ScienceDirect

Contents lists available at **sciencedirect.com**
Journal homepage: **www.elsevier.com/locate/jval**

Themed Section

The Development of a New International Generic Measure (EQ Health and Wellbeing): Face Validity and Psychometric Stages in Argentina

Federico Augustovski, MD, MSc, PhD, Fernando Argento, MSc(c), Rocío B. Rodríguez, MSc, Luz Gibbons, MSc, Clara Mukuria, MA, MSc, PhD, María Belizán, MSc

ABSTRACT

Objectives: This study aimed to present the face validity and psychometric stages performed in Spanish in Argentina, the only Spanish-speaking country of an international collaboration that undertook the construction of a new measure that can be used in economic evaluation across health, social care, and public health, the EQ Health and Wellbeing™ (EQ-HWB). We also explored the relationship among 3-level version EQ-5D (EQ-5D-3L), 5-level version EQ-5D (EQ-5D-5L), and EQ-HWB.

Methods: Face validity was based on semistructured face to face interviews of a purposive sample to explore translatability of language and concepts of 97 candidate items, translated into Argentina Spanish. The psychometric evaluation using an online panel assessed the psychometric properties of 64 items that were carried forward (floor and ceiling effects, item correlations, known-group differences in relevant prespecified subgroups by the international and local teams, exploratory and confirmatory factor analysis, and item response theory). EQ-5D-3L, EQ-5D-5L, and EQ-HWB correlations were explored.

Results: In the face validity stage, 24 interviews with carers, general public, patients, and users of social services were included. Most items showed adequate face validity. In the psychometric assessment, 497 participants were recruited (64% reporting a long-term health condition). Most of the items showed adequate psychometrics in an Argentinian context. EQ-5D-3L and EQ-5D-5L had strong correlations, and EQ-HWB was moderately correlated to EQ visual analog scale. The Argentina team recommended 23 of the final 25 items.

Conclusions: The assessment of Spanish items contributed to the overall development of EQ-HWB and helped inform the design of an internationally relevant 25-item and a short 9-item measure intended to be used in economic evaluations.

Keywords: Argentina, EQ Health and Wellbeing, face validity, psychometric, quality of life, quality-adjusted life-year, wellbeing.

VALUE HEALTH. 2022; ■(■):■–■

Introduction

Health-related quality of life (HRQOL) is a widely used concept related to health and wellbeing, used as outcome measure in health systems worldwide.[1] In many countries, generic HRQOL instruments are used to describe population health, and preference-based instruments results are used for the estimation of a common metric to assess health, which together with the lifetime or survival dimension builds the quality-adjusted life-years (QALYs).[2,3] QALYs are used in cost-effectiveness analysis to guide resource allocation decisions in health. Generic preference-based measures have been found to be valid for assessing many health conditions but they have shown limited validity in some conditions (eg, EQ-5D has shown some limitations in assessing dementia or hearing and vision problems).[4,5]

In many countries, budget holders have to take care not only of healthcare decisions but also of wider issues such as social care. For social care and in some cases healthcare of long-term conditions, the outcomes of care are not only improved health per se but improved quality of life (QOL) and wellbeing for the recipients from a better meeting of their wants and needs (eg, nutrition, accommodation, relationships, independence). There are also important consequences for the QOL of informal (unpaid) carers who support the healthcare system and provide much of the social care. Measures have been developed for use in the evaluation of social care interventions, for example, the Adult Social Care Outcomes Toolkit and the CarerQoL-7D for informal carer QOL, but different measures limit within and across sector comparisons.[6,7] If the idea is to use economic evaluation to inform wider resource allocation decisions(such as in the health and social care sectors), the metric should not only incorporate HRQOL issues but also accommodate relevant aspects for this wider scope, such as including carers QOL relevant dimensions or dimensions that are more important to the social care dimension of life such as autonomy or safety.[8,9]

1098-3015/$36.00 - see front matter Copyright © 2022, International Society for Pharmacoeconomics and Outcomes Research, Inc. Published by Elsevier Inc.

ARTICLE IN PRESS

Bearing this in mind, an international collaborative project (estimated QALY project) was launched to develop a new generic measure, the EQ Health and Wellbeing™ (EQ-HWB), that can be used in economic evaluation across health, social care, and public health to estimate QALYs, based on the views of users and beneficiaries of these services including informal carers. The aim was to develop a long and a short measure that would be amenable to valuation and that could be used to assess health and wellbeing outcomes within and across these populations of interest to address the noted limitations with existing measures. The project was led by the University of Sheffield in the United Kingdom (UK), with the participation of researchers from 6 countries (Argentina, Australia, China, Germany, UK, and United States).[10] The project had different stages with input from a public involvement group, a virtual advisory group, and a steering group (see more details in John Brazier et al[11]). The last stages of the project aimed to identify relevant items for the final measures. Qualitative and quantitative testing methods included individuals with various physical and mental health conditions, carers, and social care users. Data collection was conducted in the 6 participating countries. Argentina was the only Spanish-speaking country. For more information and methods about the international collaboration and its different stages, you can access other articles related to the project in this themed issue.[12-14]

The objective of this article is to describe in detail the data collection process and results in Argentina of both the face validity and psychometric stages and to highlight how the local findings were considered for the final instrument. The secondary objectives of this article, which are specific to the Argentinean psychometric component, are to compare the 3-level version EQ-5D (EQ-5D-3L™) with 5-level version EQ-5D (EQ-5D-5L™) versions and to explore the relationship between the new EQ-HWB measure and the EQ-5D-5L visual analog scale (VAS).

Methods

Stage 1 was a qualitative literature review that identified potential domains for the EQ-HWB which included activity, relationships, cognition, self-identity, autonomy, feelings, and physical sensations covering 32 subdomains[12] whereas stage 2 identified potential items for each subdomain (n = 97).[13] This is a mixed methods study that undertook a common core protocol developed for the international collaborative study for the next 2 stages: face validity (qualitative, stage 3 of the project) and psychometric (quantitative, stage 4).[11,13,14] Moreover, during the psychometric stage, we included as a country-specific aim the comparison between the new EQ-HWB measure and the EQ-5D-5L and EQ-5D-3L.

Materials used during the 2 stages were developed by the international study in English with the input of participant countries teams. For the study in Argentina, an independent translation company translated EQ-HWB potential items into Spanish and then back translated them into English. The researchers from Argentina reviewed the translations and proposed some editions that were the agreed with the translators until reaching final versions. Partial findings and recommendations from Argentinean study were iteratively discussed with international coordinators and researchers from other sites to inform following steps.

Face Validity

Data collection took place in Buenos Aires City, Argentina, during August and September 2018. The qualitative component aimed to explore potential users' views on the items and understand how different groups interpreted the items and the item's response choices (5 levels of severity/frequency/difficulty) in an Argentinian context. A 7-day recall period was adopted to because this was considered relevant for wellbeing-related items, for example, autonomy or self-identity, while allowing for a short enough time to prevent recall bias. For the face validation of the 97 initial candidate items, we conducted semistructured one-to-one cognitive interviews. We used a generic project-specific interview guideline that explored whether items measure what was intended to measure and whether they were relevant and understandable.[13] Guidelines were translated by local researchers. Interviewers received a virtual training for data collection, analysis, and findings report from the central coordination.

Participants were recruited using different strategies: researchers contacted known individuals; a snowball sampling was adopted asking participants to help researchers to identify further individuals (particularly social care users); and finally recruiting participants at health promotion public facilities in the city of Buenos Aires ("Estaciones Saludables"). Individuals were included in the sample to meet quotas of characteristics such as age, sex, caregiver role, and social services users. Participants were assigned different domains (with their set of items) to be tested by interviewees because it was not possible for participants to review all the items. We intended that each domain and therefore each item were reviewed by at least 12 participants, which is appropriate in the context of qualitative research (Table 1). Interviews lasted approximately 1 hour. The interviewer took notes in a Word® structured table during the interviews and also audiotaped the interviews. Audio files were uploaded to the qualitative data analysis software ATLAS.TI 8.0. Interviews were partially transcribed with focus on relevant quotes related to key aims of face validation. Matrices were developed using notes and completed by listening to the audio. The matrices included comments regarding ambiguities in the wording of items or their response options, irrelevant items, embarrassing or stressful items, and complicated or confusing items. Matrices completion for the first 2 interviews was performed by 3 researchers to agree on what to report and then each interview registered by one of the researchers. During group meetings, 3 researchers categorized each item in an internal scale, assigning a score to each, 1 being "it has no problems," 2 being "it has some problem/s that can be modified," with specific detail, and "3" being "it is very problematic—would not be recommended." Findings were shared with the international study leaders for the selection of the items to be included in the next stage. Selection and wording of final items were decided by the multicountry collaboration study taking into account the findings of all participant countries. Items were modified where it was deemed that this improved face validity, for example, adding examples, but no further face validation was undertaken on modifications. More detailed description of the items' construction is published elsewhere.[13]

Psychometric Assessment

Recruitment was performed by a market research specialist company that has a panel of potential participants from August to October 2019. The aim of the psychometric stage was to test the items that were taken forward from face validity in terms of their psychometric validity.[14] This included assessment of individual item performance and dimensionality of the items. The questionnaire used to this stage included the EQ-HWB test items taken forward from the face validity stage, background information tailored to Argentina context and other health and wellbeing questionnaires, and the EQ-5D-3L and EQ-5D-5L for the Argentine specific objectives.

Table 1. Characteristics of the study population in the face validity and psychometric stages in Argentina.

Variable	n	%
Face validity (n = 24)		
Age,* median (IQR)	55.5 (37.5-70.5)	
Female	15	62.5
Education		
High school incomplete or less	6	25
High school complete	8	33
Tertiary or incomplete university	1	4
Tertiary or complete university	9	38
Carer	8	33.3
Social services users	8	33.3
General public	8	33.3
Psychometric validity (n = 497)		
Age, median (IQR)	34 (28-45)	
Female	201	40.4
Education		
High school incomplete or less	22	4.4
High school complete	93	18.7
Tertiary or university incomplete	115	23.1
Tertiary or university complete	6	1.2
Residence		
CABA o Greater Buenos Aires	271	54.5
Pampa	136	27.4
Northwest	34	6.8
Northeast	14	2.8
Cuyo	28	5.6
Patagonian	14	2.8
Main medical coverage		
Public	94	18.9
Social security	168	33.8
Private health insurance	235	47.3
Medical expenses		
Nothing or do not know	109	22
Yes	388	78
Last month, median (IQR)	$1.832	($700-$4550)
WHOQOL-BREF (Do you have enough money for the things you need?)		
Not at all/a little	64	31.5
Moderately	252	50.7
Mostly	46	9.3
Completely	42	8.5
Carer	339	68.2
Hours cared		
1-19	145	45.2
20-49	110	34.3
≥50	66	20.6
Social care	287	57.8
Having no health condition	92	18.5
Long-term condition	317	63.8
Asthma	68	13.7
Arthritis	36	7.2
Heart conditions	62	12.5
Stroke	6	1.2
Under- or overactive thyroid	46	9.2
Bronchitis/emphysema	18	3.6
Liver disease	14	2.8
Cancer	95	19.1

continued on next page

Table 1. Continued

Variable	n	%
Diabetes	145	29.2
Epilepsy	17	3.4
High blood pressure	63	12.7
Irritable bowel syndrome	41	8.3
Depression	83	16.7
Other physical health	30	6.0
Other mental health	4	0.8
Any physical health	283	56.9
Any mental health	162	32.6
EQ-5D-3L, mean (SD)[†]	0.573	0.310
EQ-5D-5L, mean (SD)[‡]	0.879	0.144
SWEMWBS, mean (SD)	24.2	6.0

CABA indicates City of Buenos Aires; EQ-5D-3L, 3-level version of EQ-5D; EQ-5D-5L, 5-level version of EQ-5D; IQR, interquartile range; SWEMWBS, Short Warwick-Edinburgh Mental Wellbeing Scale; WHOQOL-BREF, World Health Organization Quality of Life - Brief instrument.
*Range 24 to 91.
[†]Using Uruguay EQ-5D-5L value set.
[‡]Using Argentina EQ-5D-3L value set (the only available).

Sample

For both factor and item response theory (IRT), the international collaboration sample size for Argentina was aimed at 500 participants.[15] To include different groups of interest, we aimed that 100 of the subjects should not have long-standing health conditions, whereas the remaining 400 should have any of the following conditions: cancer ($n \geq 50$), depression or anxiety ($n \geq 50$), heart conditions ($n \geq 50$), decreased mobility ($n \geq 50$), sensory disturbances ($n \geq 50$), asthma or chronic obstructive pulmonary disease (n = 50-100), diabetes (n = 50-100), hypertension or obesity (n = 50-100), and any other health condition that is not identified in this category ($n \sim 200$-300). Conditions were selected to represent long-term physical and health conditions that were considered relevant in the local context and that allowed some overlap with the international teams. Ideally, it was sought that at least 100 of the people previously identified be users of social assistance and at least 100 be caregivers. All these conditions were self-reported. All participants were older than 18 years of age and able to read and complete the questionnaire.

An online survey was conducted for psychometric testing of items that qualified for the psychometric stage, the EQ-5D-5L and the EQ-5D-3L, the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS), financial management, health coverage, and out-of-pocket spending healthcare costs. EQ-5D-5L and EQ-5D-3L are both generic 5 item instruments (the EQ-5D-5L has 5-level response options and EQ-5D-3L 3-level responses) and a VAS item commonly used to measure patient-reported QOL. The EQ-5D-3L was scored with Argentina's social values and EQ-5D-5L was scored with Uruguay's social values given that there are no published EQ-5D-5L local value sets.[16,17] Both EQ-5D versions were analyzed with the VAS and the EQ-5D index. VAS ranges from 0 to 100; the higher the score, the higher the QOL. For EQ-5D, index 0 represents death, 1 represents perfect health status, and values below 0 represent states worse than death. The SWEMWBS is a measure with 7 items covering positive mental wellbeing that was scored by summing across items or using Rasch scores. We use the summative score method for our analysis, which is scored from 7 to 35, where higher scores represent higher mental wellbeing.[18] Financial management was evaluated with the item used in the QOL scale, the World Health Organization Quality of Life - Brief developed by the World Health Organization that has been adapted to the Spanish language.[19] The health coverage module was assessed by the module used by the National Survey of Health in Argentina and out-of-pocket spending by the questions used in the National Household Survey. A total of 6 versions of the questionnaire were used where EQ-HWB items were randomized in groups based on response options. The EQ-5D-3L and EQ-5D-5L were alternated and appeared either in the middle of the EQ-HWB items or at the end after all the other questionnaires.

Continuous variables were expressed as mean and SD or interquartile range (IQR) according to its distribution. Discrete variables were expressed as a percent and 95% confidence interval (CI). Psychometric assessment explored distribution of responses including ceiling and floor effects and known-group differences—that compared specific subgroups with healthy comparators identified based on the respondent reporting no health conditions, no problems in the EQ-5D-5L dimensions, and having an EQ-VAS score > 0.8—based on effect sizes, using Cohen's D cutoffs for different categories. Dimensionality was explored using a combination of exploratory factor analysis and confirmatory factor analysis (CFA) accounting for the categorical nature of variables and potential correlation between factors (we use the CFA model of the UK for our analysis). Poor fitting items were identified using information functions and disordered item responses using category characteristic curves from IRT models. Differential item functioning was also tested across key groups in the IRT analysis. In addition, item correlation within subdomains and with items in other subdomains was assessed to identify where items did not fit based on rho < 0.5 and rho > 0.7, respectively, based on expectations of strength of correlations in related and unrelated items.

For EQ-5D-5L and EQ-5D-3L comparison, descriptive analyses were undertaken, additionally using the so-called misery index (the sum of the states—from 5 to 15 in the EQ-5D-3L and from 5 to 25 for the EQ-5D-5L), where 5 is having the best status in all 5 domains; and 15 or 25 is being in the "pits" state, with all dimensions in the worst status. The association between EQ-HWB instrument summative scores—estimated by adding the individual Likert-type scores (1-5) of the 25 items—and EQ-VAS scores

Table 2. Narrative of the development of the items from the face validity and psychometric stages in Argentina throughout the EQ-HWB final questionnaire.

Subdomain	Narrative of the findings of the FV stage and its contribution for the development of EQ-HWB
Domain: activity	
Meaningful/enjoyable	Six items included in FV. Three of those items were included in psychometric stage. Similar wording to 1 item in EQ-HWB.
Daily activities	Three items included in FV: • Answer choices were confusing for 1 item. • Need to include examples of daily activities • D-scale for response options is recommended. Two items were included in psychometric stage; one of them with wording (examples of daily activities) and response options editions (D-scale). One item was selected for EQ-HWB.
Self-care	Six items included in FV: • Three items had wording ambiguities. • One item was considered very problematic. • Personal needs are understood as many different needs, not only referred to hygiene or self-care, it is suggested to mention activities related to the domain in a straightforward manner in the statement. Three items included in psychometric stage; all with some wording and the response options were edited in one item. Recommendations from FV stage are followed in EQ-HWB.
Mobility	Two items included in FV, one related to mobility inside the home and the other, to outside the home: • D-scale for response options is recommended. Same items but with suggested wording editions and D-scale were included in psychometric stage. EQ-HWB merged inside and outside mobility.
Communication	Three items regarding difficulties in communication were included in FV: • Ambiguities in the wording • Response options do not answer the question. • Some problems with the meaning of "communication" (could include physical, emotional, and contextual issues) One item related to hearing and speech was included in psychometric stage. No questions included in the EQ-HWB.
Hearing	Two questions included in FV: • D-scale for response options is recommended. • It is suggested to switch the items order by placing the item referring to vision before the one referring to hearing as need of lenses is more frequent than hearing aid. Psychometric included one of the items with some wording editions and D-scale responses. EQ-HWB includes the items switching the order between seeing and hearing.
Seeing	Two questions included in FV: • It is recommended to use D-scale. One item with some wording editions and D-scale responses included in psychometric. EQ-HWB includes the items switching the order between seeing and hearing.
Domain: relationships	
Support	Four items were included in the FV stage, all about positive feelings regarding "to be supported": • wording ambiguities about the meaning of "to be support" and the translation of that expression. Psychometric stage included one item used at the FV and added a second one referring to a negative feeling (lack of support). EQ-HWB includes the item included in the psychometric stage.
Positive relationships	Three items were included in the FV stage, 2 of them were about positive feelings and one about negative feelings: • wording ambiguities in one item. Psychometric stage included one item. EQ-HWB does not include any item from this subdomain.
Lonely	Total of 4 items were included in the FV stage under "lonely" and "lonely and belonging" subdomains. Psychometric stage included 3 items. EQ-HWB includes one of the items tested in previous stages.

continued on next page

ARTICLE IN PRESS

Table 2. Continued

Subdomain	Narrative of the findings of the FV stage and its contribution for the development of EQ-HWB
Stigma	Total of 5 items included in the FV stage under "stigma" and "stigma/belonging" subdomains: • One item was considered very problematic. Psychometric stage included the 4 unproblematic items from the FV. EQ-HWB includes one of the items that performed well in previous stages.
Domain: cognition	
Concentration	Four items were included in the FV stage: • One was problematic and response options were considered inappropriate. The 3 no problematic items were included in psychometric, all with wording editions. EQ-HWB includes 2 items.
Memory	Two items were included in the FV stage: • For response options even S or F scales were considered appropriate. One item included in psychometric as well as in EQ-HWB without editions.
Confusion	One item was included in FV: • Some problems with the understanding; to be confused have may have different meanings in Spanish; suggestion to use examples. Same item was included in psychometric stage but not in EQ-HWB.
Domain: self-identity	
Confidence	FV stage included a total of 10 items organized under the subdomains confidence, respect and dignity, and self-worth: • Two items were considered problematic and having ambiguities. Psychometric stage included 2 items under the confidence subdomains and also 2 items under the self-worth subdomain. EQ-HWB includes one of the items tested in previous stages under the subdomain self-worth.
Domain: autonomy	
Cope	Four items were included in the FV stage: • The expression "To cope" is very difficult to be translated into Spanish. Some respondents proposed to use "hacer frente" rather to "arreglármelas." To cope may include also day-to-day situations and not only problems. • Two items were considered very problematic. Three items included in psychometric; all with wording editions. EQ-HWB includes one item with editions from the FV stage.
Control	Three item were included in the FV stage: • Two were partially understood but suggested to include a clarifications about the meaning of controlling daily life. • One item was considered very problematic. Three items included in psychometric; all with wording editions. EQ-HWB includes one item with the addition of more explanation.
Domain: feelings and emotions	
Happiness	Six items were included in the FV stage including negative and positive feelings: • Two were difficult to understand or ambiguous. • The term "unhappy" has different connotation in Argentinean culture and is used as an insult. • Response options were considered inappropriate in one item. For the psychometric stage 3 of the previous items were included and also a forth item was included following findings in FV (differentiating Happy and cheerful). EQ-HWB includes one items that was well understood during the FV.
Hope	Four items were included in the FV stage including negative and positive feelings: • One was not accepted. • One had some wording problems that made it difficult to understand. Psychometric stage included a total of 4 items: 2 without editions; one had mayor editions; a forth item was included following findings in FV. EQ-HWB includes one items that was included in the FV stage with no editions.
Safety	Five items were included in the FV stage including negative and positive feelings: • Four items had some problems regarding the meaning. • Literal translation of "unsafe" can be perceived as a lack of self-confidence, so it is recommended to add more detail. Psychometric stage included 4 items. EQ-HWB includes one items that was considered problematic in the FV but with further clarification.

continued on next page

Table 2. Continued

Subdomain	Narrative of the findings of the FV stage and its contribution for the development of EQ-HWB
Anxiety	Five items were included in the FV stage including negative and positive feelings: • One item was considered difficult to be understood. Psychometric stage included 3 items without editions. EQ-HWB includes one item despite it had some difficulties during the FV stage.
Anger	Five items were included in the FV stage: • Two items were considered difficult to be understood. EQ-HWB includes one item that was well understood in the FV.
Domain: physical sensations	
Pain	Two items were included in the FV stage: • some wording recommendation for one item. Psychometric stage included both items without editions. EQ-HWB includes same items.
Discomfort	Two items were included in the FV stage: • respondents remarked the importance to include examples as the items are including. Psychometric stage included both items without editions. EQ-HWB includes same items.
Energy	Three items were included in the FV stage. Psychometric stage included 2 items without editions. EQ-HWB includes one of the items.
Sleep	One item was included in the FV stage. Psychometric stage and EQ-HWB included same item without editions.

Note. Because most of the items were considered unproblematic (well understood and accepted) during the FV stage, we detail only the problematic items and recommendations that arose from the interviews.
See final items in Table 3 (and Supplemental Material Spanish versions).
EQ-HWB indicates EQ Health and Wellbeing; FV, face validity.

was explored. For this, of the 27 items that finally were 25 in the final measure (as 2 pairs of items from the same subdomain were combined), the 2 pairs of corresponding items were collapsed in their scoring to establish the 25 items of the final measure. In both cases, the joint item score was considered as 1 or 5 when both items coincided in this value; for the intermediate values, the highest value of any of the 2 corresponding items was used. Both additive scoring and a score transformed from 0 to 100 (score 100) were incorporated. Psychometric analysis was conducted using STATA, Mplus, and R. Additional details of the psychometric analysis are described in Peasgood et al[14] in this issue.

Ethical Aspects

The qualitative study protocol was approved by the Research Ethics Committee (Comité de Ética de Protocolos de Investigación) of the Hospital Italiano, Buenos Aires, Argentina (protocol number 3693), and the quantitative study protocol was approved by the Research Ethics Committee (Comité de Ética en Investigación) of the Centro de Educación Médica e Investigaciones Clínicas "Norberto Quirno," Buenos Aires, Argentina (Protocol number 1221). All participants provided an informed consent in both studies.

Results

Face Validity

We conducted interviews with 15 women and 9 men including 8 carers, 8 social services users, and 8 persons from the general public. Median age was 55.5 years (IQR 37.5-70.5). Level of education of participants was as follows: 6 had incomplete high school or less, 8 had finished secondary school, and 10 had some higher education (Table 1).

Each of the initial 97 candidate items grouped in domains was assessed by 12 to 13 participants. Respondents considered that all items were relevant; most of the items were well understood and accepted. Participants did not consider items embarrassing or stressful, and no missed out items that should be included were identified. Some of the candidate items had ambiguities in the wording. For example, in items related to personal needs, respondents found them open to >1 interpretation as "needs" was understood in many different aspects, not only to hygiene or self-care. For that reason, we suggested that the items should be more specific using examples or asking directly about difficulties to "wash, toilet, get dressed, eat, or care for your appearance." Other candidate items were considered confusing. For example, to get around outside the house with no difficulty was considered unclear and could be interpreted as problems with streets infrastructure, such as broken sidewalks or dog mess. For this item, we suggested a synonym for the Spanish translation more easily to be identified with physical capability (replace "trasladar" by "desplazar").

Based on participants' opinions, some wording improvements were proposed. Table 2 details the findings of the cognitive interviews and describes the pathway of candidate items development. (For more detail of this initial 97 item set and some abridged example of the online survey, see Appendix 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.12.010.) Based on the international face validity findings, 64 candidate items were taken forward for psychometric testing including 14 items that were modified and 3 new items.[13]

Psychometric Study Component

A total of 497 participants were recruited online in Argentina from August to October 2019. Most of them were from the city and the province of Buenos Aires (54.5%). The mean age was 37.4 years

ARTICLE IN PRESS

Table 3. Item-specific summary of psychometric results in Argentina (N = 497).

Subdomains	Item	Distribution		Known-group differences								Correlation		IRT	
		Ceiling	Floor	PH	MH	Dia	Can	Dep	C	HC	SC	< 0.5	> 0.7	Ord	DIF
Activity															
Meaningful/ enjoyable	I could do the things I wanted to do.	✓	✗	○	○	○	✓	✓	✗	✗	✗			✓	
Daily activities	How well were you able to do your day-to-day activities? (D)	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗		✓	
Self-care	How difficult is it for you to look after yourself? (D)	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	✗		✗	✗
Mobility	How well were you able to get around inside your home? (D)	○	✗	✓	✓	✓	✓	✓	✓	✗	✓		✗	✓	✗
	How well were you able to get around outside your home? (D)		✗	✓	✓	✓	✓	✓	✓	✗	✓				
Hearing	How well can you hear (using hearing aids if you normally wear them)? (D)	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓			-	-
Seeing	How well can you see (using your glasses or contact lenses if they are needed)? (D)	✓	✗	○	○	○	✓	○	✗	✗	✓			-	-
Relationship															
Support	I felt unsupported by other people.	✓	✓	✓	✓	○	✓	✓	✗	✓	✗	✗		✓	
Lonely	I felt lonely.	✓	✓	○	✓	○	○	✓	✗	✗	✗	✗	✗	✓	
Stigma	I felt accepted by others.	✓	✗	○	○	○	○	✓	✓	✗	✗	✗		✓	○
Cognition															
Concentrate	I found it hard to concentrate.	✓	✓	✓	✓	○	○	✓	✓	✗	✓	✗		✓	
	I had trouble thinking clearly.	✓	✗	✓	✓	○	✓	✓	✗	✗	✓			✓	
Memory	I had trouble remembering.	✓	✗	✓	✓	○	✓	✓	✗	✗	✗	✗		✓	
Self-identity															
Self-worth	I felt good about myself.	✓	✗	✓	✓	○	○	✓	✗	✗	✗	✗		✓	
Autonomy															
Cope	I felt unable to cope with my day-to-day life.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗		✓	
Control	I felt I had no control over my day-to-day life.	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗		✓	
Feelings															
Happiness	I felt sad	✓	✓	✓	✓	○	○	✓	✗	✗	✓	✗		✓	○
Hope	I felt that I had nothing to look forward to.	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗		✓	
Safety	I felt unsafe.	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓	✗		✓	
Anxiety	I felt anxious.	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗		✓	○
Anger	I felt frustrated.	✓	✓	✓	✓	○	✓	✓	✓	✗	✓			✓	
Physical															
Pain	I had physical pain. (S2)	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓		✗	✓	
	How often did you experience pain?	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓			✓	
Discomfort	I had physical discomfort. (S2)	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	-	-
	I had physical discomfort.	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗		-	-

continued on next page

Table 3. Continued

Subdomains	Item	Distribution		Known-group differences								Correlation		IRT	
		Ceiling	Floor	PH	MH	Dia	Can	Dep	C	HC	SC	< 0.5	> 0.7	Ord	DIF
Energy	I felt exhausted.	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✗		-	-
Sleep	I had problems with my sleep.	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗		-	-

Note. EQ-HWB content was reproduced by permission of EuroQol Research Foundation. Distribution < 5%; Known-group: PH; MH; Dia; Can; Dep; C; HC; SC; IRT Ord; DIF. Response options are frequency (F) for most of the items, none of the time, only occasionally, some of the time, often, most or all of the time; D, no difficulty, slight, some, a lot of difficulty, unable; severity S1, not at all, a little bit, somewhat, quite a bit, very much; S2, none, mild, moderate, severe, very severe.
Ceiling floor effects: ✗: ≥ 70% or < 5%; ○ : ≥ 50% or < 70%; ✓: ≥ 5% or < 50%. Known-group: ✓: (Cohen's D ≥ 0.5 in PH/MH; ≥ 0.2 small to moderate in C/HC), ○ (Cohen's D ≥ 0.2 to < 0.5 in PH/MH, mixed small (≥ 0.2) to large (≥ 0.8) in PH/MH), (Cohen's D < 0.2 or not statistically significant or wrong direction). Correlation: ✗: < 0.5 small or no correlation within subdomain or ≥ 0.7 large correlation with items in other subdomains. Order: ✗: Not ordered; ✓: Responses ordered. DIF: ✗: DIF present; ✓: No DIF.
C indicates carers; Can, cancer; Dep, depression; Dia, diabetes; DIF, differential item functioning; EQ-HWB, EQ Health and Wellbeing; HC, hours cared; IRT Ord, item response theory ordered responses; MH, mental health conditions; PH, physical health conditions; SC, social care users.

and the majority were men (59.6%). Approximately 64% reported that they had a long-term condition, and most reported a physical health condition (57%). Mean EQ-5D-3L and EQ-5D-5L scores were 0.57 and 0.88, respectively; mean SWEMWBS was 24.2. The majority were identified as carers (68%) and 57% benefited by social care. The online survey took a mean of 14.8 minutes (SD 7.3) to complete (see other details in Table 1).

Psychometric Performance

A summary of the results of the psychometric analysis, with all the details analyzed in Argentina of the 27-item subset that were finally included in the final long instrument, is shown in Table 3, including floor and ceiling effects, known-group validity, correlation analysis, and IRT analysis. From the 27, 2 pairs were finally merged as conceptually similar in the final 25-item EQ-HWB measure. The results of the remaining 37 items are in the Appendix 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.12.010. Results are summarized using a 3-level scale assigned by the local Argentinean research team: no problem, mixed evidence, and exhibited problems.[14]

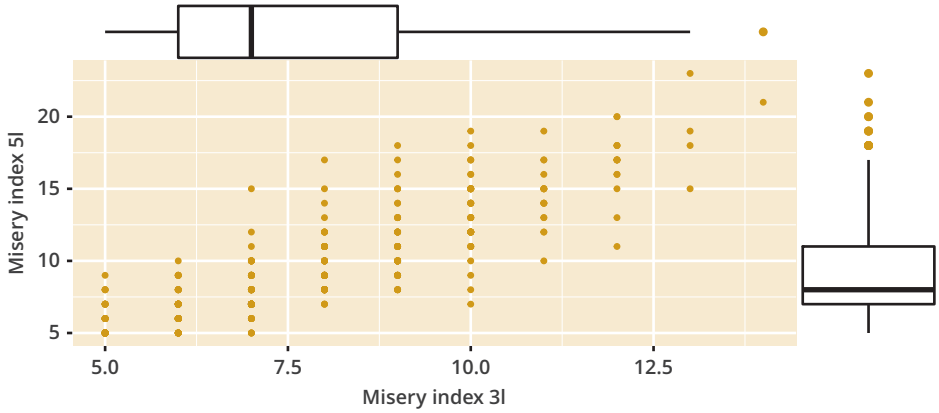
A total of 14 of 27 items showed relevant floor effect and only 1 of the 27 showed ceiling effect. The activity domain showed floor effect in all its items. The domain cognition showed floor effect in 2 of 3 items. The others domains showed low levels of floor effects.

Regarding known-group validity, all of the items could detect differences in physical and mental conditions in general and in all specific health conditions, with a moderate to large effect size. For carers, items were less able to discriminate based on providing care alone and performed worse when hours of care were taken into account. A total of 10 of the items could not detect differences for social care users.

Most items were correlated as expected with items in their domain and were not correlated strongly with items in other domains/subdomains. The IRT analysis indicated only 1 disordered item ("I was able to look after myself") and 2 items where differential item functioning was observed ("How well were you able to get around inside your home" and "How difficult is it for you to look after yourself [D]") both in the activity domain. For the CFA analysis, we used the CFA for UK data and model fit statistics for our data. Our Argentine analysis showed a good root mean square error of 0.049 (95% CI 0.047-0.051). The comparative fit index and the Tucker-Lewis index were also both good (0.956 and 0.952, respectively) when mobility was merged with self-care, energy was removed, and self-worth was merged with coping. Aggregated data of the psychometric analyses with summary data from the 6 participating countries can be seen in Peasgood et al.[14]

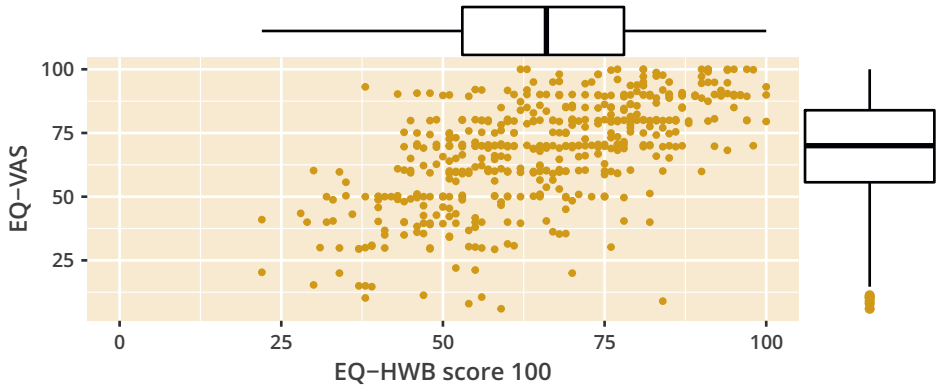
Based on the overall country recommendations and consultation with stakeholders, a 25-item version of the EQ-HWB covers activity, relationships, cognition, self-identity, autonomy, feelings,

Figure 1. Plot between misery index EQ-5D-5L and EQ-5D-3L.



EQ-5D-3L indicates 3-level version of EQ-5D; EQ-5D-5L, 5-level version of EQ-5D.

Figure 2. Plot between EQ-VAS and EQ-HWB score 100.



EQ-HWB indicates EQ Health and Wellbeing; VAS, visual analog scale.

and physical sensations with a short version EQ-HWB-S that was amenable for valuation covering activity (mobility, daily activities), relationships (loneliness), cognition (concentration/thinking clearly), feelings (sadness/depression, anxiety), and physical sensations (pain, exhaustion).[11] Argentina recommended or strongly recommended 23 of the final 25 items in the EQ-HWB. These versions of the EQ-HWB are experimental versions as defined in the Intellectual Property Policy of the EuroQol Research Foundation.

Comparison of Measures

The mean and median of the EQ-VAS were 68 (SD 20.42) and 70 (IQR 55.65-83.92), respectively. The proportion of [1,1,1,1,1] was 14.89% and 13.08% for the EQ-5D-3L and EQ-5D-5L, respectively. The proportion with at least 1 severe limitation (4 or 5) value in the EQ-5D-5L was 22.98%, and the proportion with at least 1 severe limitation[3] in the EQ-5D-3L was 22.33%. Misery index scale values were 7 (IQR 6-9) for the EQ-5D-3L and 8 (IQR 7-11) for the EQ-5D-5L with good correlation between them (tau = 0.76, $P < .05$). Figure 1 shows the correlation between the 2 misery index values, and Table 4 shows the response distribution between the 2 scales. Both the EQ-VAS and the EQ-HWB score assess QOL, so a good correlation between the 2 was expected; nevertheless, the correlation between them was moderate (tau = 0.44, $P < .05$) (Fig. 2).

Discussion

We present in our study the details of the Argentina face validity and psychometric analysis that led to the design of a new set of instruments within the EQ-HWB international collaboration project. It was the only Spanish-speaking country that participated and the only one from Latin America. Our study focuses on the "Argentina case" as a single country study and can be better read in the context of other articles in this issue that describe the overall results of the international project that led to the new EQ-HWB instrument.[11,13,14] Argentina contributed with 24 of a total of 168 face validity interviews and with of 497 of 4830 total of psychometric surveys. Partial findings and recommendations were shared and discussed iteratively with the collaborative study coordination and with the other study sites.

The face validity in Argentina found that most of the items were accepted by respondents. Findings provided suggestions for wording in the items and about the response options preferences. Most of the recommendations resulted from the Argentinean interviews were included in the final questionnaires the EQ-HWB. Most of these items showed relatively good psychometric characteristics in the Argentine population.

Our psychometric results were similar to that of other EQ-HWB countries published to date and published in a separate article in this issue.[14,20] Like in the UK and Germany, only 1 item—"How well can you hear (using hearing aids if you normally wear them)? (D)"—showed a ceiling effect, which is not unexpected in mixed populations. In our study, the floor effect was concentrated on the "activity" domain. The UK and Germany also showed floor effects in most activity items, reflecting that for most people the measure would have problems in this domain. For floor effect in the other domains/subdomains, Argentina ranked between Germany and the UK, where Germany showed in general more floor effect in the items with respect to the UK. Some of these results could be explained by the relatively young age of the Argentinean sample. Furthermore, although there was some overlap in conditions that were assessed in different countries, there were differences, for example, in the number of conditions or other groups such as informal carers that were included that could also be driving the differences.

Like in the UK and Germany analyses, almost all items could detect differences in mental and physical conditions. In carers and hours of care analyses, Argentina behaved like Germany, where almost half of the items were able to detect differences for carers and performed worse when hours of care were included. Nevertheless, the UK analysis showed improvement when hours of care were taken into account.

The study in Argentina compared with the UK and Germany showed almost half of the items with large correlation with the items in other domains/subdomains.[14] In contrast, 18 of 27 items showed poor correlation within domains/subdomains in Argentina, where only 2 in the UK and 3 in Germany showed this characteristic. For the IRT analyses, in Argentina only 2 of 27 items showed poor DIF whereas in the UK a third of the items showed this. In the case of item ordering, assessed using item information functioning curves, the analyses conducted in Argentina and in the UK had similar results. For the CFA analysis and similarly to Germany (with mobility factor merged with self-care and energy factor removed) and Australia, the UK model had good fit for our data.

As expected and having both shared and EQ-HWB additional domains, the EQ-HWB showed moderate correlation to the EQ-5D-3L and EQ-5D-5L. Regarding the EQ-5D-5L and EQ-5D-3L, Buchholz et al[21] showed a proportion of ceiling effect ([1,1,1,1,1]) in patient population of 23% (CI 17%-30%) for

Table 4. EQ-5D-3L versus EQ-5D-5L; summary of responses.

Dimension	3L	n (%)	5L	n (%)
Mobility	1	336 (67.6)	1	292 (86.9)
			2	41 (12.2)
			3	2 (0.6)
			4	1 (0.3)
			5	0 (0)
	2	153 (30.8)	1	10 (6.5)
			2	86 (56.2)
			3	48 (31.4)
			4	7 (4.6)
			5	2 (1.3)
	3	8 (1.6)	1	0 (0)
			2	0 (0)
			3	1 (12.5)
			4	4 (50.0)
			5	3 (37.5)
Self-care	1	337 (75.9)	1	348 (92.3)
			2	22 (5.8)
			3	7 (1.9)
			4	0 (0)
			5	0 (0)
	2	109 (21.9)	1	10 (9.2)
			2	57 (52.3)
			3	37 (33.9)
			4	5 (4.6)
			5	0 (0)
	3	11 (2.2)	1	0 (0)
			2	2 (18.2)
			3	3 (27.3)
			4	3 (27.3)
			5	3 (27.3)
Usual activities	1	264 (53.1)	1	211 (79.9)
			2	45 (17)
			3	7 (2.7)
			4	1 (0.4)
			5	0 (0)
	2	219 (44.1)	1	30 (13.7)
			2	122 (55.7)
			3	52 (23.7)
			4	15 (6.8)
			5	0 (0)
	3	14 (2.8)	1	0 (0)
			2	0 (0)
			3	4 (28.6)
			4	6 (42.9)
			5	4 (28.6)
Pain	1	169 (34)	1	109 (64.5)
			2	51 (30.2)

continued on next page

Table 4. Continued

Dimension	3L	n (%)	5L	n (%)
			3	7 (4.1)
			4	2 (1.2)
			5	0 (0)
	2	292 (58.8)	1	23 (7.9)
			2	167 (57.2)
			3	85 (29.1)
			4	13 (4.5)
			5	4 (1.4)
			1	0 (0)
	3	36 (7.2)	2	6 (16.7)
			3	9 (25.0)
			4	19 (52.8)
			5	2 (5.6)
Anxiety depression	1	180 (36.2)	1	126 (70)
			2	46 (25.6)
			3	7 (3.9)
			4	1 (0.6)
			5	0 (0)
	2	237 (47.7)	1	21 (8.9)
			2	121 (51.1)
			3	73 (30.8)
			4	18 (7.6)
			5	4 (1.7)
	3	80 (16.1)	1	0 (0)
			2	6 (7.5)
			3	11 (13.8)
			4	47 (58.8)
			5	16 (20.0)
EQ-5D-3L and EQ-5D-5L distribution				
EQ-5D-3L [1,1,1,1,1]			74	14.89%
EQ-5D-3L with at least one 3 value			111	22.33%
Misery index EQ-5D-3L (median and IQR)			7	(6-9)
EQ-5D-5L [1,1,1,1,1]			65	13.08%
EQ-5D-5L with at least one 4 or 5 value			114	22.94%
Misery index EQ-5D-5L (median and IQR)			8	(7-11)
EQ-5D VAS	(median and IQR)		70	(55.65-83.92)
	(mean and SD)		68	20.42
EQ-HWB (additive scoring) (mean and SD)			90.73	16.38
EQ-HWB (0-100 scoring) (mean and SD)			65.73	16.38

EQ-5D-3L indicates 3-level version of EQ-5D; EQ-5D-5L, 5-level version of EQ-5D; EQ-HWB, EQ Health and Wellbeing; IQR, interquartile range; VAS, visual analog scale.

EQ-5D-3L and 18% (CI 13%-24%) for EQ-5D-5L. Although our study showed a lower ceiling effect in the EQ-5D-3L (out of the CI), in the EQ-5D-5L it matches the lowest CI. Another finding is the relevant difference of values when using the EQ-5D-5L value set from Uruguay[17] (the nearest and most similar country with a EQ-5D-5L value set) or the EQ-5D-3L values from Argentina,[16] although this difference is mostly attributable to the difference between the value sets.[22]

Strengths

In this article, we show in detail both a face validity (qualitative) analysis and a quantitative analysis describing the psychometric properties in Argentina of this newly developed instrument. These stages were part of an international collaborative protocol that assures consistency in the conduction and analyses. We were the only Spanish-speaking country and culture of

the international group (that also included Australia, China, Germany, UK, and US), which can help to promote the validity in our country and probably other Spanish-speaking countries or cultures (ie, *latinos*). This multicultural and multilingual methodology of instrument development—as opposed to being a single country and language endeavor—can help to show from the beginning of the project that the measure will be more applicable and that would support wider use of such measure.[23] From the 25 final items of the long measure, the Argentina team recommended or strongly recommended 23 of them.

Limitations

Our study was not powered to detect specific differences among particular study groups. Additionally, because of the design of both the qualitative and the quantitative stages, these stages cannot guarantee that the results are widely valid for the Argentine population. For example, the psychometric stage included a significantly younger sample than the face validity stage. Online data collection focusing on those who are able to self-complete rather than interview-administered may also have had an impact on the results. Nevertheless, Argentine results were not used in isolation to make decisions about the overall measure and more older participants were included in the other countries. In addition, known-group difference analysis applied in the Argentine component that compared healthy people with people with some health condition showed good discrimination in several groups, although it did not assess its discriminant ability within specific disease groups. Additionally, because the study had a cross-sectional design, changes over time could not be assessed. This initial study undertook by an international collaboration in several languages had only Argentina as a Spanish-speaking country and culture. Future studies could explore the cultural and linguistic adaptation of the instruments to other Spanish-speaking contexts.

Conclusions

This study shows the detailed contributions to face validity and psychometrics of 1 site for the development of a new wellbeing generic measure. Collaborative studies require local capacity and effective and responsive coordination. The assessment of candidate items in the only Spanish-speaking country of the collaboration informed the design of an internationally relevant EQ-HWB 25-item measure and a short EQ-HWB-S 9-item measure intended to be used in economic evaluations.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.12.010.

Article and Author Information

Accepted for Publication: December 19, 2021

Published Online: XXX

doi: https://doi.org/10.1016/j.jval.2021.12.010

Author Affiliations: Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina (Augustovski, Argento, Rodríguez, Gibbons, Belizán); School of Health and Related Research, University of Sheffield, Sheffield, England, UK (Mukuria).

Correspondence: Federico Augustovski, MD, MSc, PhD, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Ravignani 2024, Ciudad de Buenos Aires (1414), Argentina. Email: faugustovski@iecs.org.ar

Author Contributions: *Concept and design:* Gibbons, Mukuria, Belizán
Acquisition of data: Rodríguez, Mukuria
Analysis and interpretation of data: Augustovski, Argento, Rodríguez, Gibbons, Mukuria, Belizán
Drafting of the manuscript: Augustovski, Argento, Rodríguez, Gibbons, Belizán
Critical revision of the paper for important intellectual content: Augustovski, Gibbons, Mukuria, Belizán
Statistical analysis: Augustovski, Argento, Gibbons
Administrative, technical, or logistic support: Rodríguez
Supervision: Augustovski, Gibbons, Mukuria, Belizán

Conflict of Interest Disclosures: All authors reported their institutions receiving grants from the UK Medical Research Council (grant number 170620) and the EuroQol Research Foundations. No other conflicts were reported. The views expressed in this article are those of the authors and not necessarily of our funders, the National Institute for Health and Care Excellence, the Department of Health and Social Care, or those acknowledged.

Funding/Support: This study is independent research funded by the UK Medical Research Council (grant number 170620) and the EuroQol Research Foundation.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgment: The authors thank the National Institute for Health and Care Excellence for highlighting the methodological research need to the Medical Research Council that resulted in the funding call entitled "Beyond the QALY," which led to this research being funded. The authors acknowledge the support of the National Institute for Health Research Yorkshire and Humber Applied Research Collaboration (formerly CLAHRC) and the National Institute for Health Research Clinical Research Network. The authors acknowledge the invaluable contributions of members of the project steering group, advisory group, and public and patient involvement and engagement groups and Julie Johnson for project administration. The authors also thank members of the EuroQol Group Association for their input at plenary and academy meetings and the EuroQol office for their support. We acknowledge the contribution of all the patients, social care users, and informal carers who took part in all the studies across the different countries. Finally, we would also like to extend our acknowledgments to all the members of the international estimated quality-adjusted life-year consortium: John Brazier, Tessa Peasgood, Jill Carlton, Janice Connell, Stacey Rand, Nancy Devlin, Karen Jones, Aki Tsuchiya, Rosemary Lovett, Bhash Naidoo, Donna Rowen, Juan Carlos Rejon-Parrilla (United Kingdom) Ole Marten, Simone Kreimeier, Kristina Ludwig, Wolfgang Greiner (Germany) Brendan Mulhern, Lidia Engel (Australia), Nan Luo, and Zhihao Yang (China).

REFERENCES

1. Romero M, Vivas-Consuelo D, Alvis-Guzman N. Is Health Related Quality of Life (HRQoL) a valid indicator for health systems evaluation? *Springerplus*. 2013;2(1):664.
2. Wouters OJ, Naci H, Samani NJ. QALYs in cost-effectiveness analysis: an overview for cardiologists. *Heart*. 2015;101(23):1868–1873.
3. Wichmann AB, Adang EM, Stalmeier PF, et al. The use of Quality-Adjusted Life Years in cost-effectiveness analyses in palliative care: mapping the debate through an integrative review. *Palliat Med*. 2017;31(4):306–322.
4. Finch AP, Brazier JE, Mukuria C, Bjorner JB. An exploratory study on using principal-component analysis and confirmatory factor analysis to identify bolt-on dimensions: the EQ-5D case study. *Value Health*. 2017;20(10):1362–1375.
5. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18(9):1–224.
6. Netten A, Burge P, Malley J, et al. Outcomes of social care for adults: developing a preference-weighted measure. *Health Technol Assess*. 2012;16(16):1–166.

ARTICLE IN PRESS

7. Brouwer WB, van Exel NJ, van Gorp B, Redekop WK. The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. *Qual Life Res*. 2006;15(6):1005–1021.
8. Dixon S, Walker M, Salek S. Incorporating carer effects into economic evaluation. *Pharmacoeconomics*. 2006;24(1):43–53.
9. Techniques to include carers' quality of life in economic evaluation. University of Birmingham. https://www.birmingham.ac.uk/research/applied-health/research/care-econ/index.aspx. Accessed February 2, 2022.
10. Extending the QALY project homepage. https://scharr.dept.shef.ac.uk/e-qaly/. Accessed February 2, 2022.
11. Brazier J, Peasgood T, Mukuria C, et al. The EQ-HWB: overview of the development of a measure of health and well-being and key results. *Value Health*. In press.
12. Mukuria C, Connell J, Carlton J, et al. A qualitative review on domains of quality of life important for patients, social care users and informal carers to inform the development of the EQ-HWB. *Value Health*. In press.
13. 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 Health and Wellbeing (EQ-HWB). *Value Health*. In press.
14. Peasgood T, Mukuria C, Brazier J, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ Health and Wellbeing (EQ-HWB). *Value Health*. In press.
15. Comfrey AL, Lee HB. *A First Course in Factor Analysis*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.
16. Augustovski FA, Irazola VE, Velazquez AP, Gibbons L, Craig BM. Argentine valuation of the EQ-5D health states. *Value Health*. 2009;12(4):587–596.
17. Augustovski F, Rey-Ares L, Irazola V, et al. An EQ-5D-5L value set based on Uruguayan population preferences: Report of the First Experience in Latin America. LB - 7eRo. *Value Health*. 2015;18(7):A810–A811.
18. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. *Health Qual Life Outcomes*. 2009;7:15.
19. Congost-Maestre N, et al. La traducción de la escala de calidad de vida de la OMS: el cuestionario WHOQOL-BREF. *J Spec Transl*. 2018;30:312–334.
20. Peasgood T, Mukuria C, Brazier J, et al. Developing a new generic classifier of quality of life: initial results from the Extending the QALY (EQALY) psychometric surveys. Paper presented at the 36th EuroQol Plenary Meeting; September 18-21, 2019; Brussels, Belgium. https://scharr.dept.shef.ac.uk/e-qaly/wp-content/uploads/sites/13/2018/11/Brazier-EQ-plenary-E-QALY-paper-2018.pdf. Accessed February 2, 2022.
21. Buchholz I, Janssen MF, Kohlmann T, Feng YS. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. *Pharmacoeconomics*. 2018;36(6):645–661.
22. Tejada RA, Gibbons L, Belizán M, Gutierrez EL, Reyes N, Augustovski FA. Comparison of EQ-5D values sets among South American countries. *Value Health Reg Issues*. 2021;26:56–65.
23. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract*. 2011;17(2):268–274.