Head-to-head comparison of the EQ-HWB-S, ReQoL, EQ-5D-5L measure in Australian general population and mental health wellbeing

Authors: Thao Thai[1*], Cathrine Mihalopoulos[1], Lidia Engel[1]

[1]Monash University Health Economics Group, School of Public Health and Preventive Medicine, Monash University, Level 4, 553 St Kilda Road, Melbourne VIC 3004, Australia

***Corresponding author: Thao Thai**

Email: thao.thai@monash.edu

Phone: +61 3 9903 0932

Address: Public Health and Preventive Medicine, Monash University, Level 4, 553 St Kilda Road, Melbourne VIC 3004, Australia

Acknowledgments

The authors would like to thank all the participants in this research.

Abstract

Objectives:

The EQ-Health and Wellbeing short version (EQ-HWB-S), is a new measure specifically developed for assessing the health and wellbeing outcomes in economic evaluations of interventions in healthcare, social care, and public health from the perspective of service users, patients and their caregivers. Given its novelty, it is imperative to gather additional evidence regarding its psychometric performance across diverse general populations. Furthermore, despite some limited evidence regarding the psychometric properties of the EQ-HWB-S in areas such as social care and among cancer survivors, its performance in mental health and wellbeing domains lacks comprehensive investigation, despite its extensive coverage of psychosocial items. This study seeks to explore the direct comparison of the psychometric performance of the EQ-HWB-S with the EQ-5D-5L and Recovering Quality of Life-Utility Index (ReQoL-UI) in both the general population and individuals with mental health problems.

Methods:

Data were used from an online survey administered to a representative sample of Australian general population aged 18 years and more. Representativeness was assessed by age, gender and geographic location. Respondents completed the EQ-HWB-S, ReQoL-UI, EQ-5D-5L and Kessler psychological distress 6 items (K6) in a random order, followed by demographic questions, questions about current and past experience with mental health problems. We compared the measures in terms of convergent validity using Spearman's rank correlation based on a validated clinical measure of psychological distress (Kessler 6 items). Correlation coefficients were interpreted as weak (≤ 0.4), moderate (0.4–0.7) and strong (> 0.7). Known-group validity was assessed between groups with and without mental health problems using self-reported mental illness and the K6 scores with a clinical threshold of 13 or higher.

Results:

The psychometric analysis was conducted using a representative sample of 2,024 Australians in terms of age, gender and geographic location. The EQ-HWB-S, ReQoL and EQ-5D-5L demonstrated convergent validity when compared against the K6. Notably, the EQ-HWB-S exhibited superior convergent validity, with all items of K6 moderately or highly correlated with the psychological items of EQ-HWB-S. Regarding the known-group validity, most EQ-HWB-S, ReQoL, EQ-5D-5L items demonstrated discriminative ability among known groups with mental health problems.

Conclusion:

Preliminary findings using a large representative sample of Australians aged 18 and above support the legitimacy of the EQ-HWB-S as a potential outcome measure for mental health and well-being. Specifically, our findings show a convergence between EQ-HWB-S and the K6 while the EQ-HWB-S performed well in differentiating among individuals with mental health problems.

Introduction

The quality-adjusted life-year (QALY) has emerged as a prominent metric in health technology assessment (HTA) and decision-making processes within healthcare systems internationally (Department of Health, 2016; National Institute for Health Care Excellence, 2013). QALYs offer a means to gauge the outcomes of interventions by integrating both the length of life and health-related quality of life (HRQoL) into a single measure. This is typically represented on a utility scale, where a score of 1 signifies perfect health and 0 represents being dead (Brazier et al., 2016). Nevertheless, while QALYs provide valuable insights into health outcomes, they have inherent limitations in fully capturing the broader spectrum of wellbeing (Brazier et al., 2022). Wellbeing extends beyond mere health, encompassing domains such as social relationships, safety, control, and autonomy. These aspects may be influenced by interventions but might not be fully captured by conventional health-related measures (Peasgood et al., 2019). However, incorporating broader measures of wellbeing can provide a more comprehensive understanding of the impacts of interventions while enabling comparisons across various interventions and sectors, which ultimately better addresses the diverse needs and preferences of individuals and communities (Brazier et al., 2022).

To enable cross-sector comparison, the EQ-Health and Wellbeing (EQ-HWB) measure was specifically developed for assessing the health and wellbeing outcomes for use in economic evaluation of interventions in healthcare, social care, and public health from the perspective of service users and their caregivers (Brazier et al., 2022). The EQ-HWB was crafted through a comprehensive process, which included reviewing qualitative studies, conducting fieldwork, and assessing its face validity and psychometric properties in six different countries (Argentina, Australia, China, Germany, the United Kingdom, and the United States) to guide the final item selection (Carlton et al., 2022; Peasgood et al., 2022). There are two available versions of the EQ-HWB: an experimental 25-item long version and a shorter 9-item version known as EQ-HWB-S for use in economic evaluations. The EQ-HWB captures the health and wellbeing outcomes across domains of 1) activity, 2) autonomy, 3) cognition, 4) feelings and emotions, 5) relationships, 6) physical sensations, and 7) self-identity. A pilot valuation set has recently been developed for the EQ-HWB-S, reflecting the UK's general population preferences towards the states described by the EQ-HWB-S (Mukuria et al., 2023a). Since EQ-HWB-S is a relatively new measure, there is a need to assess its psychometric performance as compared to one of the most frequently used generic measures, the EQ-5D-5L.

Additionally, as a health and wellbeing measure, it is critical to assess the EQ-HWB-S' psychometric performance in gauging health and wellbeing, especially within the context of mental health. Historically, EQ-5D-5L, one of the most frequently used generic health measures in economic evaluation of interventions (Roberts et al., 2019), was often used in mental health populations although its validity in serious mental illness was often questioned (Brazier, 2010; Brazier et al., 2014). Recently, the Recovering Quality of Life (ReQoL) measure was developed, a condition-specific measure that captures the quality of life outcomes for people over 16 years of age with mental health problems (Keetharuth et al., 2018). The Recovering Quality of Life- Utility Index (ReQoL-UI) classification system covers seven domains: activity, hope, belonging and relationship, self-perception, well-being, autonomy and physical health (Keetharuth et al., 2021).

There is potential overlap between the psychosocial domains covered by EQ-HWB-S and ReQoL-UI, especially in the context of mental health and wellbeing, such as aspects related to social relationships ("I felt lonely" in ReQoL and "Feel lonely?" in EQ-HWB-S) and a sense of control ("I felt unable to cope with my day-to-day life" in ReQoL versus "Feel you had no control over your day-to-day life?" in EQ-HWB-S). Therefore, it is important to investigate how these measures perform psychometrically when applied to individuals with mental health conditions.

This study aims to examine and compare the psychometric properties of EQ-HWB-S with the EQ-5D-5L and the ReQoL-UI in terms of response distributions, convergent validity, and known-group validity in a representative sample of Australian general population and people with any self-reported mental health condition and psychological distress measured by the Kessler-6 (K6). With its comprehensive coverage both health and wellbeing, EQ-HWB-S exhibits promising potential as a measure for evaluating outcomes in economic evaluation of mental health interventions. Unlike the ReQoL-UI, which is tailored specifically for individuals with mental health conditions, EQ-HWB-S stands out as a generic preference-based quality of life measure offering advantages for cross-sectors comparisons. Thus, our research not only generates valuable insights into the psychometric performance of EQ-HWB-S utilising a substantial and representative Australian sample, but also provides evidence for its suitability as an appropriate outcome measure for mental health interventions.

Methods

Measures

The EQ Health and Wellbeing short version

The EQ-HWB has been developed to capture a broad range of health and wellbeing outcomes for economic evaluation of interventions in a wider range of health and social care including informal carers. There are two version of the instruments. A longer profile measure with 25-items and the short version with 9-items (Brazier et al., 2022). The EQ-HWB-S captures the health and wellbeing outcomes across domains of 1) activity, 2) autonomy, 3) cognition, 4) feelings and emotions, 5) relationships, and 6) physical sensations. Responses to items in the EQ-HWB are captured on a 5-point Likert frequency scale *("Not at all", "Only occasionally", "Some of the time", "Often", "Most or all of the time"*), or difficulty scale *("No difficulty", "Slight", "Some", "A lot of", "Unable"*). The recall period of the EQ-HWB refers to the *"last 7 days"*. A pilot value set has been crafted to evaluate the instrument according to the preferences of the UK public {Mukuria, 2023 #16027}.

The Recovering Quality of Life measure

The Recovering Quality of Life (ReQoL) measure is a newly developed tool designed to measure the quality of life outcomes for people over 16 years of age with mental health problems (Keetharuth et al., 2018). The questionnaire was developed to be used as a routinely collected patient-reported outcome measure (PROMS) in the UK but can also be used to derive QALYs based on the value set developed for the UK population using Time Trade-off method (Keetharuth et al., 2021). The ReQoL-UI classification system covers seven domains (Table 1): activity, hope, belonging and relationship, self-perception, well-being, autonomy and physical health (Keetharuth et al., 2021). The same five response levels are attached to 6 mental health items ranging from the *"never"* to *"most or all of the time"* while response levels attached to the physical health items ranges from *"no problems"* to *"very severe problems"*.

The EQ-5D-5L

The EQ-5D-5L, a widely utilised generic health measure in economic evaluations of interventions (Roberts et al., 2019), covers five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each item comes with five levels of severity where "1" stands for no problems, "2" for mild problems, "3" for moderate problems, "4" for severe

problems, and "5" for unable/extreme problems (Herdman et al., 2011). Additionally, the EQ-5D-5L includes a Visual Analog Scale, allowing respondents to self-rate their health on a vertical scale from 0 (worst imaginable health) to 100 (best imaginable health). A value set has been developed using a larger sample of Australian general population using Discrete Choice Experiments {Norman, 2023 #16136}.

Kessler Psychological Distress Scale (K6)

Kessler Psychological Distress Scale (K6) is a clinical assessment for the risk of serious mental illness in the general population (Prochaska et al., 2012). K6 consists of self-reported 6-items of psychological distress measuring the frequency of experiencing six different feelings or experiences including *"nervous"*, *"hopeless"*, *"restless of fidgety"*, *"so depressed that nothing could cheer you up"*, *"that everything was an effort"*, *and "worthless"*. Response options ranges from *"None"* to *"All of the time"*.

Survey Questions and Instruments

Respondents filled out the EQ-HWB-S, EQ-5D-5L, ReQoL-UI and K6 questionnaires, along with supplementary questions regarding socioeconomic and health status, caregiver responsibilities, and overall health and life satisfaction. The health-related questions included self-reported questions about chronic physical and mental health conditions. The sequence of the three quality of life questionnaires was randomised for each participant to reduce potential order effects. The survey was pilot tested among researcher team, members of Monash University Health Economics Group ensure readability and comprehensibility.

Recruitment

Data were collected from a sample of Australian general population aged 18 and above sourced an online panel, Pureprofile (www.pureprofile.com). Quota sampling based on age, sex and jurisdiction was used to ensure the representativeness of the sample. Respondents were financially rewarded following completion of the survey based on Pureprofile's policy.

Data analysis

Participant characteristics and responses were analysed using descriptive analysis in STATA. Where applicable, characteristics were compared with the Australian population data derived from Wave 19 of the Household Income and Labour Dynamics Australia (HILDA) survey

except for age distribution, which was derived from Australian Bureau of Statistics. . QoL instruments were scored to produce level sum scores and a total utility score where appropriate using value sets to enable further analysis of validity.

Floor and ceiling effects

Floor and ceiling effects were examined at the item and instrument level as the proportion of respondents reporting "no problems" (or the equivalent option on the response scale) for each item (item level) and in all items of each measure (instrument level). Items with > 50% of respondents in the ceiling were identified as potentially problematic, using a threshold established by a previous study (Peasgood et al., 2022). At the instrument level, 15% was adopted as a threshold to indicate potential problems (Monteiro et al., 2022). Floor effects were identified when a high proportion of respondents reported their HRQOL using the most extreme end of the response scale. A 5% threshold was used to indicate floor effect problems at the item level (Monteiro et al., 2022).

Convergent validity was evaluated whether instruments of similar concept agree with each other. The EQ-HWB-S items was compared to existing measures that include the EQ-5D-5L, ReQoL-UI and K6. The relationship between measures was based on Spearman's rank absolute correlation strength (ACS) coefficient. We expected stronger correlation between items covering similar dimensions among measures (e.g. autonomy in ReQoL-UI and EQ-HWB-S). Decision rules were based on the following categories for evidence of correlation: > 0.6 = very strong, ≥ 0.5 to < 0.6 = strong, < 0.5 to ≥ 0.3 =moderate, and < 0.3 =weak. Statistical significance levels were also derived for all correlations.

Known group differences were examined by investigating *a priori* hypothesised relationships between the EQ-HWB-S items and respondent characteristics, where we hypothesise that EQ-HWB-S scores will be lower for respondents who: (a) report poorer general health, (b) report low satisfaction with their health and their life, (c) self-report any chronic physical health condition, (d) self-report any mental health condition, (e) have psychological distress measured by the K6 scoresSpecifically, the hypotheses for each group were classified as follow.

(a) Respondents with poorer general health, defined as those whose general health was reported as fair or poor, will have lower EQ-HWB-S scores compared to respondents who reported their general health as good, very good or excellent.

(b&c) Respondents who reported that their health and life satisfaction was less than 5 on a scale of 0 to 10 will have lower EQ-HWB-S scores compared to those reporting higher or equal to 5.
(d) Respondents with a chronic health condition (which lasted or expected to last more than 6 months) would have lower EQ-HWB-S scores compared to those without.
(e) Respondents who have self-reported that they have one or more mental health condition(s) would have lower EQ-HWB-S scores compared to those who do not have mental a health condition.
(f)) Respondents who have a K6 score of 19 or higher (clinical threshold) would have lower EQ-HWB-S scores compared to respondents who have the K6 scores less than 19.

Known-group validity examines how EQ-HWB-S scores vary among groups that are anticipated to have different characteristics. This is assessed by employing Cohen's d standardized absolute effect sizes (AES), which capture the difference in mean scores between two adjacent severity subgroups, divided by the standard deviation of scores for the less severe subgroup. Cohen's d of 0.2 are normally considered small, 0.5 moderate, and 0.8 large. In addition to this, the non-parametric Mann-Whitney test is utilised to determine if there exists a statistically significant difference between two known groups. If the EQ-HWB-S demonstrates the capability to distinguish between multiple known groups, it provides evidence of known-group validity.

Ethics

The survey used for data collection was approved by the Ethical Review Committee of Monash University (Ref No: 39392).

Results

Data and Sample

Table 1 summarizes the sample's sociodemographic characteristics. In total, 2024 Australians participated in the survey. Our surveyed population exhibited a higher proportion of individuals aged 55 and above compared to national demographics. Additionally, there is a slight overrepresentation of males and a slightly lower proportion of individuals born in Australia. Furthermore, the surveyed participants tended to have higher levels of education. However, their marital status closely aligned with that of the national population.

Response Distribution

Table 2 illustrates the response distribution for the nine items of the EQ-HWB-S across the full general population, as well as subgroups with and without mental health conditions. Notably, items 3–9 (exhaustion, loneliness, cognition, anxiety, sad/depressed, control, and pain) demonstrate a more balanced distribution of responses, whereas items 1 (mobility) and 2 (activities) exhibit skewed distributions across all presented samples. Specifically, item 1 (mobility) indicates that 85% and 82% of participants reported no difficulty with mobility in the full sample and the sample with mental health conditions, respectively. Similarly, item 2 (activities) reveals that 64% and 58.3% of participants reported no difficulty with day-to-day activities in the full sample and the sample with mental health conditions, respectively. Moreover, the response option "none of the time" was the most frequently selected response for items related to loneliness (37.1%), anxiety (23.6%), sad/depression (32.1%), and control (41%), while it was the least selected response for exhaustion (8.8%). Conversely, all items were detected to have floor effects at the item level except for the exhaustion item with 5.1% participants reporting the most severe level -“most of the time” in the full general population sample. However, seven items out of nine did not exhibit floor effects in the sample with mental health conditions including exhaustion (13.4%), loneliness (6.9%), cognition (6.9%), anxiety (10%), sad/depression (7.1%), and control (5.2%).

Ceiling and Floor Effects at the instrument level

The EQ-HWB-S demonstrates no ceiling or floor effects with only 8% participants reporting the best health state (111111111) while no participant reported the worst health state (555555555).

Convergent validity- item level

Table 4 summarises the associations between items of EQ-5D-5L, ReQoL-UI, K6 compared to items of EQ-HWB-S. Overlapping dimensions were moderately or strongly correlated between EQ-HWB-S and EQ-5D-5L, for example “mobility” and “pain” in EQ-HWB-S and “mobility” in EQ-5D-5L. Similar results were also displayed between EQ-HWB-S and ReQoL-UI where there were strong correlation between “loneliness” in EQ-HWB-S and “belonging and relationship” in ReQoL-UI. Table 4 provides a clear overview of the associations between items of EQ-HWB-S, EQ-5D-5L, ReQoL-UI with items of K6 on their performance in mental

health. As expected, items 3-8 including exhaustion, loneliness, cognition, anxiety, sad/depression and control were moderately or strongly correlated with all items of the K6. As for the EQ-5D-5L, strong correlation was only observed between anxiety/depression item and all items of the K6. When examining the ReQoL-UI, all mental health items were moderately correlated with all K6 items except the 'hope' item, which was weakly correlated with the item "…restless or fidgety" of the K6.

Convergent validity- preference weight and sum scores

The EQ-HWB-S sum score was strongly correlated with the sum scores of EQ-5D-5L, ReQoL-UI and K6. As expected, the sum score of K6 was strongly correlated with that of EQ-HWB-S and ReQoL-UI while only moderately correlated with that of EQ-5D-5L. The EQ-HWB-S preference weighted score was strongly correlated with those of EQ-5D-5L and the ReQoL-UI (Table 4).

Known Group Validity

Table 5 summarises the findings of known-group validity of EQ-HWB-S across five known-groups. Both the EQ-HWB-S preference-weighted score or sum-score were significantly different between known groups in the hypothesised directions with *p*-values of both t-test and Mann-Whitney less than 0.001. Effect sizes, although significant, were mostly small.

Discussion

This is the first study comparing the psychometric performance of the EQ-HWB-S with the ReQoL-UI, EQ-5D-5L and K6 in a representative sample of Australian general population aged 18 years and over and people with mental health conditions. Overall, the EQ-HWB-S performed well in this large sample of Australians and people with mental health conditions, supporting its legitimacy as a potential outcome measure for mental health and wellbeing. Specifically, the EQ-HWB-S exhibited superior convergent validity with the K6 and performed well in differentiating among individuals in the known-groups analyses using level sum-scores and preference-weighted scores.

Although no ceiling or floor effects were detected on the sum scores of EQ-HWB-S, response distributions show that some items were highly skewed. Specifically, item "mobility" and "activities" have 85% and 64% of participants reporting "no problems", respectively in the full

general population sample. This finding is consistent with the limited literature on the ceiling effects of the EQ-HWB-S (Bailey et al., 2024). In contrast, none items exhibited floor effects at the item level in the full general population sample except the exhaustion items which was consistent with previous study {Monteiro, 2022 #16129;Peasgood, 2022 #16130}.

The item-level analysis of convergent validity revealed a surprising strength of the EQ-HWB-S compared to the ReQoL-UI, despite the latter's specific design for individuals with mental health problems. Remarkably, all items of the K-6 exhibited moderate to high correlations with the psychological items of the EQ-HWB-S. This finding, combined with the EQ-HWB-S' strong agreement with other measures and its significant ability to identify individuals with and without mental health problems, underscores its potential as a valuable outcome measure in this domain. Serving as a generic preference-based measure, this finding constitutes a significant contribution to the literature, offering compelling evidence for EQ-HWB-S' suitability in facilitating cross-sector comparisons for informed decision-making.

The robust correlation observed between EQ-HWB-S and EQ-5D-5L implies its suitability for capturing quality of life outcomes in general populations. Given that the EQ-HWB-S encompasses a wider spectrum of wellbeing outcomes, it presents distinct advantages over the EQ-5D-5L in national population surveys, particularly when aiming to comprehensively assess the dynamic health and wellbeing trends of the population.

Strengths and limitations

Our study has several strengths. Firstly, we provide a direct comparison of EQ-HWB-S with EQ-5D-5L, ReQoL-UI in a substantial and representative of Australian general population, supporting the generalizability of the results in promoting the use of EQ-HWB-S to measure population norms of quality of life. Secondly, our results of psychometric performance in mental health were based both on self-reporting questions and psychological screening tool (Kessler-6). However, our study is not without limitations. Although we endeavoured to conduct a psychometric comparison using preference weighted scores, we must acknowledge that the utility values for the EQ-HWB-S and the ReQoL-UI were derived from the UK population, while those for the EQ-5D-5L were from the Australian population, due to the lack of availability of the former utility weights from Australian populations. It is unclear how our results would have differed if Australian value sets were available for the EQ-HWB-S and the

ReQoL-UI, which may better reflect the preferences of Australian populations. However, the findings based on sum-scores demonstrate similar results, providing reassurance regarding the psychometric properties of the EQ-HWB-S. Secondly, we acknowledge the exclusive inclusion of the K6 in our convergent analysis. Since the K6 primarily focuses on psychological distress, it may be potential overlooking other mental health problems detectable by alternative specific clinical questionnaires, such as the General Anxiety Disorder-7 or Patient Health Questionnaire-9. Although there is significant evidence of the validity of K6 in diagnosing a wide range of mental health problems (Sakurai et al., 2011; Umucu et al., 2021), we urge future research to explore the convergent analysis of EQ-HWB-S with these questionnaires to furnish more comprehensive evidence of its suitability as a mental health measure.

Conclusion

Our findings using a large representative sample of Australians aged 18 and above support the legitimacy of the EQ-HWB-S as a potential outcome measure for economic evaluation and populations studies. Specifically, our findings show a convergence between EQ-HWB-S and EQ-5D-5l, ReQoL-UI, and the K6 while the EQ-HWB-S performed well in differentiating among individuals with known groups including mental health problems. More in-depth investigation on construct validity such as factor analysis and item response analysis are imperative for a more definitive assessment of the psychometric properties of the EQ-HWB-S.

Tables and Figures

Table 1: Descriptive statistics of the full sample

Variables	N(%)	Comparable Australian data (%)
Age (years)		
18-24	159 (7.86)	11
25-34	288 (14.23)	18.6
35-44	272 (13.44)	17.6
45-54	263 (12.99)	16.3
55-64	499 (24.65)	15
+65	543 (26.83)	21.5
Gender		
Male	1134 (56.03)	47.4
Female	890 (43.97)	52.6
Country of birth - Australia	1578 (77.96)	79.8
Highest education level		
Primary/Secondary school	625 (30.88)	39.4
III/IV or Diploma	616 (30.43)	33
University degree	748 (36.96)	27.5
Gross household income per annum		
< A$52 000	793 (39.19)	
A$52,000-$159,999	770 (38.04)	
> A$130,000	294 (14.52)	
Marital status		
Single	498 (24.6)	23.6
Married/De facto	1196 (59.09)	59
Separated/divorced	231 (11.42)	12.6
Widow	91 (4.5)	4.9
Employment		
Employed (full-time, part-time, casual, self-employed)	1076 (53.16)	
Retired, housework duties including carer tasks r	760 (37.55)	
Student, unemployed or other	182 (8.99)	

Notes: Comparable Australian data derived from Wave 19 of the Household Income and Labour Dynamics Australia (HILDA) survey except age distribution derived from Australian Bureau of Statistics

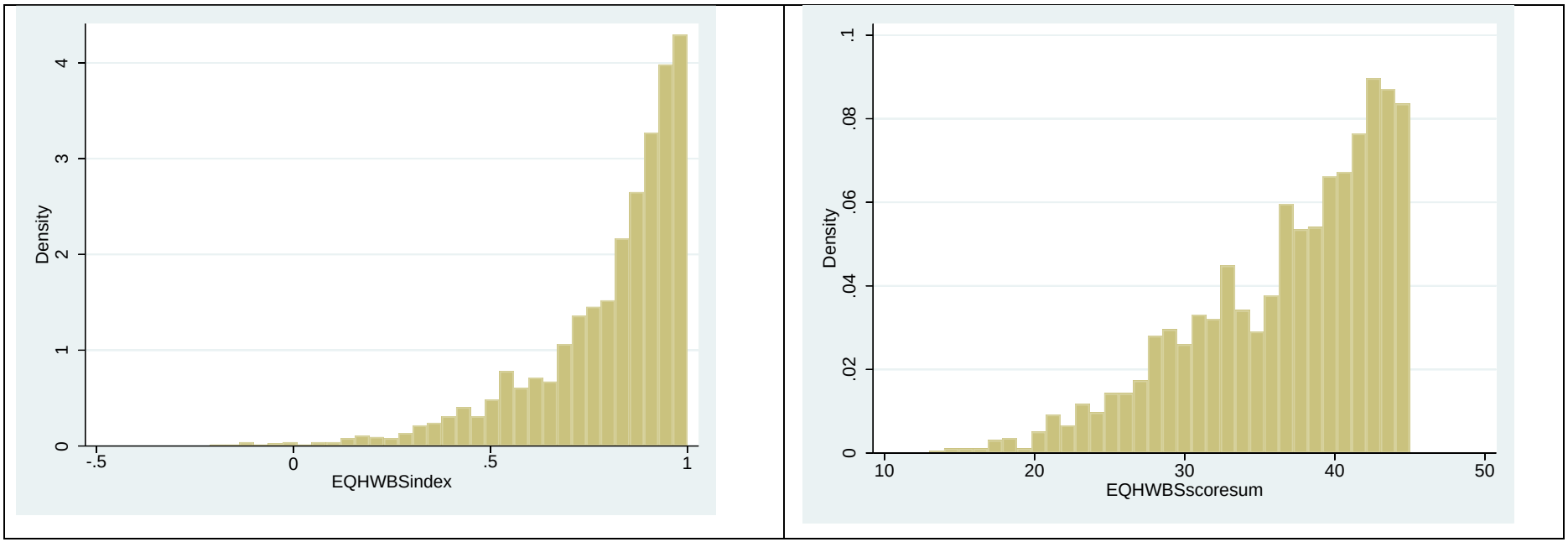

Figure 1: Distribution of EQ-HWB-S preference-weight score and sum scores

Table 1: Frequency of observations by EQ-HWB-S item for full, general population with and without any mental health samples (N = 2024)

	Full sample	**General population without a mental health condition**	**General Population with any Mental Health condition**	**Percentage difference (%)**
	***n* = 2024**	***n* = 1443**	***n* = 581**	
	Number (%)	**Number (%)**	**Number (%)**	
Mobility				
No difficulty	1538 (76.0)	1160 (80.4)	378 (65.1)	15.33
Slight difficulty	260 (12.9)	159 (11.0)	101 (17.4)	-6.36
Some difficulty	167 (8.3)	89 (6.2)	78 (13.4)	-7.26
A lot of difficulty	55 (2.7)	32 (2.2)	23 (4.0)	-1.74
Unable	4 (0.2)	3 (0.2)	1 (0.2)	0.04
Activities				
No difficulty	1316 (65.0)	1056 (73.2)	260 (44.8)	28.43
Slight difficulty	403 (19.9)	239 (16.6)	164 (28.2)	-11.67
Some difficulty	211 (10.4)	102 (7.1)	109 (18.8)	-11.69
A lot of difficulty	84 (4.2)	42 (2.9)	42 (7.2)	-4.32
Unable	10 (0.5)	4 (0.3)	6 (1.0)	-0.75
Exhaustion				
None of the time	510 (25.2)	454 (31.5)	56 (9.6)	21.82
Only occasionally	703 (34.7)	567 (39.3)	136 (23.4)	15.88

Sometimes	440 (21.7)	261 (18.1)	179 (30.8)	-12.72
Often	267 (13.2)	135 (9.4)	132 (22.7)	-13.36
Most or all of the time	104 (5.1)	26 (1.8)	78 (13.4)	-11.63
Loneliness				
None of the time	980 (48.4)	839 (58.1)	141 (24.3)	33.87
Only occasionally	465 (23.0)	324 (22.5)	141 (24.3)	-1.82
Sometimes	371 (18.3)	198 (13.7)	173 (29.8)	-16.06
Often	151 (7.5)	65 (4.5)	86 (14.8)	-10.3
Most or all of the time	57 (2.8)	17 (1.2)	40 (6.9)	-5.7
Cognition				
None of the time	866 (42.8)	769 (53.3)	97 (16.7)	36.59
Only occasionally	575 (28.4)	417 (28.9)	158 (27.2)	1.71
Sometimes	382 (18.9)	197 (13.7)	185 (31.8)	-18.19
Often	154 (7.6)	52 (3.6)	102 (17.6)	-13.96
Most or all of the time	47 (2.3)	8 (0.6)	39 (6.7)	-6.16
Anxiety				
None of the time	816 (40.3)	753 (52.2)	63 (10.8)	41.34
Only occasionally	554 (27.4)	416 (28.8)	138 (23.8)	5.08
Sometimes	387 (19.1)	203 (14.1)	184 (31.7)	-17.6
Often	201 (9.9)	63 (4.4)	138 (23.8)	-19.38

Most or all of the time	66 (3.3)	8 (0.6)	58 (10.0)	-9.43
Sad/depression				
None of the time	936 (46.3)	838 (58.1)	98 (16.9)	41.2
Only occasionally	530 (26.2)	374 (25.9)	156 (26.9)	-0.93
Sometimes	344 (17.0)	175 (12.1)	169 (29.1)	-16.96
Often	168 (8.3)	51 (3.5)	117 (20.1)	-16.61
Most or all of the time	46 (2.3)	5 (0.4)	41 (7.1)	-6.71
Control				
None of the time	1128 (55.7)	956 (66.3)	172 (29.6)	36.65
Only occasionally	415 (20.5)	260 (18.0)	155 (26.7)	-8.66
Sometimes	303 (15.0)	160 (11.1)	143 (24.6)	-13.52
Often	123 (6.1)	42 (2.9)	81 (13.9)	-11.03
Most or all of the time	55 (2.7)	25 (1.7)	30 (5.2)	-3.43
Pain				
No physical pain	755 (37.3)	623 (43.2)	132 (22.7)	20.45
Mild pain	873 (43.1)	616 (42.7)	257 (44.2)	-1.54
Moderate pain	301 (14.9)	163 (11.3)	138 (23.8)	-12.45
Severe pain	80 (4.0)	39 (2.7)	41 (7.1)	-4.36
Very severe pain	15 (0.7)	2 (0.1)	13 (2.2)	-2.1

Table 3: : Convergent validity of EQ-5D-5L and ReQoL-UI compared to EQ-HWB-S at the item level

EQ-HWB-S dimensions	Mobility	Activities	Exhaustion	Loneliness	Cognition	Anxiety	Sad/depression	Control	Pain
EQ-5D-5L dimensions									
Mobility	0.66	**0.63**	0.25	0.10	0.18	0.14	0.18	0.23	0.57
Self-care	0.53	0.54	0.24	0.16	0.22	0.18	0.22	0.28	0.39
Usual activities	0.55	0.73	0.37	0.23	0.32	0.27	0.31	0.37	0.55
Pain/discomfort	0.46	0.54	0.31	0.15	0.24	0.19	0.23	0.27	0.83
Anxiety/depression	0.19	0.33	0.49	**0.59**	**0.61**	**0.74**	**0.73**	**0.59**	0.26
ReQoL-UI dimensions									
Activity	0.17	0.28	0.41	0.47	0.48	0.50	0.54	0.50	0.19
Belonging and relationship	0.12	0.23	0.39	0.80	0.50	0.54	0.61	0.48	0.15
Choice and autonomy	0.21	0.34	0.49	0.52	0.58	0.59	0.61	**0.56**	0.21
Hope	0.21	0.30	0.34	0.49	0.46	0.46	0.56	0.48	0.18
Self-perception	0.14	0.24	0.45	0.49	0.54	0.58	0.56	0.48	0.18
Well-being	0.16	0.27	0.43	0.52	0.50	0.56	0.62	0.53	0.19
Physical health	**0.51**	0.59	0.36	0.23	0.31	0.28	0.32	0.34	**0.65**
Kessler-6 dimensions									
. . . so sad that nothing could cheer you?	-0.23	-0.33	-0.48	-0.63	-0.58	-0.64	-0.75	-0.60	-0.22
. . . nervous?	-0.16	-0.25	-0.51	-0.54	-0.60	-0.77	-0.61	-0.53	-0.17
. . . restless or fidgety?	-0.17	-0.26	-0.47	-0.50	-0.57	-0.62	-0.54	-0.49	-0.20
. . . hopeless?	-0.21	-0.34	-0.52	-0.62	-0.61	-0.66	-0.71	-0.65	-0.20
. . . that everything was an effort?	-0.29	-0.43	-0.62	-0.50	-0.64	-0.59	-0.62	-0.60	-0.33

. . . worthless?	-0.22	-0.32	-0.48	-0.61	-0.59	-0.61	-0.69	-0.61	-0.22

Table 4: Convergent validity of EQ-HWB-S, EQ-5D-5L and ReQoL-UI compared to K6 at the item level

Kessler-6 dimensions	. . . so sad that nothing could cheer you?	. . . nervous?	. . . restless or fidgety?	. . . hopeless?	. . . that everything was an effort?	. . . worthless?
EQ-HWB-S dimensions						
Mobility	-0.23	-0.16	-0.17	-0.21	-0.29	-0.22
Activities	-0.33	-0.25	-0.26	-0.34	-0.43	-0.32
Exhaustion	**-0.48**	**-0.51**	**-0.47**	**-0.52**	**-0.62**	**-0.48**
Loneliness	**-0.63**	**-0.54**	**-0.50**	**-0.62**	**-0.50**	**-0.61**
Cognition	**-0.58**	**-0.60**	**-0.57**	**-0.61**	**-0.64**	**-0.59**
Anxiety	**-0.64**	**-0.77**	**-0.62**	**-0.66**	**-0.59**	**-0.61**
Sad/depression	**-0.75**	**-0.61**	**-0.54**	**-0.71**	**-0.62**	**-0.69**
Control	**-0.60**	**-0.53**	**-0.49**	**-0.65**	**-0.60**	**-0.61**
Pain	-0.22	-0.17	-0.20	-0.20	-0.33	-0.22
EQ-5D-5L dimensions						
Mobility	-0.19	-0.12	-0.12	-0.16	-0.28	-0.18
Self-care	-0.24	-0.17	-0.16	-0.24	-0.27	-0.23
Usual activities	-0.29	-0.23	-0.24	-0.32	-0.42	-0.30
Pain/discomfort	-0.21	-0.14	-0.20	-0.19	-0.30	-0.21
Anxiety/depression	**-0.68**	**-0.65**	**-0.56**	**-0.64**	**-0.58**	**-0.63**
ReQoL-UI dimensions						
Activity	**-0.52**	**-0.47**	**-0.41**	**-0.55**	**-0.52**	**-0.51**
Belonging and relationship	**-0.58**	**-0.50**	**-0.48**	**-0.57**	**-0.49**	**-0.56**
Choice and autonomy	**-0.61**	**-0.57**	**-0.49**	**-0.65**	**-0.59**	**-0.60**
Hope	**-0.58**	**-0.44**	-0.39	**-0.60**	**-0.47**	**-0.63**
Self-perception	**-0.53**	**-0.55**	**-0.45**	**-0.59**	**-0.51**	**-0.57**
Well-being	**-0.58**	**-0.50**	**-0.44**	**-0.59**	**-0.54**	**-0.55**

Physical health	-0.31	-0.24	-0.26	-0.31	-0.41	-0.30

Table 2: Convergent validity analysis at instrument levels

	EQ-HWB-S		**EQ-5D-5L**		**ReQoL**		**Kessler-6**	
	Preference-weighted score	**Sum-score**	***Preference-weighted score***	**Sum-score**	***Preference-weighted score***	**Sum-score**	***Preference-weighted score***	**Sum-score**
EQ-HWB-S								
Preference-weighted score	1.00	-						
Sum-score	-	1.00						
EQ-5D-5L								
Preference-weighted score	0.77	-	1.00	-				
Sum-score	-	0.73	-	1.00				
ReQoL-UI								
Preference-weighted score	0.80	-	0.67	-	1.00			
Sum-score	-	0.79	-	0.58		1.00		
Kessler-6								
Preference-weighted score	-	-	-	-	-	-	-	-
Sum-score	-	-0.83		-0.56		-0.80	-	1.00

Table 5: Known Group Validity of EQ-HWB-S across known groups

	Yes		**No**		**Mean**	**Mann-Whitney**	**Cohen's d5**
	n	**Mean (SD)**	***n***	**Mean (SD)**	**difference**		
Poor/Fair general health status							
Preference-weighted score	425	0.59 (0.24)	1599	0.85 (0.14)	0.264***	< 0.001	1.59

Sum-score	425	30.71 (6.89)	1599	38.76 (5.35)	8.047***	< 0.001	1.41
Low Health satisfaction							
Preference-weighted score	304	0.54 (0.24)	1720	0.84 (0.15)	0.306***	< 0.001	1.86
Sum-score	304	29.29 (6.65)	1720	38.44 (5.53)	9.157***	< 0.001	1.60
Low Life satisfaction							
Preference-weighted score	238	0.52 (0.24)	1786	0.84 (0.16)	0.318***	< 0.001	1.89
Sum-score	238	27.93 (6.10)	1786	38.29 (5.62)	10.36***	< 0.001	1.82
Self-reported chronic physical health conditions							
Preference-weighted score	1293	0.75 (0.21)	731	0.88 (0.13)	0.127***	< 0.001	0.68
Sum-score	1293	35.74 (6.86)	731	39.41 (5.30)	3.674***	< 0.001	0.58
Self-reported mental conditions							
Preference-weighted score	581	0.66 (0.23)	1443	0.86 (0.15)	0.197***	< 0.001	1.12
Sum-score	581	31.86 (6.71)	1443	39.16 (5.23)	7.299***	< 0.001	1.28
Kessler-6 (clinical threshold > 19 for psychological distress)							
Preference-weighted score	1327	0.89 (0.12)	250	0.48 (0.22)	0.403***	< 0.001	2.94
Sum-score	1327	40.50 (3.88)	250	26.32 (5.05)	14.17***	< 0.001	3.47

*p<0.05,**p<0.01,***p<0.001

References

Bailey, C., Dalziel, K., Jones, R., Hiscock, H., Devlin, N. J., & Peasgood, T. (2024, Jan 18). The Validity of the EuroQol Health and Wellbeing Short Version (EQ-HWB-S) Instrument in Parents of Children With and Without Health Conditions. *Pharmacoeconomics*. https://doi.org/10.1007/s40273-024-01351-5

Brazier, J. (2010). Is the EQ–5D fit for purpose in mental health? *British Journal of Psychiatry*, *197*(5), 348-349. https://doi.org/10.1192/bjp.bp.110.082453

Brazier, J., Connell, J., Papaioannou, D., Mukuria, C., Mulhern, B., Peasgood, T., Jones, M. L., Paisley, S., O'Cathain, A., Barkham, M., Knapp, M., Byford, S., Gilbody, S., & Parry, G. (2014, May). 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*, *18*(34), vii-viii, xiii-xxv, 1-188. https://doi.org/10.3310/hta18340

Brazier, J., Peasgood, T., Mukuria, C., Marten, O., Kreimeier, S., Luo, N., Mulhern, B., Pickard, A. S., Augustovski, F., Greiner, W., Engel, L., Belizan, M., Yang, Z., Monteiro, A., Kuharic, M., Gibbons, L., Ludwig, K., Carlton, J., Connell, J., Rand, S., Devlin, N., Jones, K., Tsuchiya, A., Lovett, R., Naidoo, B., Rowen, D., & Rejon-Parrilla, J. C. (2022, 2022/04/01/). The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. *Value in Health*, *25*(4), 482-491. https://doi.org/https://doi.org/10.1016/j.jval.2022.01.009

Brazier, J., Ratcliffe, J., Saloman, J., & Tsuchiya, A. (2016). *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford University Press. https://doi.org/10.1093/med/9780198725923.001.0001

Carlton, J., Peasgood, T., Mukuria, C., Connell, J., Brazier, J., Ludwig, K., Marten, O., Kreimeier, S., Engel, L., Belizán, M., Yang, Z., Monteiro, A., Kuharic, M., Luo, N., Mulhern, B., Greiner, W., Pickard, S., & Augustovski, F. (2022, 2022/04/01/). Generation, Selection, and Face Validation of Items for a New Generic Measure of Quality of Life: The EQ-HWB. *Value in Health*, *25*(4), 512-524. https://doi.org/https://doi.org/10.1016/j.jval.2021.12.007

Department of Health, A. G. (2016). *Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee, Version 5.0, September 2016*. https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011, 2011/12/01). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727-1736. https://doi.org/10.1007/s11136-011-9903-x

Keetharuth, A. D., Brazier, J., Connell, J., Bjorner, J. B., Carlton, J., Taylor Buck, E., Ricketts, T., McKendrick, K., Browne, J., Croudace, T., & Barkham, M. (2018). Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. *The British Journal of Psychiatry, 212*(1), 42-49. https://doi.org/10.1192/bjp.2017.10

Keetharuth, A. D., Rowen, D., Bjorner, J. B., & Brazier, J. (2021, 2021/02/01/). Estimating a Preference-Based Index for Mental Health From the Recovering Quality of Life Measure: Valuation of Recovering Quality of Life Utility Index. *Value in Health, 24*(2), 281-290. https://doi.org/https://doi.org/10.1016/j.jval.2020.10.012

Monteiro, A. L., Kuharic, M., & Pickard, A. S. (2022, 2022/04/01/). A Comparison of a Preliminary Version of the EQ-HWB Short and the 5-Level Version EQ-5D. *Value in Health, 25*(4), 534-543. https://doi.org/https://doi.org/10.1016/j.jval.2022.01.003

Mukuria, C., Peasgood, T., McDool, E., Norman, R., Rowen, D., & Brazier, J. (2023a, 2023/02/18/). Valuing the EQ Health and Wellbeing Short Using Time Trade-Off and a Discrete Choice Experiment: A Feasibility Study. *Value in Health*. https://doi.org/https://doi.org/10.1016/j.jval.2023.02.008

Mukuria, C., Peasgood, T., McDool, E., Norman, R., Rowen, D., & Brazier, J. (2023b, 2023/07/01/). Valuing the EQ Health and Wellbeing Short Using Time Trade-Off and a Discrete Choice Experiment: A Feasibility Study. *Value in Health*, *26*(7), 1073-1084. https://doi.org/https://doi.org/10.1016/j.jval.2023.02.008

National Institute for Health Care Excellence. (2013). NICE Process and Methods Guides. In *Guide to the Methods of Technology Appraisal 2013*. National Institute for Health and Care Excellence (NICE).

Peasgood, T., Carlton, J., & Brazier, J. E. (2019). A Qualitative Study of the Views of Health and Social Care Decision-makers on the Role of Wellbeing in Resource Allocation Decisions in the UK. *Economies*.

Peasgood, T., Mukuria, C., Brazier, J., Marten, O., Kreimeier, S., Luo, N., Mulhern, B., Greiner, W., Pickard, A. S., Augustovski, F., Engel, L., Gibbons, L., Yang, Z., Monteiro, A. L., Kuharic, M., Belizan, M., & Bjørner, J. (2022, 2022/04/01/). Developing a New Generic Health and Wellbeing Measure: Psychometric Survey Results for the EQ-HWB. *Value in Health*, *25*(4), 525-533. https://doi.org/https://doi.org/10.1016/j.jval.2021.11.1361

Prochaska, J. J., Sung, H. Y., Max, W., Shi, Y., & Ong, M. (2012, Jun). Validity study of the K6 scale as a measure of moderate mental distress based on mental health treatment need and utilization. *Int J Methods Psychiatr Res*, *21*(2), 88-97. https://doi.org/10.1002/mpr.1349

Roberts, S. L. E., Healey, A., & Sevdalis, N. (2019, 2019/07/15). Use of health economic evaluation in the implementation and improvement science fields—a systematic literature review. *Implementation Science*, *14*(1), 72. https://doi.org/10.1186/s13012-019-0901-7

Sakurai, K., Nishi, A., Kondo, K., Yanagida, K., & Kawakami, N. (2011, 2011/08/01). Screening performance of K6/K10 and other screening instruments for mood and anxiety disorders in Japan. *Psychiatry and clinical neurosciences*, *65*(5), 434-441. https://doi.org/https://doi.org/10.1111/j.1440-1819.2011.02236.x

Umucu, E., Fortuna, K., Jung, H., Bialunska, A., Lee, B., Mangadu, T., Storm, M., Ergun, G., Mozer, D. A., & Brooks, J. (2021, 2022/01/01). A National Study to Assess Validity and Psychometrics of the Short Kessler Psychological Distress Scale (K6). *Rehabilitation Counseling Bulletin*, *65*(2), 140-149. https://doi.org/10.1177/00343552211043261