health and well-being aspects, in contrast to the EQ-5D-5L, which focuses on specific HRQoL dimensions.

EQ Health and Well-being (EQ-HWB)/EQ Health and Well-being Short Form (EQ-HWB-S)

The EQ Health and Well-being (EQ-HWB) and its shorter version, EQ-HWB-S, are standardized measures designed to assess health and well-being.[18-20] The EQ-HWB is a 25-item profile measure covering seven domains: activity, relationships, cognition, self-identity, autonomy, feelings, and physical sensations. Each item is described in terms of 5 levels, which use frequency, severity, or difficulty. The EQ-HWB-S, a shorter 9-item version, serves primarily as a classifier for valuation purposes. Both measures aim to provide well-being assessment, for patients and caregivers, complementing rather than replacing existing tools such as the EQ-5D. They use a 7-day recall period, which allows respondents to reflect on their health and well-being over the past week. Currently, the EQ-HWB and EQ-HWB-S are Experimental Versions, available exclusively to research collaborators, with ongoing work to further refine and validate them for wider use.[24]

The EQ-HWB-S index values originated from a UK pilot study using the EuroQol Valuation Technology protocol, adapted specifically for this measure.[25] In the pilot value set of the EQ-HWB-S, scores range from −0.384 to 1, with 1 representing "full health and well-being"—a broader concept than "full health" as assessed by EQ-5D-5L. Although there is no established method for nonpreference-based scoring of the EQ-HWB yet, ongoing methodological studies are exploring this aspect.[26] Based on this study, the EQ-HWB-S was calculated as a single-level summary score (LSS) and EQ-HWB into 3 subscales: (1) activities LSS (3 items: Day-to-day activities, Getting around inside and outside, and Personal care) with a score range of 3 to 15, (2) Pain/Discomfort LSS (4 items: Pain [frequency and severity] and Discomfort [frequency and severity]) with a score range of 4 to 20, and (3) psychosocial well-being LSS (16 items: Sleep, Exhausted, Lonely, Unsupported, Remembering, Concentrating/thinking clearly, Anxious, Unsafe, Frustrated, Depressed, Look Forward, Control, Cope, Accepted, Feelgood, and Do things wanted to do) with a score range of 16 to 80. Higher scores on the EQ-HWB LSS and its subscales LSS indicate worse health and well-being.

EQ-5D

The EQ-5D-5L is a generic preference-based measure designed to measure HRQoL. It includes a 5-dimension descriptive system: Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression.[27] Each dimension offers 5 response levels, from "no problems" to "unable to/extreme" problems, yielding 3125 possible health states. Additionally, it includes a VAS (EQ VAS), for respondents to rate their health from zero (worst imaginable) to 100 (best imaginable).[28] The EQ-5D-5L, focusing on current health status, uses a "today" recall period. The EQ-5D-5L index score assigns a value to each health state, with a value of 1 signifying "full health" and 0 indicating a state of being "dead," based on the country-specific general population preferences.[29] For this study, the EQ-5D-5L index values were calculated using a scoring function from Italian population-based preference weights.[23]

Statistical Analysis

This study involved a secondary analysis of collected data, focusing on the measures index scores and item level, including

Table 1. Respondent characteristics.

Demographic Characteristics (N = 1182)	N (%)
Gender	
Male	606 (51.3)
Female	575 (48.7)
Other	1 (0.1)
Age	
Overall (Mean, SD)	48.29 (16.1)
18-44	475 (40.2)
45-65	462 (39.1)
65≥	245 (20.7)
Education	
Primary school diploma	77 (6.5)
High school diploma (technical, professional)	637 (53.9)
Three-year degree	125 (10.6)
Specialized or single-cycle degree	269 (22.8)
Master's degree	58 (4.9)
PhD	16 (1.4)
Employment	
Employed/self-employed	637 (41.2)
Student	112 (12.7)
Retired	234 (9.5)
Unemployed	92 (7.8)
Household	96 (8.1)
Other	11 (0.9)
Annual gross household salary (considering all income)	
Less than €14 000	93 (7.9)
€14 000-€34 999	463 (39.2)
€35 000-€55 999	313 (26.5)
€56 000-€90 999	133 (11.3)
€91 000 or more	13 (1.1)
Rather not respond	167 (14.1)
Marital status	
Unmarried	350 (29.6)
Married or living with a partner	727 (61.5)
Divorced or separated	78 (6.6)
Widowed	27 (2.3)

Clinical characteristics	N (%)
Do you suffer from one or more diseases?	
Yes	461 (39.0)
One disease	259 (21.9)
Two diseases	126 (10.7)
Three or more	75 (6.4)
Diseases	
Arthritis	68 (5.8)
Asthma/Chronic Obstructive Pulmonary Disorder	58 (4.9)
Cancer	32 (2.7)
Depression	54 (4.6)
Diabetes	62 (5.3)
Hepatitis	6 (0.5)
Cardiovascular Disease	136 (11.5)
Hearing issues	51 (4.3)
Multiple Sclerosis	6 (0.7)
Other diseases	291 (24.6)

Social care and caregiving	N (%)
Social care users ("Have you received social assistance services from a national, local or non-governmental body in the last 2 years?"), Yes	42 (3.6)
Caregivers ("Do you care for a family member, friend, or acquaintance who is not self-sufficient [due to a disease]?")	185 (15.7)

Table 1. Continued

Health and well-being	Mean (SD)	(Min-Max)	Median	Mode
EQ-HWB-S Index (Mukuria et al)	0.87 (0.12)	(0.11-1.00)	0.90	1
EQ-HWB Subscales summary scores*				
EQ-HWB (pain/discomfort)	6.75 (2.53)	(4-17)	6	4
EQ-HWB (psychosocial health)	26.76 (8.22)	(16-64)	25	20
EQ-HWB (activities)	3.36 (1.09)	(3-12)	3	3
EQ-5D-5L index (Finch et al)	0.93 (0.11)	(−0.23-1)	0.96	1
EQ VAS	81.83 (13.53)	(20-100)	85	90

Perfect health				N (%)
EQ-HWB-S Index (111111111)				88 (7.5)
EQ-5D-5L Index (11111)				410 (34.7)

EQ-HWB Act indicates EQ Activities LSS; EQ-HWB, EQ Health and Well-being; EQ-HWB-S, EQ Health and Well-being short form; EQ-HWB PD, EQ pain/discomfort level summary score; EQ-HWB MH, EQ psychosocial LSS; HRQL, Health-Related Quality of Life.
*EQ-HWB subscales and EQ-HWB-S LSS have been scored as level summary scores, EQ-5D-5L with utility values set for Italy and EQ-HWB-S with pilot data utility value set for the United Kingdom. More details on scoring in Methods.

response pattern distributions, ceiling/floor effects, correlation strength between related items/constructs, and known-group comparison. Respondent characteristics were detailed using descriptive statistics. Continuous variables were reported as means and standard deviations (SD), whereas categorical variables were presented as frequencies and proportions. With 1182 participants, the sample size surpassed Fayer's minimum requirement of 100 for validity studies.[30] The analyses were carried out using SAS Version 9.4 and RStudio Version 2021.09.

Response patterns

Floor/ceiling effects were evaluated both at the item and measure level, to determine their potential impact on the sensitivity to changes over time or differences between groups.[31] Response patterns were analyzed using absolute and relative frequencies for each item level.[32] A floor/ceiling effect at the item level was identified if 50% of respondents reported "no problems" for an item.[33] At the instrument level, a floor/ceiling effect was considered if 15% of respondents achieved the highest possible score.[31] Because the ceiling effect depends on the health status of the sample,[34] we expected some ceiling effects in this general population. Based on previous research, we hypothesized that items on the EQ-HWB-S would exhibit lower ceiling effects than those on the EQ-5D-5L.[22] Ceiling effects were further examined among those reporting full health on the EQ-5D-5L.

Convergent validity

Convergent validity, assessed through correlation analyses, evaluated the strength of the relationship between measures.[35] The correlation between EQ-HWB and EQ-5D-5L scores, as well as individual items, was examined using Spearman rank-order correlation. The interpretation of correlations (r_s) followed Cohen's guidelines: strong ($\geq$0.50), moderate (0.30-0.49), weak (0.10-0.29), and none ($<$0.10), with strong correlations suggesting related concepts measured by the measures.[36] We expected moderate-to-strong correlations between overlapping dimensions, such as EQ-HWB's item "Getting around inside and outside" and EQ-5D-5L's "Mobility," or EQ-HWB's "Activities" and EQ-5D-5L's "Usual activities." Furthermore, we anticipated at least moderate correlations between EQ-5D-5L and EQ-HWB-S index scores.

Known-groups comparison

Known-group comparison was used to evaluate the ability of individual items and aggregate scores (such as LSS sub-scales, index) to differentiate between groups that were expected to vary in terms of disease presence, social care use, and caregiver status.[37] Groups were identified by presence of self-reported chronic health conditions, having symptomatic chronic health conditions, recent social care usage, and caregiver role.[22] We hypothesized that individuals with chronic health conditions, social care users, and caregivers would report lower index scores on both EQ-5D-5L and EQ-HWB-S, as well as higher scores on the HWB LSS, indicative of poorer health and well-being compared with those not in these groups. We further compared the EQ-HWB-S index and subscale LSS scores with the EQ-5D-5L index scores and EQ VAS in terms of ES. Effect sizes (ES) for each known-group were determined with the thresholds: small (d = 0.2), medium (d = 0.5), large (d = 0.8), and very large (d = 1.40).[38] To calculate 95% confidence intervals, we applied bias-corrected and accelerated (BCA) bootstrap method using R's bootES package, involving 2000 replications with replacement for resampling.[39]

Results

Summary Statistics

The study sample included 1182 respondents, with 461 (39%) indicating the presence of 1 or more chronic conditions, 185 (15.7%) serving as caregivers, and 42 (3.5%) utilizing social care services in the previous 2 years (Table 1, with detailed demographic characteristics of social care users and carers are provided in Appendix Tables S3 and S4 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.03.002). Of the total sample, 606 (48.7%) were male, and the mean age was 48.3 (SD 16.1) years old. The 3 most common self-reported health conditions among the respondents were cardiovascular disease (n = 136, 11.5%), diabetes (n = 62, 5.2%), and asthma/chronic obstructive pulmonary disorder (n = 58, 4.9%). The mean EQ-5D-5L index scores were 0.93 (SD 0.11), and the EQ-HWB-S index score was 0.87 (SD 0.12). The EQ-HWB subscales scored were as follows: Psychosocial health LSS 26.76 (SD 8.22), Pain/Discomfort LSS 6.75 (SD 2.53), and Activities LSS 3.36 (SD 1.09).

Table 2. Distribution of EQ-HWB responses among respondents reporting full health on EQ-5D.

EQ-HWB Item	Response (N = 410)				
	N (%)	N (%)	N (%)	N (%)	N (%)
	No difficulty	Slight difficulty	Some difficulty	Much difficulty	Unable
See	331 (80.7)	61 (14.9)	18 (4.4)	0 (0.0)	0 (0.0)
Hear	377 (91.9)	25 (6.1)	6 (1.5)	2 (0.5)	0 (0.0)
Getting around inside and outside*	408 (99.5)	2 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)
Day-to-day activities*	400 (97.6)	8 (2.0)	1 (0.2)	1 (0.2)	0 (0.0)
Personal care	406 (99.0)	3 (0.7)	1 (0.2)	0 (0.0)	0 (0.0)
	None of the time	Only occasionally	Sometimes	Often	Most of the time
Sleep	343 (83.7)	38 (9.3)	26 (6.3)	3 (0.7)	0 (0.0)
Exhausted*	387 (94.4)	19 (4.6)	4 (1.0)	0 (0.0)	0 (0.0)
Lonely*	280 (68.3)	87 (21.2)	38 (9.3)	5 (1.2)	0 (0.0)
Unsupported	304 (74.2)	72 (17.6)	31 (7.6)	2 (0.5)	1 (0.2)
Remembering	269 (65.6)	100 (24.4)	37 (9.0)	4 (1.0)	0 (0.0)
Concentrating/thinking clearly*	291 (71.0)	65 (15.9)	51 (12.4)	3 (0.7)	0 (0.0)
Anxious*	275 (67.1)	91 (22.2)	42 (10.2)	2 (0.5)	0 (0.0)
Unsafe	325 (79.3)	51 (12.4)	27 (6.6)	7 (1.7)	0 (0.0)
Frustrated	308 (75.1)	62 (15.1)	37 (9.0)	3 (0.7)	0 (0.0)
Depressed*	329 (80.2)	49 (12.0)	25 (6.1)	6 (1.5)	1 (0.2)
Look Forward	388 (94.6)	15 (3.7)	6 (1.5)	1 (0.2)	0 (0.0)
Control*	213 (52.0)	111 (27.1)	61 (14.9)	22 (5.4)	3 (0.7)
Cope	236 (57.6)	102 (24.9)	59 (14.4)	12 (2.9)	1 (0.2)
Accepted[†]	3 (0.7)	3 (0.7)	11 (2.7)	98 (23.9)	295 (72.0)
Feel good[†]	1 (0.2)	12 (2.9)	126 (30.7)	271 (66.1)	0 (0.0)
Do things wanted to do[†]	2 (0.5)	15 (3.7)	55 (13.4)	146 (35.6)	192 (46.8)
Pain (frequency)	277 (67.6)	90 (22.0)	42 (10.2)	1 (0.2)	0 (0.0)
	No	Mild	Moderate	Severe	Very Severe
Pain (severity)*	276 (67.3)	109 (26.6)	22 (5.4)	3 (0.7)	0 (0.0)
	None of the time	Only occasionally	Sometimes	Often	Most of the time
Discomfort (frequency)	347 (84.6)	42 (10.2)	19 (4.6)	2 (0.5)	0 (0.0)
	No	Mild	Moderate	Severe	Very Severe
Discomfort (severity)	348 (84.9)	55 (13.4)	7 (1.7)	0 (0.0)	0 (0.0)

EQ-HWB indicates EQ Health and Well-being; EQ-HWB-S, EQ Health and Well-being Short.
*Part of the E-HWB-S.
[†]Reverse coded for summary score.

Floor and Ceiling Effects

Most of EQ-HWB items (17 out of 25) and 6 out of 9 EQ-HWB-S items, along with all the EQ-5D-5L items, exhibited ceiling effects (Appendix Table S1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.03.002). More specifically, the proportion of respondents opting for the highest response option (no problem or equivalent) varied greatly for different items. For EQ-HWB, this proportion ranged from 37.3% for item "Do things wanted to do," to 94.8% for "Personal Care." For the EQ-5D-5L, the range was between 56.7% for "Pain/Discomfort" and 95.7% for "Self-care."

Regarding conceptually overlapping items, in EQ-HWB 39.4% of respondents classified "Pain frequency" and 38.9% classified "Pain severity" as "no problem," whereas 56.8% reported no problems with "Pain/Discomfort" on the EQ-5D-5L. In contrast, a greater number of respondents reported "no problems" with "Discomfort frequency" (72.4%) and "Discomfort severity" (71.5%) in the EQ-HWB items. Fewer respondents reported "no problem" on the EQ-HWB equivalent "Anxious" (46.5%) item compared with the EQ-5D-5L item "Anxiety/Depression" (58.7%). In contrast, the EQ-HWB item "Depressed" had a higher proportion of respondents (68.4%) choosing the "no problem" option.

Table 3. Correlation coefficients between EQ-5D-5L and EQ-HWB items.

EQ-HWB	EQ-5D-5L				
	Mobility	Self-Care	Usual Activities	Pain/ Discomfort	Anxiety/ Depression
See	0.10	0.08	0.11	0.15	0.12
Hear	0.13	0.07	0.11	0.12	0.02
Getting around inside and outside*	0.61	0.33	0.41	0.29	0.09
Day-to-day activities*	0.36	0.31	0.53	0.32	0.18
Personal care	0.28	0.51	0.38	0.24	0.11
Sleep	0.11	0.18	0.23	0.14	0.28
Exhausted*	0.14	0.18	0.22	0.15	0.18
Lonely*	0.09	0.04	0.11	0.15	0.16
Unsupported	0.03	0.08	0.14	0.12	0.30
Remembering	0.03	0.09	0.15	0.11	0.66
Concentrating/ thinking clearly*	0.05	0.03	0.14	0.11	0.40
Anxious*	0.03	0.05	0.11	0.11	0.50
Unsafe	0.08	0.08	0.16	0.11	0.34
Frustrated	-0.02	0.07	0.09	0.01	0.29
Depressed*	0.01	0.10	0.12	0.10	0.28
Look Forward	0.14	0.13	0.17	0.11	0.16
Control*	0.11	0.09	0.19	0.22	0.26
Cope	0.10	0.09	0.19	0.22	0.24
Accepted	-0.03	0.04	0.07	0.05	0.24
Feel good	0.06	0.07	0.13	0.14	0.42
Do things wanted to do	0.11	0.10	0.16	0.12	0.24
Pain (frequency)	0.33	0.20	0.29	0.60	0.22
Pain* (severity)	0.30	0.20	0.28	0.55	0.23
Discomfort (frequency)	0.16	0.10	0.22	0.17	0.19
Discomfort (severity)	0.16	0.10	0.21	0.19	0.20

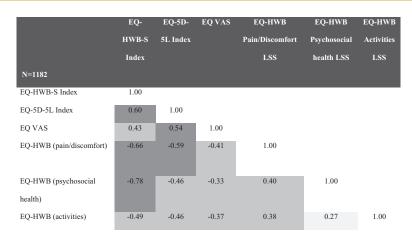EQ-HWB=EQ Health and Well-being; EQ-HWB-S=EQ Health and Well-being Short; all results are p <.0001.

* part of the EQ-HWB-S

Correlation coefficient

(Cohen's)

none	< 0.10
weak	0.10-0.29
moderate	0.30-0.49
strong	≥ 0.50-1

Table 4. Correlation coefficient between EQ-HWB and EQ-5D-5L.

N=1182	EQ-HWB-S Index	EQ-5D-5L Index	EQ VAS	EQ-HWB Pain/Discomfort LSS	EQ-HWB Psychosocial health LSS	EQ-HWB Activities LSS
EQ-HWB-S Index	1.00					
EQ-5D-5L Index	0.60	1.00				
EQ VAS	0.43	0.54	1.00			
EQ-HWB (pain/discomfort)	-0.66	-0.59	-0.41	1.00		
EQ-HWB (psychosocial health)	-0.78	-0.46	-0.33	0.40	1.00	
EQ-HWB (activities)	-0.49	-0.46	-0.37	0.38	0.27	1.00

EQ-HWB = EQ Health and well-being, EQ-HWB-S = EQ Health and well-being short form. LSS = Level Summary Score; EQ-HWB subscales have been scored as level summary scores, EQ-5D-5L with utility values set for Italy and EQ-HWB-S with pilot data utility value set for UK. More details on scoring in Methods.

Correlation coefficient (Cohen's)

none	< 0.10
weak	0.10-0.29
moderate	0.30-0.49
strong	≥ 0.50-1

At the instrument level, the EQ-HWB-S did not show a ceiling effect, with only 88 respondents (7.5%) reporting full health and well-being compared with the EQ-5D-5L in which 410 respondents (34.7%) reported full health. Among these 410 individuals who reported full health on the EQ-5D-5L, certain EQ-HWB items still demonstrated a ceiling effect. Nonetheless, some items were endorsed for having at least some problems; for example, 48.1% of respondents reported feeling at least occasionally that they had "No Control" of their life (Table 2). No floor effect was detected.

Convergent Validity

Convergent validity was evidenced by the strength of correlation between EQ-HWB and EQ-5D-5L items (Table 3). As expected, strong correlations were observed between conceptually similar items, specifically: EQ-HWB "Getting around inside and outside" and EQ-5D "Mobility" ($r_s = 0.61$), EQ-HWB "Day-to-day activities" and EQ-5D "Usual Activities" ($r_s = 0.53$), and EQ-HWB "Personal Care" and EQ-5D "Self-care" ($r_s = 0.51$). The strongest correlation was found between EQ-HWB item "Remembering" and EQ-5D-5L item "Anxiety/Depression" ($r_s = 0.66$), whereas the weakest correlation was between EQ-HWB items "Frustrated" and "Accepted" and EQ-5D-5L items (Mobility, Self-care, Usual Activities, and Pain/Discomfort) all with $r_s < 0.1$.

Notably, EQ-5D item *Anxiety/Depression* showed a strong correlation with EQ-HWB items "Remembering" ($r_s = 0.66$) and "Anxious" ($r_s = 0.50$), whereas having only a moderate correlation with the item "Depressed" ($r_s = 0.28$). Similarly, EQ-5D item "Pain/Discomfort" had a strong correlation with EQ-HWB items "Pain

severity" ($r_s = 0.55$) and "Pain frequency" ($r_s = 0.60$), whereas it was weakly correlated with EQ-HWB "Discomfort frequency" ($r_s = 0.17$) and "Discomfort severity" ($r_s = 0.19$). At the instrument level (Table 4), moderate to strong correlations were observed between EQ-HWB-S index and EQ-5D-5L ($r_s = 0.60$) and EQ VAS ($r_s = 0.43$). Out of EQ-HWB subscales, Pain/Discomfort LSS showed strongest correlation with EQ-5D-5L ($r_s = -0.66$), whereas EQ-HWB-S index with Psychosocial LSS ($r_s = -0.78$). All correlation coefficients were statistically significant ($P < .001$).

Known-groups Comparison

Both EQ-HWB/EQ-HWB-S and EQ-5D-5L items showed similar performance in differentiating among individuals with and without chronic health conditions, with ES ranging from 0.2 to 0.5 (Appendix Table S2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.03.002). The EQ-HWB items demonstrating the highest ES in differentiating individuals with symptomatic health conditions from the healthy group were: Getting around inside and outside, Day-to-day activities, Pain frequency and severity, and Discomfort frequency and severity. Corresponding EQ-5D-5L items showing similar distinctions included were Mobility, Usual Activities, Pain/Discomfort.

In terms of index scores, both EQ-HWB-S and EQ-5D-5L showed comparable discriminatory power based on the presence of symptomatic chronic conditions (ES d = 0.68, 95% CI 0.55-0.81 for EQ-HWB-S vs d = 0.71, 95% CI 0.60-0.81 for EQ-5D-5L), as shown in Tables 5 and 6. When evaluating individual self-reported chronic health conditions, both the EQ-HWB-S and EQ-5D-5L index scores exhibited similar small to medium ES, ie, 0.2 to 0.5

Table 5. Mean scores for EQ-HWB and EQ-5D-5L by caregiver status and comorbidities.

		N	EQ-HWB index score	EQ-5D-5L index score	EQ-HWB psychosocial subscale LSS	EQ-HWB pain/ discomfort subscale LSS	EQ-HWB activities subscale LSS
			Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Caregiver	No	997	0.88 (0.12)	0.93 (0.11)	26.47 (8.11)	6.71 (2.52)	3.37 (1.12)
	Yes	185	0.86 (0.11)	0.93 (0.08)	28.29 (8.62)	6.94 (2.54)	3.30 (0.82)
Social care user	No	1140	0.88 (0.12)	0.93 (0.10)	26.57 (8.14)	6.68 (2.46)	3.33 (1.02)
	Yes	42	0.78 (0.18)	0.85 (0.22)	31.76 (8.87)	8.66 (3.35)	3.97 (2.15)
Any disease	Yes	461	0.84 (0.15)	0.88 (0.15)	27.65 (9.12)	7.81 (2.78)	3.69 (1.50)
	No	721	0.90 (0.09)	0.96 (0.06)	26.18 (7.53)	6.07 (2.08)	3.14 (0.60)
Arthritis	Mild/ asymptomatic	33	0.83 (0.13)	0.90 (0.10)	26.90 (9.02)	7.54 (2.76)	3.51 (1.06)
	Moderate	31	0.79 (0.21)	0.82 (0.17)	27.61 (10.42)	9.25 (2.70)	4.06 (2.14)
	Severe	4	0.48 (0.36)	0.62 (0.34)	39.50 (17.36)	11.75 (4.64)	7.25 (4.34)
Asthma/COPD	Mild/ asymptomatic	24	0.85 (0.14)	0.92 (0.07)	25.91 (8.09)	8.79 (2.90)	3.58 (1.38)
	Moderate	29	0.85 (0.16)	0.89 (0.16)	28.86 (11.41)	8.34 (2.64)	3.41 (1.70)
	Severe	5	0.60 (0.23)	0.72 (0.26)	33.60 (17.21)	14.20 (2.04)	5.00 (1.00)
Cancer	Mild/ asymptomatic	13	0.85 (0.12)	0.86 (0.31)	25.46 (8.09)	7.00 (1.47)	3.61 (1.93)
	Moderate	13	0.82 (0.16)	0.90 (0.08)	27.07 (7.81)	8.53 (3.38)	4.23 (1.87)
	Severe	6	0.86 (0.06)	0.88 (0.08)	25.50 (8.16)	8.16 (2.78)	3.16 (0.40)
Depression	Mild/ asymptomatic	34	0.75 (0.19)	0.80 (0.21)	33.79 (8.82)	9.44 (2.86)	4.29 (2.276)
	Moderate	18	0.62 (0.22)	0.77 (0.18)	41.61 (11.42)	9.38 (3.69)	4.61 (2.00)
	Severe	2	0.28 (0.15)	-0.12 (0.16)	49.00 (4.24)	10.50 (0.70)	7.50 (6.36)
Diabetes	Mild/ asymptomatic	21	0.85 (0.13)	0.88 (0.14)	23.95 (7.17)	7.52 (2.56)	3.52 (0.87)
	Moderate	36	0.83 (0.15)	0.87 (0.16)	26.61 (8.59)	8.50 (2.90)	4.25 (2.08)
	Severe	5	0.89 (0.08)	0.93 (0.07)	22.60 (3.78)	7.20 (0.83)	3.80 (1.09)
Hearing issue	Mild/ asymptomatic	32	0.80 (0.17)	0.85 (0.18)	27.50 (9.32)	8.53 (2.73)	4.18 (2.24)
	Moderate	15	0.82 (0.18)	0.89 (0.14)	27.73 (10.62)	7.60 (2.26)	4.00 (2.17)
	Severe	4	0.57 (0.31)	0.74 (0.23)	44.50 (16.90)	10.50 (6.40)	5.50 (2.64)
Cardiovascular disease	Mild/ asymptomatic	73	0.87 (0.11)	0.90 (0.1)	25.39 (7.69)	7.17 (2.51)	3.47 (1.42)
	Moderate	52	0.81 (0.15)	0.86 (0.19)	28.03 (9.59)	8.11 (2.59)	3.90 (1.63)
	Severe	11	0.88 (0.11)	0.90 (0.16)	24.54 (8.58)	6.54 (2.54)	3.54 (0.82)
Other diseases	Mild/ asymptomatic	102	0.88 (0.09)	0.93 (0.07)	25.85 (7.65)	6.86 (2.44)	3.32 (0.78)
	Moderate	153	0.82 (0.15)	0.87 (0.12)	29.56 (9.85)	8.19 (2.62)	3.72 (1.34)
	Severe	36	0.66 (0.22)	0.72 (0.29)	33.36 (11.18)	10.33 (3.48)	5.80 (2.98)

Note. EQ-HWB subscales have been scored as level summary scores, EQ-5D-5L with utility values set for Italy and EQ-HWB-S with pilot data utility value set for UK. More details on scoring in Methods.
EQ-HWB indicates EQ Health and Well-being, EQ-HWB-S, EQ Health and Well-being short form. LSS, Level Summary Score.

across conditions, except when distinguishing between those with and without depression, in which both measures revealed a very large ES (d >1.4). Among the EQ-HWB LSS subscales, sensitivity varied depending on the condition; for example, the EQ-HWB Psychosocial LSS was most sensitive to depression, showing a large ES (d = 1.35, 95% Cl 0.98-1.12), whereas the Pain/Discomfort LSS was most sensitive to asthma/COPD with a large ES (d = 0.97, 95% Cl 0.64-1.30). The Activities LSS subscale was most sensitive to respondents with hearing issues with large ES (d = 0.85, 95% Cl 0.31-1.4).

The EQ-HWB-S index score demonstrated a slightly higher ES than the EQ-5D-5L index score when differentiating between social care users and non-users (ES d = 0.84, 95% Cl 0.40-1.30 vs d = 0.74, 95% CL 0.16-1.33). Differences in index scores for both

EQ-5D-5L and EQ-HWB-S based on caregiver status were trivial (ES d < 0.2), whereas small ES was noted for EQ-HWB Psychosocial LSS subscale (d = 0.22, 95% Cl 0.06-0.38).

Discussion

This study is among of the first to provide evidence on the psychometric properties of the revised experimental version of EQ-HWB/EQ-HWB-S in comparison with the EQ-5D-5L. Initial evidence supports the validity of the EQ-HWB/EQ-HWB-S as an outcome measure. Our findings showed that the EQ-HWB-S had fewer respondents reporting ceiling effects than the EQ-5D-5L, likely because of its broader coverage of health and well-being

Table 6. Effect Sizes for EQ-HWB and EQ-5D-5L index scores based on caregiver status, chronic conditions, and social care users.

Group	N Yes	EQ-HWB-S Index score	EQ-5D-5L Index score	EQ-HWB Psychosocial subscale LSS	EQ-HWB Pain/discomfort subscale LSS	EQ-HWB Activities subscale LSS
Caregivers	185	-0.14 (-0.30, 0.01)	-0.02 (-0.16, 0.10)	0.22 (0.06, 0.38)	0.09 (-0.06, 0.25)	-0.06 (-0.17, 0.08)
Social Care users	42	-0.84 (-1.30, -0.39)	-0.74 (-1.33, -0.16)	0.64 (0.32, 0.97)	0.79 (0.38, 1.20)	0.59 (0.07, 1.21)
Chronic condition (any)	461	-0.56 (-0.67, -0.44)	-0.70 (-0.79, -0.60)	0.18 (0.06, 0.30)	0.73 (0.60, 0.85)	0.52 (0.41, 0.62)
Symptomatic condition (any)	310	-0.68 (-0.81, -0.55)	-0.71 (-0.81, -0.60)	0.45 (0.30, 0.60)	0.89 (0.73, 1.04)	0.61 (0.51, 0.70)
Arthritis	67	-0.77 (-1.16, -0.38)	-0.78 (-1.16, -0.40)	0.16 (-0.13, 0.48)	0.78 (0.49, 1.08)	0.62 (0.19, 1.04)
Asthma/COPD	58	-0.38 (-0.77, -0.03)	-0.37 (-0.73, -0.04)	0.17 (-0.16, 0.52)	0.97 (0.64, 1.30)	0.25 (-0.07, 0.66)
Cancer	32	-0.31 (-0.73, 0.02)	-0.46 (-1.16, -0.03)	-0.08 (-0.39, 0.28)	0.44 (0.09, 0.82)	0.40 (-0.08, 1.01)
Depression	54	-1.73 (-2.22, -1.21)	-1.75 (-2.29, -1.09)	1.35 (0.98, 1.70)	1.16 (0.80, 1.49)	1.15 (0.58, 1.71)
Diabetes	62	-0.29 (-0.59, -0.01)	-0.48 (-0.84, -0.15)	-0.18 (-0.40, 0.09)	0.55 (0.28, 0.83)	0.60 (0.2, 1.00)
Cardiovascular disease	136	-0.24 (-0.44, -0.05)	-0.47 (-0.70, -0.24)	-0.06 (-0.28, 0.13)	0.33 (0.15, 0.52)	0.30 (0.07, 0.54)
Hearing issue	51	-0.79 (-1.24, -0.34)	-0.73 (-1.17, -0.29)	0.27 (-0.01, 0.66)	0.69 (0.36, 1.03)	0.85 (0.31, 1.42)

EQ-HWB = EQ Health and well-being, EQ-HWB-S = EQ Health and well-being short form. LSS = Level Summary Score; EQ-HWB subscales have been scored as level summary scores, EQ-5D-5L with utility values set for Italy and EQ-HWB-S with pilot data utility value set for UK. More details on scoring in Methods.

Effect Size

none	<0.20
small	0.20-0.49
medium	0.50-0.79
large	0.80-1.39
very large	1.40>

dimensions. The EQ-HWB-S, with its 9 items, may capture health states more comprehensively than the 5-item EQ-5D-5L. Moreover, the difference in the number of items—9 in the EQ-HWB-S vs 5 in the EQ-5D-5L—also contributes to the lower incidence of maximum score reporting in the EQ-HWB. The observed ceiling effect in the EQ-5D is consistent with previous studies showing more pronounced ceiling effects in general public samples, which typically have better health status.[34] Prior research found that between 31% to 63% of general public sample participants reported "11111" (full health) on the EQ-5D-5L descriptive system.[34] Thus, our finding of 35% lies at the lower bound of the range.

A high degree of convergence was observed across conceptually related EQ-HWB items and EQ-5D, which supports EQ-HWB measures' construct validity.[35] Considering that the EQ-HWB and EQ-5D-5L share concepts, with items covering mobility, activities, pain, anxiety, and depression in both measures, a significant level of overlap was expected. Indeed, moderate to strong correlations were found. However, there are conceptual difference in items such as "Mobility." Although the EQ-5D-5L specifically asks about an individual's ability to walk, the EQ-HWB's approach to mobility is broader, encompassing not only walking but also the ability to get around with or without aids. This distinction

reflects that the EQ-HWB's Mobility domain potentially offers a more comprehensive assessment, which may be more relevant for the older population. Some EQ-HWB items that were not correlated with EQ-5D-5L indicated that the EQ-HWB covers wider aspects of health and well-being, which may be relevant for broader QoL, such as social care and evaluating the impact on caregivers, despite these concepts not being considered necessary by health technology agencies that solely focus on health.[22]

Interestingly, there were some differences in how the EQ-HWB items related to equivalent composite items in the EQ-5D-5L. The development of the EQ-HWB aimed to avoid combining conceptually different aspects, such as anxiety and depression or pain and discomfort. The EQ-5D item Anxiety/Depression correlated strongly with the EQ-HWB item Anxiety, but weakly with Depression. Similar trends were observed for the other EQ-5D-5L composite item, Pain/Discomfort, showing a strong correlation with EQ-HWB items' Pain (severity and frequency) and weak correlation with Discomfort (frequency and severity). These results support recent studies suggesting that EQ-5D's composite dimensions are inherently ambiguous and that the EQ-5D-5L Pain/Discomfort item is mainly endorsed to report pain.[40] Additionally, discomfort is not solely pain-related; it is an umbrella term encompassing around 100 different non-pain physical sensations, mental problems, or feelings.[41,42] Moreover, the distinct recall periods of the 2 instruments potentially contribute to these variations. The EQ-5D-5L assesses health based on the day of administration, offering a "today'" perspective, whereas the EQ-HWB uses a 7-day recall period, providing view of health and well-being over a week. This difference in recall periods may suggest that respondents might report varying levels of anxiety and depression over the previous week, as opposed to their state on the day of the survey. Furthermore, prior research has shown that the level of agreement between reported health states can be particularly low for the Anxiety/Depression dimension when different recall periods are used.[43] Our study supports the notion that using separate items for anxiety, depression, physical pain, and physical discomfort in the EQ-HWB may be more appropriate for capturing these concepts, especially when considering their variability over time.

An intriguing finding was the strongest correlation between EQ-5D-5L item Anxiety/Depression and EQ-HWB item Remembering. Although anxiety and depression are distinct concepts from memory, symptoms of depression are closely associated with subjective memory complaints rather than actual memory performance.[44,45] Furthermore, depressive symptoms have been well documented as one of the candidate explanations for a weak relationship between self-assessment of memory and actual memory performance.[46]

In terms of known-groups, EQ-HWB/EQ-HWB-S and EQ-5D-5L items performed similarly, yielding small to medium ES, ie, 0.2 to 0.5 among those with and without chronic health conditions and social care users. The EQ-HWB LSS was the only measure able to differentiate caregivers and non-caregivers with small ES. These results indicate moving from EQ-HWB to EQ-HWB-S entails a loss of a small magnitude in discriminative properties for caregivers, and there is some merit in EQ-HWB measure subscales. In a recent study comparing the EQ-HWB-S and EQ-5D-5L measures in the United States, the discriminative ability of EQ-HWB-S was notably higher in differentiating between caregiver and noncaregiver populations, in contrast to its performance observed in an Italian cohort.[22] This observation may be influenced by cultural differences in the perception and practice of caregiving. In Italy, the strong tradition of familial bonds and mutual support within families could lead to a different caregiving experience, potentially perceived as a natural part of family life. Conversely, in the United States, where individualistic values are more pronounced, caregivers might experience caregiving roles differently. In future research, it would be valuable to explore various known group validity anchors for the assessment of psychometric properties of different measures in caregivers, particularly in their ability to differentiate caregivers based on caregiver burden.

There are limitations to this study that need to be acknowledged. First, its cross-sectional design restricts the ability to assess the measures' sensitivity to changes over time, a key aspect for evaluating health conditions and interventions in the health sector. The interviewer-assisted mode of administration might have influenced responses because research indicates that self-completion often yields lower well-being and health status scores, particularly among women.[47] Social desirability bias may also have played a role.[48] The recruitment aimed to broadly represent the Italian adult population but may not fully capture the diverse experiences of social care users and caregivers. Group categorization—chronic health condition, caregivers, and social care users—was based on self-reported data, which, although common in health research, can introduce biases without medical record verification. Moreover, the relatively small caregiver and social care user sample sizes may affect the study's validity results, although the sample provides a reasonable basis for preliminary exploration in the Italian context. Future research with larger sample sizes is recommended. Furthermore, the EQ-HWB was completed at interviews' end, potentially affecting responses because of fatigue. The EQ-HWB-S value set, derived from a UK pilot study, may not reflect Italian population preferences. Additionally, the correlations between EQ-5D-5L and EQ-HWB/EQ-HWB-S items should be interpreted with caution because of potential shared method variance or measurement error.[49] Finally, we must highlight that the EQ-HWB currently holds the status of an experimental measure. This designation implies that aspects of the instrument, including its wording and response options, may change with ongoing validation and research.

In conclusion, our study provides initial evidence supporting the validity of EQ-HWB/EQ-HWB-S as an outcome measure for the Italian population. The results suggest a high degree of convergence between conceptually related EQ-5D and EQ-HWB items. The EQ-HWB-S performed comparably to EQ-5D-5L among patient groups and was better in differentiating social care users. A slight decrease in discriminative properties for caregivers was observed when transitioning from EQ-HWB to EQ-HWB-S. Further research is needed to fully assess the performance of these measures in caregivers, especially using different anchors for caregiver burden. Additionally, further investigation is needed to evaluate EQ-HWB's performance in contexts beyond health, particularly in relation to social care and support services. Overall, this study contributes to the growing body of literature on the psychometric properties and performance of EQ-HWB/EQ-HWB-S, underscoring its potential as a valid measure for use in health and social care.

Author Disclosures

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

Supplemental Material

Article and Author Information

Author Affiliations: Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA (Kuharić, Pickard); Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois Chicago, IL, USA (Kuharić); Maths in Health BV, Amsterdam, The Netherlands (Pickard); Division of Population Health, University of Sheffield, Sheffield, England, UK (Mukuria); EuroQol Office, EuroQol Research Foundation, Rotterdam, The Netherlands (Finch).

Correspondence: Maja Kuharić, Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, 625 N Michigan Avenue, Chicago, IL, USA. Email: maja.kuharic@northwestern.edu

REFERENCES

1. Brazier J, Ara R, Rowen D, Chevrou-Severac H. A review of generic preference-based measures for use in cost-effectiveness models. *Pharmacoeconomics*. 2017;35(suppl 1):21–31.
2. Kuharic M, Joshi M, Zecic F, Pickard A. PCR198 contemporary trends in the use of generic patient-reported outcome measures in pediatric clinical trials: an analysis of Clinicaltrials. Gov. *Value Health*. 2022;25(12):S428.
3. Dawoud D, Lamb A, Moore A, et al. Capturing what matters: updating NICE methods guidance on measuring and valuing health. *Qual Life Res*. 2022;31(7):2167–2173.
4. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. *Eur J Health Econ*. 2018;19(4):557–570.
5. Payakachat N, Ali MM, Tilford JM. Can the EQ-5D detect meaningful change? A systematic review. *Pharmacoeconomics*. 2015;33(11):1137–1154.
6. Tosh J, Brazier J, Evans P, Longworth L. A review of generic preference-based measures of health-related quality of life in visual disorders. *Value Health*. 2012;15(1):118–127.
7. Papaioannou D, Brazier J, Parry G. How to measure quality of life for cost effectiveness analyses in personality disorders: a systematic review. *J Pers Disord*. 2013;27(3):383–401.
8. Brazier J, Connell J, Papaioannou D, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technol Assess*. 2014;18(34):vii-188.
9. Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. *Value Health*. 2011;14(6):907–920.
10. Hounsome N, Orrell M, Edwards RT. EQ-5D as a quality of life measure in people with dementia and their carers: evidence and key issues. *Value Health*. 2011;14(2):390–399.
11. Kuspinar A, Mayo NE. A review of the psychometric properties of generic utility measures in multiple sclerosis. *Pharmacoeconomics*. 2014;32(8):759–773.
12. Brazier J, Tsuchiya A. Improving cross-sector comparisons: going beyond the health-related QALY. *Appl Health Econ Health Policy*. 2015;13(6):557–565.
13. Peasgood T, Mukuria C, Carlton J, et al. What is the best approach to adopt for identifying the domains for a new measure of health, social care and carer-related quality of life to measure quality-adjusted life years? Application to the development of the EQ-HWB? *Eur J Health Econ*. 2021;22(7):1067–1081.
14. Makai P, Brouwer WB, Koopmanschap MA, Stolk EA, Nieboer AP. Quality of life instruments for economic evaluations in health and social care for older people: a systematic review. *Soc Sci Med*. 2014;102:83–93.
15. Finch AP, Mulhern B. Where do measures of health, social care and wellbeing fit within a wider measurement framework? Implications for the measurement of quality of life and the identification of bolt-ons. *Soc Sci Med*. 2022;313:115370.
16. Finch AP, Brazier J, Mukuria C. Selecting bolt-on dimensions for the EQ-5D: testing the impact of hearing, Sleep, cognition, energy, and relationships on preferences using pairwise choices. *Med Decis Mak*. 2021;41(1):89–99.
17. Mulhern BJ, Sampson C, Haywood P, et al. Criteria for developing, assessing and selecting candidate EQ-5D bolt-ons. *Qual Life Res*. 2022;31(10):3041–3048.
18. Carlton J, Peasgood T, Mukuria C, et al. Generation, selection, and face validation of items for a new generic measure of quality of life: the EQ-HWB. *Value Health*. 2022;25(4):512–524.
19. 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.
20. 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.
21. Masutti S, Falivena C, Purba FD, Jommi C, Mukuria C, Finch AP. Content validity of the EQ-HWB and EQ-HWB-S in a sample of Italian patients, informal caregivers and members of the general public. *J Patient Rep Outcomes*. 2024;8(1):36.
22. 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.
23. Finch AP, Meregaglia M, Ciani O, Roudijk B, Jommi C. An EQ-5D-5L value set for Italy using videoconferencing interviews and feasibility of a new mode of administration. *Soc Sci Med*. 2022;292:114519.
24. EuroQol Research Foundation. Intellectual property. experimental version. https://euroqol.org/eq-5d-instruments/ip-protection/experimental-version/. Accessed May 28, 2023.
25. 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.
26. Feng YS, Kohlmann T, Peasgood T, Engel L, Mulhern B, Pickard AS. Scoring the EQ-HWB-S: can we do it without value sets? A non-parametric item response theory analysis. *Qual Life Res*. 2024;21. https://doi.org/10.1007/s11136-024-03601-7.
27. Gudex C. The descriptive system of the EuroQOL instrument. In: Kind P, Brooks R, Rabin R, eds. *EQ-5D Concepts and Methods: a Developmental History*. Berlin, Germany: Springer; 2005:19–27.
28. Rabin R, Oemar M, Oppe M. EQ-5D-3L User Guide Basic Information on How to Use the EQ-5D-3L Instrument. EuroQol Group. Published 2011. http://euroqol.org. Accessed May 28, 2023.
29. Devlin N, Parkin D, Janssen B. *Methods for Analysing and Reporting EQ-5D Data*. Cham (CH), Switzerland: Springer; 2020.
30. Fayers P, Machin D. Introduction. In: *Quality of Life Assessment, Analysis and Interpretation*. UK: John Willy & Sons; 2000:3–10.
31. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.
32. Hughes DJ. Psychometric validity: establishing the accuracy and appropriateness of psychometric measures. In: Irwing P, Booth T, Hughes DJ, eds. *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*. Hoboken, NJ: Wiley Blackwell; 2018:751–779.
33. Cappelleri JC, Bushmakin AG, Gerber RA, et al. Psychometric analysis of the Three-Factor Eating Questionnaire-R21: results from a large diverse sample of obese and non-obese participants. *Int J Obes (Lond)*. 2009;33(6):611–620.
34. Konnopka A, Koenig HH. The "no problems"-problem: an empirical analysis of ceiling effects on the EQ-5D 5L. *Qual Life Res*. 2017;26(8):2079–2084.
35. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: a Practical Guide to Their Development and Use*. 5th ed. Oxford: Oxford University Press; 2015.
36. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155–159.
37. Mukuria C, Rowen D, Peasgood T, Brazier J. An empirical comparison of well-being measures used in the UK. Project report. Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU) (27). Policy Research Unit in Economic Evaluation of Health and Care Interventions. https://eprints.whiterose.ac.uk/99499/. Accessed May 28, 2023.
38. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen*. 2012;141(1):2.
39. Kirby KN, Gerlanc D. BootES: an R package for bootstrap confidence intervals on effect sizes. *Behav Res Methods*. 2013;45(4):905–927.
40. McDonald R, Mullett TL, Tsuchiya A. Understanding the composite dimensions of the EQ-5D: an experimental approach. *Soc Sci Med*. 2020;265:113323.

41. Rencz F, Janssen MF. Analyzing the pain/discomfort and anxiety/depression composite domains and the meaning of discomfort in the EQ-5D: a mixed-methods study. *Value Health*. 2022;25(12):2003–2016.

42. Rencz F, Mukuria C, Bató A, Poór AK, Finch AP. A qualitative investigation of the relevance of skin irritation and self-confidence bolt-ons and their conceptual overlap with the EQ-5D in patients with psoriasis. *Qual Life Res*. 2022;31(10):3049–3060.

43. Spronk I, Geraerds AJLM, Bonsel GJ, de Jongh MAC, Polinder S, Haagsma JA. Correspondence of directly reported and recalled health-related quality of life in a large heterogeneous sample of trauma patients. *Qual Life Res*. 2019;28(11):3005–3013.

44. Schweizer S, Kievit RA, Emery T, Henson RN, Henson RN. Symptoms of depression in a large healthy population cohort are related to subjective memory complaints and memory performance in negative contexts. *Psychol Med*. 2018;48(1):104–114.

45. Perini G, Cotta Ramusino M, Sinforiani E, Bernini S, Petrachi R, Costa A. Cognitive impairment in depression: recent advances and novel treatments. *Neuropsychiatr Dis Treat*. 2019;15:1249–1258.

46. Yoon JS, Charness N, Boot WR, Czaja SJ, Rogers WA. Depressive symptoms as a predictor of memory complaints in the PRISM sample. *J Gerontol B Psychol Sci Soc Sci*. 2019;74(2):254–263.

47. Pudney S. An experimental analysis of the impact of survey design on measures and models of subjective wellbeing. ISER Working Paper Series. IISER. Published 2010. https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2010-20. Accessed May 28, 2023.

48. Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health (Oxf)*. 2005;27(3):281–291.

49. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol*. 2003;88(5):879–903.