Article:
Pangestu, S., Purba, F.D., Setyowibowo, H. et al. (3 more authors) (2025) The psychometric properties of the EQ-HWB and EQ-HWB-S in patients with breast cancer: a comparative analysis with EQ-5D-5L, FACT-8D, and SWEMWBS. Value in Health, 28 (3). pp. 449-459. ISSN 1098-3015

https://doi.org/10.1016/j.jval.2024.12.003

ScienceDirect

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

Preference-Based Assessments

The Psychometric Properties of the EQ-HWB and EQ-HWB-S in Patients With Breast Cancer: A Comparative Analysis With EQ-5D-5L, FACT-8D, and SWEMWBS

Stevanus Pangestu, MBA,* Fredrick Dermawan Purba, PhD,* Hari Setyowibowo, PhD, Yohana Azhar, PhD, Clara Mukuria, PhD, Fanni Rencz, PhD

ABSTRACT

Objectives: The EQ Health and Wellbeing (EQ-HWB) is a new generic measure that captures constructs beyond health-related quality of life, with a 25-item long form and a shorter 9-item version (EQ-HWB-S). This study aimed to assess the psychometric performance of both versions in breast cancer, which is the most prevalent cancer worldwide, and compare them with other instruments.

Methods: A longitudinal survey in Indonesia (2023-2024) with 300 female patients used the EQ-HWB, 5-level EQ-5D (EQ-5D-5L), Functional Assessment of Cancer Therapy – General (from which Functional Assessment of Cancer Therapy Eight Dimension [FACT-8D] was derived), and Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS, from which the Short WEMWBS was derived). Distributional characteristics, convergent validity, known-group validity (Student's *t* test or analysis of variance), test-retest reliability, and responsiveness were assessed.

Results: All patients reported problems in at least 1 EQ-HWB item. The EQ-HWB-S index (11%) had a lower ceiling than the EQ-5D-5L (35%) and the Short WEMWBS (15.3%), but not the FACT-8D (5%). EQ-HWB-S index values correlated strongly with EQ-5D-5L (r = 0.73) and FACT-8D index values (r = 0.70), whereas EQ-HWB level sum scores correlated strongly with Functional Assessment of Cancer Therapy – General (r = 0.69) and moderately with WEMWBS (r = 0.49). The EQ-HWB and EQ-HWB-S discriminated across known groups comparably with the EQ-5D-5L and FACT-8D with large effect sizes according to EuroQol visual analog scale groups, number of symptoms, and general health and exhibited excellent instrument-level test-retest reliability (intraclass correlations, 0.79-0.83) and acceptable responsiveness (standardized response means, |0.24| to |0.97|).

Conclusions: This study represents one of the first validations of the EQ-HWB and EQ-HWB-S in any clinical population. Both instrument versions demonstrate robust psychometric performance. The EQ-HWB-S can be recommended to inform resource allocation decisions of breast cancer treatments.

Keywords: breast cancer, EQ-HWB, EQ-5D, FACT-G, health-related quality of life, psychometrics, well-being.

VALUE HEALTH. 2025; 28(3):449–459

Highlights

- The EQ Health and Wellbeing (EQ-HWB) is a 25-item experimental instrument that was developed to capture constructs beyond health-related quality of life. The shorter form (EQ-HWB-S) is a preference-accompanied measure comprising a subset of 9 EQ-HWB items. Very limited evidence is available about their psychometric performance in clinical populations and from low- or middle-income countries.
- We tested the psychometric properties of the EQ-HWB and EQ-HWB-S in female patients with breast cancer in Indonesia. Neither exhibited ceiling or floor effects. The EQ-HWB-S demonstrated good convergence with corresponding items of the 5-level EQ-5D (EQ-5D-5L) and the cancer-specific Functional Assessment of Cancer Therapy Eight Dimension (FACT-8D) measures, comparable known-group validity with their index values, and excellent instrument-level test-retest reliability. EQ-HWB items of exhaustion, pain, discomfort, and sleep seem particularly relevant for this population.
- Our study provides a comprehensive assessment of the psychometric properties of EQ-HWB and EQ-HWB-S in breast cancer, which is the most common cancer worldwide. In some psychometric properties, the EQ-HWB-S outperforms the generic EQ-5D-5L (eg, ceiling) and cancer-specific FACT-8D (eg, test-retest reliability). The EQ-HWB-S may be recommended to be used in health technology assessments of breast cancer interventions.

Introduction

EQ Health and Wellbeing (EQ-HWB) is a recently developed generic instrument that goes beyond traditional measures of health-related quality of life (HRQoL) to encompass carer-related and social care-related quality of life.[1] The EQ-HWB underwent a rigorous development process, including conceptual and domain identification using literature reviews, face validation of candidate items using interviews with a wide range of stakeholders (ie, patients, social care users, caregivers, and health technology assessment experts), psychometric testing, and final item selection through consultation rounds and pilot valuation.[1–5] Two versions of the measure have been constructed: a long 25-item version (EQ-HWB) and a short 9-item version (EQ-HWB-S). The EQ-HWB is a profile measure, whereas the EQ-HWB-S is a preference-accompanied measure comprising a subset of EQ-HWB items, primarily designed as a self-classifier for economic evaluations in healthcare, social care, and public health interventions.[1]

Currently, the EQ-HWB and EQ-HWB-S

*The first 2 authors are equal contributors to this work and are designated as co-first authors.

are experimental instruments, and so far, only a few studies have reported on their measurement performance. The content validity of EQ-HWB and EQ-HWB-S has been documented in various populations from 4 countries, namely patients with cancer, caregivers, older adults, and other general populations.[6-8] Six studies from Australia, China, Italy, and the United States have validated the EQ-HWB and EQ-HWB-S in surgery patients, caregivers, social care users, and general populations with/without chronic conditions, with all assessing distributional characteristics, and convergent and known-group validity.[8-13] In some of these validation studies, the psychometric performance of EQ-HWB/EQ-HWB-S has been compared with various other preference-accompanied measures such as the Adult Social Care Outcomes Toolkit for Carers, Care-Related Quality of Life, and 5-level EQ-5D (EQ-5D-5L).[9-12] Furthermore, one study from Australia has provided test-retest reliability and responsiveness evidence on the EQ-HWB-S from a caregiver population.[8]

Little is known about the psychometric performance of EQ-HWB in clinical populations, particularly with longitudinal designs and in-person administration. Nearly all validation studies have been conducted on online panel samples or using a cross-sectional design.[9-11,13,14] Further psychometric evidence is necessary, including test-retest reliability, responsiveness, comparisons with condition-specific measures, and validations across diverse patient populations, countries, and cultures. The instruments' performance in languages beyond those of its development countries (ie, English, Spanish, Mandarin Chinese, and German) is also underexplored. Additional evidence from other Asian countries is warranted, given that the face validation and psychometric testing stages of the instrument's development in Asia were only conducted in China.[12]

Therefore, the purpose of this study is to assess the psychometric performance of the EQ-HWB and EQ-HWB-S in Indonesian patients with breast cancer, which is the most common form of cancer in 157 countries, including Indonesia.[15,16] A previous qualitative study, primarily involving patients with breast cancer, validated the content of EQ-HWB-S for cancer outcomes, highlighting its relevance in capturing patients' experiences.[7] A secondary aim is to compare their psychometric properties with other measures widely used in cancer populations, specifically the generic EQ-5D-5L, Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS), from which the Short WEMWBS (SWEMWBS) can be derived, and the cancer-specific Functional Assessment of Cancer Therapy – General (FACT-G), from which FACT Eight Dimension (FACT-8D) can be derived.

Methods

Study Design and Patients

A longitudinal data collection was conducted from September 2023 to March 2024 at the Oncology Division of Hasan Sadikin General Hospital in Bandung, Indonesia, after approval from the hospital's Research Ethics Committee (LB.02.01/X.6.5/284/2023). Under the oversight of the chief oncologist and nurses, a team of 3 graduate research assistants recruited patients in the hospital's oncology clinic waiting area. The participant inclusion criteria were (1) female patients aged 18 years or older, (2) diagnosed as having any type or stage of breast cancer, (3) undergoing active treatment, (4) cognitively able to complete the survey, (5) fluent in Indonesian, and (6) provided an informed consent. Patients undergoing their first cycle of therapy were not included. Our aim was for the first half of the recruited patients to be in an active treatment cycle at baseline and to be invited to complete a follow-up questionnaire during their next cycle (group 1). The remaining half were in their final treatment cycle at baseline and were invited to complete the follow-up at their post-treatment consultation (group 2). Every patient was compensated 100 000 Indonesian rupiah (approximately 6.3 US dollars) after completing each of the baseline and follow-up questionnaires.

Survey Instruments

Two distinct paper questionnaires were prepared for data collection: one for the patients and the other for the nurses. The patients' questionnaire used the paper-and-pencil, self-completion version of standardized measures in the official Indonesian versions presented in a fixed order: EQ-HWB, EQ-5D-5L, Functional Assessment of Chronic Illness Therapy – Comprehensive Score for Financial Toxicity, WEMWBS, and FACT-G. The Functional Assessment of Chronic Illness Therapy – Comprehensive Score for Financial Toxicity responses have been reported elsewhere.[17] All questions or items from the outcome measures were mandatory, except for the FACT-G "satisfied with sex life" item. Patients were also asked about their sociodemographic characteristics (age, marital status, education, employment, household size, and net monthly household income), caregiver use, general health, and symptoms experienced over the past 7 days. Self-reported symptoms were queried using a binary (yes/no) format, featuring 30 predefined symptoms (eg, nausea, fatigue, and hair loss) alongside an open-ended "other" option. In the follow-up questionnaire, patients were asked to rate the change in their health status using the Global Rating of Change (GRC) scale. To assess the questionnaire's feasibility and comprehensibility, a pilot test was performed with 5 patients and no further changes were made thereafter. During data collection, the research assistants distributed the questionnaire to patients and provided explanations during completion when needed. Oncology nurses provided the following clinical information on patients based on the hospital's electronic records: stage and type of breast cancer, cancer duration, metastasis status and sites, comorbidities (eg, hypertension), and current treatments (eg, immunotherapy and chemotherapy).

Health and Well-Being Measures

EQ-HWB

The EQ-HWB is a 25-item profile measure with 7 high-level domains: activity, feelings and emotions, cognition, relationships, autonomy, self-identity, and physical sensations.[1] The items 5-level response scales, which describe (1) frequency (none of the time [1], only occasionally [2], sometimes [3], often [4], most or all of the time [5]), (2) severity (no [1], mild [2], moderate [3], severe [4], very severe [5]), or (3) difficulty (no difficulty [1], slight difficulty [2], some difficulty [3], a lot of difficulty [4], unable [5]). The recall period is "in the last 7 days." The 25-item EQ-HWB was administered, from which we derived the EQ-HWB-S responses. Level sum scores (LSSs) were computed for both the EQ-HWB and EQ-HWB-S by summing the responses from the 25 and 9 items, respectively. The theoretical LSS ranges of 25 to 125 (EQ-HWB) and 9 to 45 (EQ-HWB-S) were linearly rescaled to a 0 to 100 range, with higher scores indicating better health and well-being. Furthermore, an index value was computed for the EQ-HWB-S by using the UK pilot value set, with higher scores indicating better health and well-being.[4]

EQ-5D-5L

The EQ-5D-5L is a generic preference-accompanied measure of HRQoL consisting of a descriptive system and a visual analog scale (VAS), both with a "today" recall period.[18] The descriptive system has 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each with 5 levels: no problems (1), slight problems (2), moderate problems (3), severe

problems (4), and extreme problems/unable to (5). LSSs were calculated by summing the 5 digits of the health state profile. The theoretical LSS range of 5 to 25 was linearly rescaled to 0 to 100 range for analysis, similarly to the EQ-HWB. Furthermore, an index value was assigned to each health state profile by using the Indonesian value set.[19] Higher scores in both the LSS and index value indicate better HRQoL. The EQ VAS enables respondents to indicate their overall health status using a vertical line ranging from 100 ("the best health you can imagine") to 0 ("the worst health you can imagine").

FACT-G

The FACT-G is a cancer-specific HRQoL measure with 4 domains: physical, social, emotional, and functional.[20] It has 27 items, a recall period of 7 days, and 5 response options for each item: not at all (0), a little bit (1), somewhat (2), quite a bit (3), and very much (4). The FACT-G was scored in 2 ways. First, a total score was calculated from the sum of all responses that were rescaled to a scale of 0 to 100. Imputation was undertaken as recommended to score those missing the "satisfied with sex life" item.[21,22] Second, from 9 items of the FACT-G (nausea, pain, fatigue, sleep, work, worry about worsening health, sadness, and support from family and friends), index values were computed using the FACT-8D Australian value set.[23] Both the FACT-G total score and FACT-8D index value indicated better HRQoL with higher scores.

WEMWBS

The WEMWBS is a 14-item generic measure of mental well-being, focusing on subjective well-being and psychological functioning.[24] The instrument has a recall period of "over the last 2 weeks" and 5 response options for each item: none of the time (1), rarely (2), some of the time (3), often (4), and all of the time (5). The WEMWBS total score was computed by summing all the responses and rescaled to a range of 0 to 100 for analysis. Next, from the responses of the 14 items, we derived the 7-item SWEMWBS and assigned an index value using the UK value set.[25] Higher WEMWBS total score and SWEMWBS index value reflected better mental well-being.

GRC

The GRC scale was used in the follow-up questionnaire for patients to evaluate changes in their health status compared with their previous hospital visit (ie, baseline).[26] The scale comprised a 7-point horizontal numeric rating system with the following points: much worse (−3), moderately worse (−2), a little worse (−1), unchanged (0), a little better (1), moderately better (2), and much better (3). Three patient subgroups were defined using the GRC responses: improved (1-3), worsened (−1 to −3), and unchanged (0) health. Patients from group 1 with unchanged health status were considered for the test-retest reliability analysis, whereas all patients from group 2 were included in the responsiveness analysis, which was conducted for each of the 3 subgroups of patients based on GRC.

Statistical Analyses

Descriptive characteristics were used to summarize the characteristics of the patient population, responses, and scores on all instruments. Measurement properties were assessed for both the EQ-HWB and EQ-HWB-S, comparing them with the EQ-5D-5L, FACT-G, FACT-8D, WEMWBS, and SWEMWBS where relevant. The analytical framework followed earlier work on psychometric testing of preference-accompanied measures.[27-30] All analyses were conducted using Stata 18.0 (StataCorp LLC, College Station, TX). Results were deemed statistically significant at $P < .05$.

Distributional characteristics

Ceiling and floor effects were determined by assessing the proportion of patients achieving the best and worst responses on (1) each item and (2) LSSs, total scores, or index values of the measures. Thresholds of 70% were used at the item level following previous EQ-HWB studies[5,31] and 15% at the instrument level.[32] Due to the broader domains included, we predicted that EQ-HWB-S would have a lower ceiling than EQ-5D-5L and SWEMWBS,[1,10] but not necessarily in comparison with the cancer-specific FACT-8D.

Convergent and divergent validity

Convergent validity was tested to assess the strength of the relationship between items or domains aiming to measure a similar construct and between different instruments.[32] Divergent validity was used to identify where EQ-HWB items captured aspects not covered in other measures. We used Spearman's rank-order correlations between individual items or domains, whereas, at the instrument level, Pearson's correlation was used for index values and LSSs or total scores. For individual items, raw responses were recoded to indicate a better condition in the HRQoL or well-being domain with higher scores, where applicable. Absolute correlation coefficients were interpreted as none (r = 0.00-0.09), weak (r = 0.10-0.29), moderate (r = 0.30-0.49), or strong (r = 0.50 and above).[33] We hypothesized at least moderate correlations between conceptually overlapping items, for example, (1) EQ-HWB personal care and EQ-5D-5L self-care, (2) EQ-HWB sleep and FACT-8D sleeping well, (3) EQ-HWB-S cognition and WEMWBS thinking clearly, and (4) EQ-HWB-S anxiety and EQ-5D-5L anxiety/depression and FACT-G nervous.[1,9-12] At the instrument level, due to overlapping HRQoL and well-being constructs, we expected strong correlations between (1) EQ-HWB-S, EQ-5D-5L, and FACT-8D index values[9,10] and (2) EQ-HWB and EQ-5D-5L LSS, EQ VAS, and FACT-G total score[9,10] and moderate correlations between (1) EQ-HWB-S and SWEMWBS index values and (2) EQ-HWB LSS and WEMWBS total score.[5] Meanwhile, moderate, weak, or no correlations were expected across the nonoverlapping items.

Known-group validity

Known-group validity tests were used to examine the ability of EQ-HWB LSS and EQ-HWB-S index values to distinguish between groups of patients compared with the other measures. The known groups were defined by cancer stage, EQ VAS score (≥80),[5] number of comorbidities, number of self-reported symptoms, and general health. For general health, the 5-point scale response was recategorized into 3 subgroups: poor/fair, good, and very good/excellent. Student's *t* test or analysis of variance was used to compare mean differences between the known groups. Effect sizes (ESs) and their 95% confidence intervals were calculated: Cohen's d for *t* tests and eta-squared (η^2) for analysis of variance. ESs were interpreted as trivial (d = 0–0.19, $\eta^2 < 0.01$), small (d = 0.20–0.49, η^2 = 0.01–0.05), moderate (d = 0.50–0.79, η^2 = 0.06–0.13), or large (d ≥ 0.80, $\eta^2 \geq 0.14$).[33,34] We hypothesized that patients with higher cancer stage at diagnosis, EQ VAS < 80, higher number of comorbidities and self-reported symptoms, and those who rated their health as poorer would report lower scores on EQ-HWB/EQ-HWB-S, EQ-5D-5L, FACT-G/FACT-8D, and WEMWBS/SWEMWBS.

Test-retest reliability

The test-retest reliability of EQ-HWB/EQ-HWB-S, EQ-5D-5L, FACT-G/FACT-8D, and WEMWBS/SWEMWBS items was assessed using Gwet's AC2,[35] where values of 0.0 to 0.19, 0.20 to 0.39, 0.40 to 0.59, 0.60 to 0.79, and ≥0.80 were interpreted as slight, fair, moderate, strong, and almost-perfect agreement.[36] Furthermore,

Table 1. Patient characteristics at baseline.

Characteristic	Total sample (n = 300)		Group 1 (n = 32*)		Group 2 (n = 150†)	
	n or mean	% or SD	n or mean	% or SD	n or mean	% or SD
Sociodemographic characteristics						
Age	51.26	10.29	49.34	8.93	52.35	10.54
<50 years	132	44.0%	15	46.9%	64	42.7%
50 years and older	168	56.0%	17	53.1%	86	57.3%
Marital status	-	-	-	-	-	-
Married	233	77.7%	27	84.4%	32	21.3%
Single/divorced/widower	67	22.3%	5	15.6%	118	78.7%
Education	-	-	-	-	-	-
Primary or less	92	30.7%	9	28.1%	45	30.0%
Secondary	157	52.3%	18	56.3%	78	52.0%
Tertiary	51	17.0%	5	15.6%	27	18.0%
Employment status‡	-	-	-	-	-	-
Employed	55	18.3%	8	25.0%	23	15.3%
Homemaker	221	73.7%	22	68.8%	115	76.7%
Unemployed (looking for work)	4	1.3%	1	3.1%	2	1.3%
Retired	20	6.7%	1	3.1%	10	6.7%
Household size	-	-	-	-	-	-
1-2	59	19.7%	9	28.1%	27	18.0%
3-4	140	46.7%	15	46.9%	78	52.0%
5 or more	101	33.7%	8	25.0%	45	32.0%
Net monthly household income§	-	-	-	-	-	-
5 million and below	270	90.0%	27	84.4%	137	91.3%
>5 million IDR	30	10.0%	5	15.6%	13	8.7%
Use of caregiver	184	61.3%	15	46.9%	82	54.7%
Clinical characteristics						
Cancer stage at diagnosis‖	-	-	-	-	-	-
1	26	8.7%	4	12.5%	8	5.3%
2	186	62.0%	20	62.5%	90	60.0%
3	81	27.0%	8	25.0%	49	32.7%
4	5	1.7%	0	0.0%	3	2.0%
Unknown	2	0.7%	0	0.0%	0	0.0%
Breast cancer type	-	-	-	-	-	-
Invasive lobular carcinoma	140	46.7%	18	56.3%	76	50.7%
Invasive ductal carcinoma¶	117	39.0%	12	37.5%	51	34.0%
Ductal carcinoma in situ	37	12.3%	2	6.3%	22	14.7%
Lobular carcinoma in situ	3	1.0%	-	-	1	0.7%
Inflammatory breast cancer	2	0.7%	0	0.0%	0	0.0%
Mucinous carcinoma	1	0.3%	0	0.0%	0	0.0%
Disease duration (in years)	2.45	3.19	3.09	3.61	2.90	3.26
Number of metastasis sites£	-	-	-	-	-	-
0	276	92.0%	29	90.6%	140	93.3%
1	19	6.3%	3	9.4%	9	6.0%
2	4	1.3%	0	0.0%	1	0.7%
3	1	0.3%	0	0.0%	0	0.0%
Current treatment‡	-	-	-	-	-	-
Immunotherapy	253	84.3%	28	87.5%	135	90.0%
Chemotherapy	37	12.3%	6	18.8%	8	5.3%
Radiation therapy	11	3.7%	0	0.0%	7	4.7%
Stem cell or bone marrow	2	0.7%	0	0.0%	2	1.3%
Palliative care	23	7.7%	1	3.1%	14	9.3%
Unknown	2	0.7%	0	0.0%	0	0.0%
Number of comorbidities**	-	-	-	-	-	-
0	78	26.0%	8	25.0%	44	29.3%
1	123	41.0%	15	46.9%	53	35.3%
2-3	86	28.7%	8	25.0%	44	29.3%
4 or more	13	4.3%	1	3.1%	9	6.0%
Number of symptoms	-	-	-	-	-	-
0	17	5.7%	4	12.5%	5	3.3%
1-3	71	23.7%	10	31.3%	30	20.0%

continued on next page

Table 1. Continued

Characteristic	Total sample (n = 300)		Group 1 (n = 32*)		Group 2 (n = 150†)	
	n or mean	% or SD	n or mean	% or SD	n or mean	% or SD
4-6	68	22.7%	5	15.6%	42	28.0%
7-9	60	20.0%	6	18.8%	29	19.3%
10+	84	28.0%	7	21.9%	44	29.3%
Self-reported symptoms in the past week	-	-	-	-	-	-
Fatigue	175	58.3%	16	47.1%	100	66.7%
Dizziness	143	47.7%	16	47.1%	76	50.7%
Muscle pain	133	44.3%	18	52.9%	72	48.0%
Sleep problem	123	41.0%	9	26.5%	67	44.7%
Anxiety	122	40.7%	11	32.4%	67	44.7%
Hair loss	120	40.0%	10	29.4%	64	42.3%
Skin itching	106	35.3%	13	38.2%	56	37.3%
Dry mouth	101	33.7%	10	29.4%	51	34.0%
Headache	101	33.7%	14	41.2%	51	34.0%
Weight loss	95	31.7%	8	23.5%	44	29.3%

Note. Group 1 consisted of patients in active treatment cycle at baseline who were invited to complete the follow-up questionnaire during their next cycle. Group 2 consisted of those in their final treatment cycle at baseline and were invited to complete the follow-up at their post-treatment consultation. Totals may not equal 100% due to rounding adjustments.
*Included in the test-retest reliability analysis.
†Included in the responsiveness analysis.
‡Response may belong to one more category.
§IDR = Indonesian rupiah, 5 million IDR ≈ 324 US dollars.
‖0 = noninvasive, precancerous; 1 = early stage, spread to other tissue in small area; 2 = localized, tumor between 20 and 50 mm and lymph nodes involved or tumor larger than 50 mm with no lymph nodes involved); 3 = regional spread, tumor larger than 50mm with lymph nodes involved in the larger region, may have spread to skin or chest wall; 4 = metastatic, distant spread beyond the breast and nearby lymph nodes (American Joint Committee on Cancer Staging Manual 8th ed. New York, NY: Springer; 2017:589).
¶Includes subtypes: triple-negative breast cancer, luminal A, luminal B HER-2 negative, luminal B HER-2 positive, and HER-2 positive.
£Most common sites were bone, lung, and liver.
**Most common comorbidities: chronic gastritis, hypertension, obesity, and hyperlipidemia.

intraclass correlation coefficients (ICCs), using 2-way mixed effects model with absolute agreement, and their corresponding 95% confidence intervals were computed for the instrument-level analysis: EQ-HWB-S, EQ-5D-5L, FACT-8D, SWEMWBS index values, EQ-HWB, EQ-5D-5L LSSs, EQ VAS, FACT-G total, and WEMWBS total scores.[37] ICC values of 0.0 to 0.39, 0.40 to 0.59, 0.60 to 0.74, and 0.75 to 1.0 were interpreted as poor, fair, good, and excellent test-retest reliability, respectively.[38]

Responsiveness to change

Responsiveness of the EQ-HWB-S, EQ-5D-5L, FACT-8D, SWEMWBS index values, EQ-HWB, EQ-5D-5L LSSs, EQ VAS, FACT-G total, and WEMWBS total scores was assessed using the standardized response means (SRMs) and their associated 95% confidence intervals. The SRM was estimated as the mean change in scores (or indices) between baseline and follow-up divided by the change's standard deviation. The SRM values were interpreted as small (<0.50), moderate (0.50-0.79), or large (≥0.80).[33]

Results

Characteristics of the Patient Population

Of 310 patients approached, 10 declined participation due to time constraints. The final sample included 300 female patients with breast cancer aged 51.3 ± 10.3 (Table 1). Most were married (77.7%), completed secondary education or less (83.0%), were homemakers (73.7%), and had low household income (90.0%). The majority were diagnosed as having stage 2 breast cancer (62.0%), an average of 2.45 ± 3.19 years ago, with invasive lobular (46.7%) or invasive ductal (39.0%) carcinoma. At the time of the survey, most patients were receiving immunotherapy (84.3%) or chemotherapy (13.0%). The most commonly self-reported symptoms in the past week were fatigue (58.3%), dizziness (47.7%), and muscle pain (44.3%).

Follow-up questionnaires were completed by 298 patients (response rate, 99.3%): 148 from group 1 (mean follow-up, 5.8 ± 3.0 weeks; range, 1.9-13.0 weeks) and 150 from group 2 (mean follow-up, 11.6 ± 4.0 weeks; range, 4.0-25.9 weeks). Two patients died and were excluded from the test-retest reliability and responsiveness analyses. Of group 1, 32 patients (21.6%) had unchanged health status based on the GRC scale. In group 2, 75 (50.0%), 18 (12.0%), and 57 patients (38.0%) reported improved, worsened, and unchanged health, respectively.

Distributional Characteristics

Seven EQ-HWB items exhibited ceiling effects, namely "hearing" (85.7%), "getting around inside and outside" (71.0%), "personal care" (79.7%), "unsupported" (88.3%), "nothing to look forward" (81.3%), "no control over daily life" (79.7%), and "coping" (76.7%) (Table 2). In addition, "cognition" also showed a borderline ceiling effect (69.7%). Meanwhile, EQ-5D-5L exhibited ceiling effects for 3 items and FACT-G/FACT-8D for 4 items, whereas WEMWBS/SWEMWBS showed none (see Appendix Table 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003).

Overall, the patients reported good health status at baseline. Comparable mean index values of EQ-HWB-S, EQ-5D-5L, and SWEMWBS were observed at 0.84 ± 0.17, 0.85 ± 0.21, and 0.86 ± 0.15, respectively, whereas the mean FACT-8D index was lower at 0.72 ± 0.23 (see Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003). The EQ-5D-5L index showed a more skewed distribution with a few clusters, whereas EQ-HWB-S, FACT-8D, and SWEMWBS were less skewed and had greater variability. Interestingly, the observed ranges were similar across all instruments except for

Table 2. Response distribution of EQ-HWB/EQ-HWB-S items.

No.	EQ-HWB item	Response at baseline, n (%)				
		No difficulty	Slight difficulty	Some difficulty	Much difficulty	Unable
1	Sight	138 (46.0%)	134 (44.7%)	21 (7.0%)	7 (2.3%)	-
2	Hearing	257 (85.7%)	30 (10.0%)	8 (2.7%)	4 (1.3%)	1 (0.3%)
3	Getting around inside and outside (s)	213 (71.0%)	61 (20.3%)	20 (6.7%)	5 (1.7%)	1 (0.3%)
4	Day-to-day activities (s)	201 (67.0%)	70 (23.3%)	15 (5.0%)	11 (3.7%)	3 (1.0%)
5	Personal care	239 (79.7%)	41 (13.7%)	12 (4.0%)	3 (1.0%)	1 (0.3%)
		None of the time	**Only occasionally**	**Sometimes**	**Often**	**Most of the time**
6	Sleep	123 (41.0%)	56 (18.7%)	77 (25.7%)	25 (8.3%)	19 (6.3%)
7	Exhaustion (s)	87 (29.0%)	74 (24.7%)	99 (33.0%)	30 (10.0%)	10 (3.3%)
8	Loneliness (s)	215 (71.7%)	29 (9.7%)	40 (13.3%)	11 (3.7%)	5 (1.7%)
9	Unsupported	265 (88.3%)	10 (3.3%)	18 (6.0%)	6 (2.0%)	1 (0.3%)
10	Memory	194 (64.7%)	59 (19.7%)	41 (13.7%)	6 (2.0%)	-
11	Cognition (s)	209 (69.7%)	43 (14.3%)	43 (14.3%)	5 (1.7%)	-
12	Anxiety (s)	158 (52.7%)	58 (19.3%)	66 (22.0%)	15 (5.0%)	3 (1.0%)
13	Unsafe	199 (66.3%)	54 (18.0%)	42 (14.0%)	4 (1.3%)	1 (0.3%)
14	Frustration	252 (84.0%)	23 (7.7%)	21 (7.0%)	3 (1.0%)	1 (0.3%)
15	Sadness or depression (s)	192 (64.0%)	57 (19.0%)	41 (13.7%)	8 (2.7%)	2 (0.7%)
16	Nothing to look forward	244 (81.3%)	31 (10.3%)	21 (7.0%)	3 (1.0%)	1 (0.3%)
17	No control over daily life (s)	239 (79.7%)	29 (9.7%)	23 (7.7%)	5 (1.7%)	4 (1.3%)
18	Coping	230 (76.7%)	36 (12.0%)	29 (9.7%)	3 (1.0%)	2 (0.7%)
19	Accepted by others (r)	24 (8.0%)	3 (1.0%)	29 (9.7%)	50 (16.7%)	194 (64.7%)
20	Feel good about self (r)	13 (4.3%)	14 (4.7%)	34 (11.3%)	56 (18.7%)	183 (61.0%)
21	Do things one wanted to do (r)	14 (4.7%)	13 (4.3%)	91 (30.3%)	61 (20.3%)	121 (40.3%)
22	Pain (frequency)	113 (37.7%)	58 (19.3%)	89 (29.7%)	29 (9.7%)	11 (3.7%)
		No	Mild	Moderate	Severe	Very severe
23	Pain (severity) (s)	115 (38.3%)	104 (34.7%)	64 (21.3%)	13 (4.3%)	4 (1.3%)
		None of the time	**Only occasionally**	**Sometimes**	**Often**	**Most of the time**
24	Discomfort (frequency)	105 (35.0%)	73 (24.3%)	81 (27.0%)	33 (11.0%)	8 (2.7%)
		No	**Mild**	**Moderate**	**Severe**	**Very severe**
25	Discomfort (severity)	107 (35.7%)	114 (38.0%)	61 (20.3%)	17 (5.7%)	1 (0.3%)

Note. Overall, 300 patients completed the baseline survey.
EQ-HWB indicates EQ Health and Wellbeing; EQ-HWB-S, EQ Health and Wellbeing short form; (r), reverse coded for the level sum scores; (s), EQ-HWB-S items.

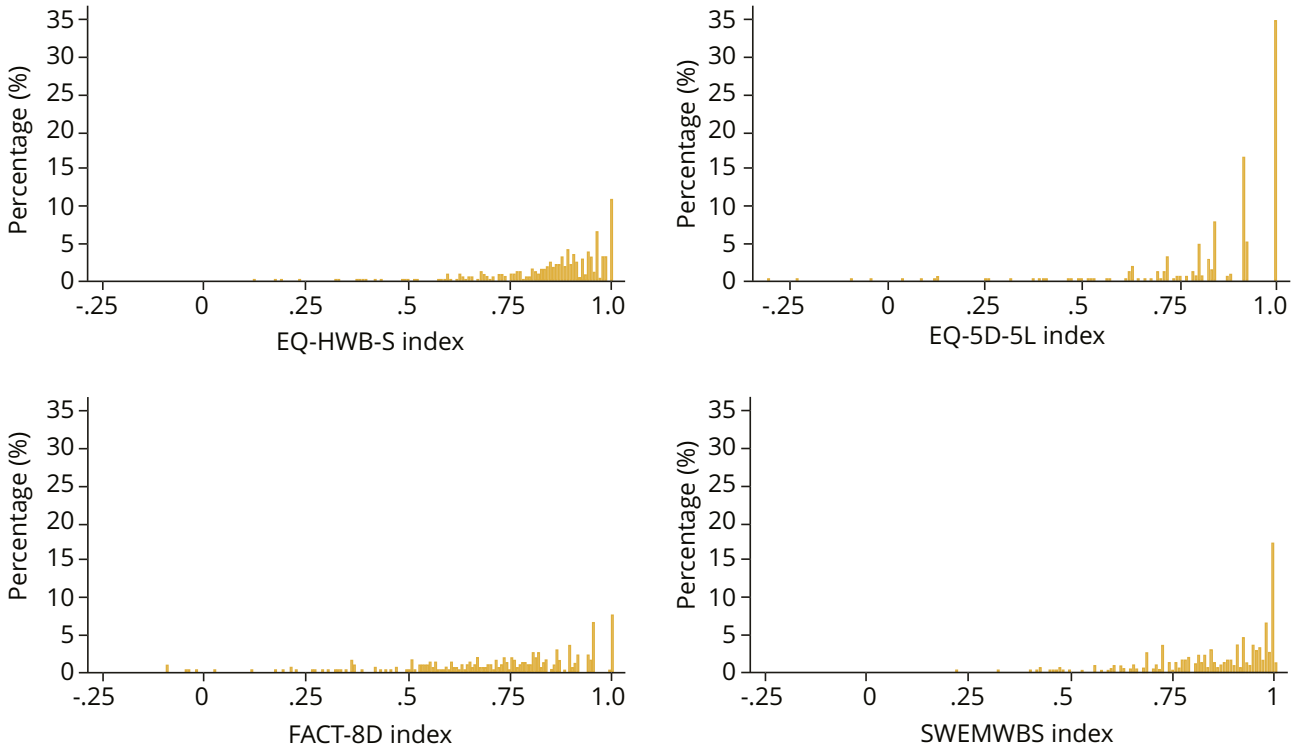
SWEMWBS, which did not exhibit negative values in this sample (Fig. 1).

The EQ-HWB-S and FACT-8D indicated no ceiling effects at the instrument level (10.7% and 5.0%) (see Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003). However, patients with no problems in any items of EQ-HWB-S reported some problems in 7 of the FACT-8D dimensions, for example, "work" (62.0%), "sleep" (46.0%), and "fatigue" (22.0%) (see Appendix Table 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003). Furthermore, EQ-5D-5L and SWEMWBS demonstrated ceiling effects of 35.0% and 15.3%, respectively. The EQ-HWB LSS, FACT-G total, and WEMWBS total scores did not exhibit ceiling effects. The mean EQ-HWB LSS, EQ-5D-5L LSS, EQ VAS, FACT-G total, and WEMWBS total scores were 83.52 ± 11.76, 90.67 ± 12.57, 81.18 ± 15.63, 76.48 ± 13.73, and 76.02 ± 17.85, respectively (see Appendix Table 4 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003).

Convergent and Divergent Validity

The correlations between EQ-HWB/EQ-HWB-S and corresponding items of EQ-5D-5L ranged from 0.31 to 0.64, with FACT-G/FACT-8D from 0.20 to 0.66, and with WEMWBS/SWEMWBS from 0.31 to 0.35 (Table 3). The strongest correlations were observed between EQ-HWB sleep and FACT-8D sleeping well (r = 0.66), EQ-HWB pain frequency and EQ-5D-5L pain/discomfort (r = 0.64), and EQ-HWB personal care and EQ-5D-5L self-care (r = 0.62). Some EQ-HWB items correlated varyingly with the composite domains of EQ-5D-5L, for example, a strong correlation between EQ-5D-5L anxiety/depression and EQ-HWB anxiety but moderately with EQ-HWB sadness/depression and strong correlation between EQ-5D-5L pain/discomfort and EQ-HWB pain (severity and frequency) but moderately with EQ-HWB discomfort (severity and frequency). Furthermore, in most cases, none or weak correlations were observed between nonoverlapping items (see

Figure 1. Distribution of the EQ-HWB-S, EQ-5D-5L, FACT-8D, and SWEMWBS index values.



EQ-HWB indicates EQ Health and Wellbeing; EQ-HWB-S, EQ Health and Wellbeing short form; EQ-5D-5L, 5-level EQ-5D; FACT-8D, Functional Assessment of Cancer Therapy Eight Dimension; SWEMWBS, Short Warwick-Edinburgh Mental Wellbeing Scale.

Appendix Table 5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003).

At the instrument level, strong correlations were observed between EQ-HWB-S index and EQ-5D-5L (r = 0.73) and FACT-8D (r = 0.70) and between EQ-HWB LSS and EQ-5D-5L LSS (r = 0.65), EQ VAS (r = 0.50), and FACT-G total score (r = 0.69). Moderate correlations were shown between EQ-HWB-S and SWEMWBS index (r = 0.34) and between EQ-HWB LSS and WEMWBS total score (r = 0.49) (Table 4).

Known-Group Validity

The EQ-HWB and EQ-HWB-S significantly distinguished known groups, with large ESs for EQ VAS, general health, and number of symptoms, and small ES for comorbidities (Table 4). All instruments, except SWEMWBS, performed similarly for EQ VAS known groups (large ESs) (see Appendix Table 6 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003). EQ-HWB-S performed similarly to FACT-8D and outperformed EQ-5D-5L index and SWEMWBS for general health (large vs moderate ESs). For the number of symptoms, EQ-HWB-S performed comparably with EQ-5D-5L and FACT-8D and better than SWEMWBS (large vs moderate ESs). For the cancer stage, EQ-HWB-S was comparable with EQ-5D-5L index and better than FACT-8D and SWEMWBS (borderline small vs trivial ESs). EQ-HWB-S performed better than EQ-HWB in all comparisons except for the number of symptoms.

Test-Retest Reliability

Across EQ-HWB, 7 items demonstrated almost-perfect agreement, 10 strong agreement, 5 moderate agreement, and 3 fair agreement (see Appendix Table 7 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.12.003). The best-performing items included "unsupported," "getting around inside and outside," "personal care," "frustration," "nothing to look forward," and "no control over daily life" (Gwet's AC2, 0.87-0.97), whereas the lowest-performing in terms of test-retest reliability were "anxiety," "exhaustion," and "sleep" (Gwet's AC2, 0.32-0.35). In comparison, Gwet's AC2 ranged from 0.64 to 0.97 for EQ-5D-5L, 0.24 to 0.90 for FACT-G, and 0.29 to 0.69 for WEMWBS.

At the instrument level, the EQ-HWB-S index and EQ-HWB LSS demonstrated excellent reliability with ICCs of 0.83 and 0.79, respectively (Table 5). In comparison, the EQ-5D-5L had higher ICCs at 0.89 (LSS) and 0.88 (index value), whereas FACT-8D index and FACT-G total scores were lower at 0.77 and 0.76, respectively. Notably, the SWEMWBS index and WEMWBS total score had the lowest ICCs among the instruments at 0.53 and 0.50, respectively.

Responsiveness

Small to moderate responsiveness to change was demonstrated by the EQ-HWB-S index in the improved (SRM, 0.24), worsened (SRM, −0.68), and unchanged subgroups of patients (SRM, 0.35). Meanwhile, the EQ-HWB LSS performed better and exhibited small to large responsiveness: improved (SRM, 0.29), worsened (SRM, −0.97), and unchanged subgroups (SRM, 0.59) (Table 5). In comparison, other measures also displayed small responsiveness in the improved health subgroup, except for EQ VAS, which exhibited moderate responsiveness (SRM, 0.62). In the worsened health subgroup, 3 other measures also exhibited large responsiveness: FACT-8D (SRM, −1.72), EQ-5D-5L LSS (SRM, −1.13), and EQ VAS (SRM, −1.10). In the unchanged health subgroup, all except EQ-HWB LSS exhibited small responsiveness.

Table 3. Convergent validity results.

Instrument score or item	Correlation with EQ-HWB/EQ-HWB-S item or score	Instrument score or item	Correlation with EQ-HWB/EQ-HWB-S item or score
	EQ-HWB-S index		**EQ-HWB-S sadness or depression**
EQ-5D-5L index	0.73	EQ-5D-5L anxiety/depression	0.40
FACT-8D index	0.70	FACT-8D sad	0.45
SWEMWBS index	0.34		**EQ-HWB discomfort (frequency)**
	EQ-HWB LSS	EQ-5D-5L pain/discomfort	0.41
EQ-5D-5L LSS	0.65	FACT-8D nausea	0.24†
EQ VAS	0.50	FACT-G bothered by side effects	0.33
FACT-G total	0.69	FACT-G feel ill	0.43
WEMWBS total	0.49		**EQ-HWB discomfort (severity)**
	EQ-HWB-S anxiety	EQ-5D-5L pain/discomfort	0.44
EQ-5D-5L anxiety/depression	0.51	FACT-8D nausea	0.20†
FACT-G nervous	0.57	FACT-G bothered by side effects	0.32
FACT-G worry about dying	0.34	FACT-G feel ill	0.45
FACT-8D worry about condition	0.42		**EQ-HWB feel good about self**
	EQ-HWB-S cognition	WEMWBS feeling good	0.31
WEMWBS thinking clearly	0.35		**EQ-HWB frustration**
	EQ-HWB-S day-to-dayactivities	EQ-5D-5L anxiety/depression	0.44
EQ-5D-5L usual activities	0.60		**EQ-HWB pain (frequency)**
FACT-8D work	0.40	EQ-5D-5L pain/discomfort	0.64
	EQ-HWB-S exhaustion	FACT-G pain	0.56
FACT-8D fatigue (lack of energy)	0.43	FACT-G feel ill*	0.51
	EQ-HWB-S getting around		**EQ-HWB personal care**
EQ-5D-5L mobility	0.60	EQ-5D-5L self-care	0.62
	EQ-HWB-S loneliness		**EQ-HWB sleep**
EQ-5D-5L anxiety/depression	0.41	FACT-8D sleeping well	0.66
	EQ-HWB-S pain (severity)		**EQ-HWB unsafe**
EQ-5D-5L pain/discomfort	0.60	EQ-5D-5L anxiety/depression	0.31
FACT-G pain	0.51		
FACT-G feel ill*	0.53		

Note. Correlations are presented in absolute form. Correlations: weak (0.10-0.29), moderate (r = 0.30-0.49), or strong (r = 0.50 and above).
EQ-HWB indicates EQ Health and Wellbeing; EQ-HWB-S, EQ Health and Wellbeing short form; EQ VAS, EuroQol visual analog scale; EQ-5D-5L, 5-level EQ-5D; FACT-G, Functional Assessment of Cancer Therapy – General; FACT-8D, Functional Assessment of Cancer Therapy Eight Dimension; LSS, level sum score; SWEMWBS, Short Warwick-Edinburgh Mental Wellbeing Scale; WEMWBS, Warwick-Edinburgh Mental Wellbeing Scale.
*In the Indonesian language, the word for "ill" is an umbrella term that may also include "pain."
†Weaker than hypothesized.

Discussion

This study is the first to validate the EQ-HWB and EQ-HWB-S in breast cancer and provide evidence on the test-retest reliability and responsiveness of both instruments in a clinical population. It is also the first study providing quantitative evidence of the Indonesian version of the measures. Ceiling effects were exhibited by 7 EQ-HWB items, but not at the instrument level. The construct validity of EQ-HWB and EQ-HWB-S was supported by the high degree of convergence across multiple conceptually similar dimensions with EQ-5D-5L and FACT-8D. In known-group validity, both instrument versions discriminated with large ESs among patients grouped by their EQ VAS scores, general health, and number of symptoms, with the EQ-HWB-S displaying to be just as effective as EQ-HWB. The EQ-HWB and EQ-HWB-S performed favorably in test-retest analysis, showing at least strong reliability for almost 70% of individual items and excellent reliability at the instrument level. Evidence of responsiveness was also observed, with a particularly large effect of the EQ-HWB LSS in patients with worsened health.

The EQ-HWB and EQ-HWB-S also performed well compared with other commonly used measures in cancer—the EQ-5D-5L, FACT-G/FACT-8D, and WEMWBS/SWEMWBS. In cancer outcomes research, the EQ-5D has been shown to have limited sensitivity compared with cancer-specific preference-accompanied instruments.[7,39-41] The absence of ceiling effects in the EQ-HWB and EQ-HWB-S suggests that these instruments may be more sensitive than EQ-5D-5L, likely due to the inclusion of broader aspects and more dimensions. Furthermore, the varying degree of correlations between EQ-HWB and the composite domains of EQ-5D-5L suggests that EQ-HWB may provide better clarity at capturing patient problems.[10-12] Across various cancer types and treatments,

Table 4. Known-group validity of EQ-HWB and EQ-HWB-S.

Groups	n	EQ-HWB-S index		EQ-HWB LSS	
		Mean (SD)	Effect size (95% CI)	Mean (SD)	Effect size (95% CI)
Caregiver use					
Yes	184	0.84 (0.16)	d = 0.04 (−0.19 to 0.27)	83.53 (11.70)	d = 0.00 (−0.23 to 0.23)
No	116	0.84 (0.18)		83.51 (11.90)	
Cancer stage at diagnosis					
1-2	212	0.85 (0.15)	d = 0.21 (−0.04 to 0.46)	83.63 (11.71)	d = 0.04 (−0.21 to 0.29)
3-4	86	0.82 (0.20)		83.21 (12.07)	
EQ VAS score					
80 and above	219	0.89 (0.11)*	d = 1.18 (0.91-1.45)	86.67 (9.63)*	d = 1.10 (0.83-1.37)
<80	81	0.71 (0.23)*		75.02 (12.81)*	
Number of comorbidities					
0	78	0.87 (0.17)	η^2 = 0.031 (0.00-0.08)	86.45 (11.67)	η^2 = 0.030 (0.00-0.07)
1	123	0.85 (0.13)*		83.61 (10.29)	
2+	99	0.80 (0.20)*		81.11 (13.06)	
Number of symptoms					
0	17	0.94 (0.06)*	η^2 = 0.271 (0.19-0.35)	92.18 (8.13)*	η^2 = 0.290 (0.20-0.36)
1-3	71	0.92 (0.08)*		89.56 (8.02)*	
4-6	68	0.88 (0.11)*		86.71 (8.62)*	
7-9	60	0.86 (0.11)*		83.50 (9.81)*	
10+	84	0.70 (0.22)*		74.11 (12.58)*	
General health					
Very good/excellent	53	0.88 (0.13)*	η^2 = 0.154 (0.08-0.23)	86.68 (9.37)*	η^2 = 0.145 (0.08-0.22)
Good	166	0.88 (0.12)*		86.10 (9.88)*	
Poor/fair	81	0.73 (0.22)*		76.19 (13.55)*	

d indicates Cohen's d; EQ-HWB, EQ Health and Wellbeing long form (25 items); EQ-HWB-S, EQ-HWB short form (9 items); EQ VAS, EuroQol visual analog scale; η^2, eta-squared; LSS, level summary score (linearly transformed to 0-100 scale).
*$P < .05$.

Table 5. Test-retest reliability and responsiveness of index values, level summary scores, and EQ VAS.

Instrument score	Test-retest reliability			Responsiveness								
	Unchanged health subgroup (n = 32)			Improved health subgroup (n = 75)			Worsened health subgroup (n = 18)			Unchanged health subgroup (n = 57)		
	Mean (SD)		ICC (95% CI)	Mean (SD)		SRM (95% CI)	Mean (SD)		SRM (95% CI)	Mean (SD)		SRM (95% CI)
	Baseline	Follow-up		Baseline	Follow-up		Baseline	Follow-up		Baseline	Follow-up	
EQ-HWB-S index	0.87 (0.17)	0.88 (0.15)	0.83 (0.68-0.91)	0.88 (0.14)	0.90 (0.10)	0.24 (0.01-0.47)	0.80 (0.15)	0.75 (0.15)	−0.68 (−0.99 to −0.37)	0.83 (0.21)	0.84 (0.18)	0.35 (0.23-0.46)
EQ-5D-5L index	0.87 (0.17)	0.90 (0.15)	0.94 (0.87-0.97)	0.91 (0.14)	0.92 (0.13)	0.16 (−0.03 to 0.34)	0.83 (0.21)	0.70 (0.23)	−0.65 (−0.92 to −0.39)	0.82 (0.22)	0.84 (0.23)	0.31 (0.12-0.49)
FACT-8D index	0.78 (0.21)	0.75 (0.21)	0.77 (0.58-0.88)	0.77 (0.22)	0.81 (0.19)	0.23 (0.01-0.44)	0.64 (0.19)	0.51 (0.18)	−1.72 (−2.33 to −1.10)	0.72 (0.25)	0.72 (0.24)	0.06 (−0.21 to 0.34)
SWEMWBS index	0.90 (0.10)	0.94 (0.07)	0.53 (0.07-0.76)	0.86 (0.15)	0.89 (0.13)	0.29 (0.07-0.52)	0.82 (0.14)	0.79 (0.15)	−0.27 (−0.66 to 0.12)	0.85 (0.15)	0.86 (0.15)	0.24 (0.08-0.40)
EQ-HWB LSS*	86.38 (10.77)	86.28 (9.87)	0.79 (0.61-0.89)	86.32 (10.77)	88.37 (8.55)	0.29 (0.05-0.53)	79.72 (10.87)	74.11 (12.25)	−0.97 (−1.36 to −0.58)	83.14 (13.33)	84.74 (13.11)	0.59 (0.40-0.78)
EQ-5D-5L LSS*	92.19 (11.43)	93.75 (10.16)	0.89 (0.79-0.95)	94.33 (9.02)	95.47 (7.59)	0.22 (−0.10 to 0.53)	89.44 (13.49)	80.83 (14.78)	−1.13 (−1.80 to −0.46)	89.39 (12.36)	90.18 (13.13)	0.18 (−0.11 to 0.46)
EQ VAS	79.38 (20.31)	83.13 (15.75)	0.71 (0.48-0.85)	82.33 (14.41)	88.87 (7.56)	0.62 (0.40-0.84)	76.67 (12.72)	62.50 (14.68)	−1.10 (−1.64 to −0.57)	79.82 (16.85)	80.96 (15.63)	0.18 (−0.03 to 0.40)
FACT-G total score*	79.51 (13.54)	80.77 (13.60)	0.76 (0.56-0.87)	80.74 (12.27)	82.16 (10.44)	0.17 (0.04-0.31)	72.79 (14.06)	66.90 (15.09)	−0.74 (−1.51 to 0.02)	77.46 (15.13)	77.96 (15.50)	0.18 (0.03-0.33)
WEMWBS total score	80.41 (14.71)	85.49 (10.48)	0.50 (0.02-0.75)	77.50 (18.33)	78.98 (16.05)	0.15 (−0.05 to 0.37)	69.74 (15.97)	65.58 (16.85)	−0.55 (−1.53 to 0.42)	75.38 (18.13)	75.72 (18.02)	0.12 (−0.09 to 0.33)

Note. Test-retest reliability analysis was conducted on patients with unchanged health status from group 1, ie, patients in active treatment cycle at baseline who were invited to complete the follow-up questionnaire during their next cycle. Responsiveness analysis was conducted on patients from group 2, ie, patients in their final treatment cycle at baseline who were invited to complete the follow-up at their post-treatment consultation.
EQ-HWB indicates EQ Health and Wellbeing; EQ-HWB-S, EQ Health and Wellbeing short form (9 items); FACT, Functional Assessment of Cancer Therapy; FACT-G, FACT-General; FACT-8D, FACT Eight Dimension; ICC, intraclass correlation coefficient; LSS, level summary score; SRM, standardized response mean; SWEMWBS, Short Warwick-Edinburgh Mental Wellbeing Scale.
*Linearly transformed to a scale of 0-100.

patients often experience diminished HRQoL and well-being, specifically marked by problems in exhaustion, sleep, nausea, interpersonal relationships, and personal appearances, which may not be sufficiently captured by the EQ-5D-5L.[42–44] These HRQoL areas included in the EQ-HWB may be considered candidates for bolt-ons to improve the performance of EQ-5D-5L in (breast) cancer populations.

Before our investigation, one Australian study demonstrated favorable test-retest reliability of EQ-HWB-S in a caregiver population despite the limited sample size.[8] Our findings provide encouraging evidence about the test-retest reliability of both EQ-HWB and EQ-HWB-S. Importantly, the EQ-HWB/EQ-HWB-S showed slightly better test-retest reliability than the FACT-G/FACT-8D, which has been widely used in cancer clinical trials.[45] However, notably, 3 EQ-HWB items, of which 2 also belong to EQ-HWB-S, performed suboptimally: "sleep," "exhaustion," and "anxiety." Similarly, the FACT-8D "sleep" showed the poorest test-retest reliability among the FACT-G/FACT-8D items. The relatively long follow-up interval for the analysis may have influenced the results; however, we anchored this upon the patients' self-reported unchanged health status. In addition, patients reacting differently to treatments (ie, adverse effects) may contribute to the heterogeneity.

Some limitations in this study need to be acknowledged. First, our sample predominantly consisted of less-educated Indonesians with lower economic status, who may tend to rate their health more favorably than wealthier and more educated individuals (ie, response heterogeneity).[46] Furthermore, Asian patients are often less inclined to report health problems, including physical and mental symptoms, compared with their Western counterparts, possibly resulting in better self-reported health.[47,48] Second, the use of the Australian and (pilot) UK value sets for the FACT-8D, EQ-HWB-S, and SWEMWBS may not precisely reflect the preferences of the Indonesian population. Third, the measures in the survey were administered in a fixed order. Although this could potentially introduce bias, previous studies suggest that presentation order is not likely to significantly affect responses or may have only a small effect if present.[49,50] Fourth, the use of GRC as an anchor for responsiveness may not have fully captured the scope of changes experienced by patients, given that the EQ-HWB extends beyond health constructs. Fifth, varying recall periods among the instruments may have influenced our results, although the extent is unclear and needs further investigation.

Conclusions

This study provided psychometric evidence regarding the validity, reliability, and responsiveness of the EQ-HWB and EQ-HWB-S in a breast cancer population. The EQ-HWB and EQ-HWB-S performed comparably with the widely validated EQ-5D-5L, FACT-G, and FACT-8D, where the domains of exhaustion, pain, discomfort, and sleep may be particularly relevant in our sample. Our findings support the potential usefulness of EQ-HWB and EQ-HWB-S as patient-reported outcome measures for clinical and economic purposes in cancer populations, including their role in health technology assessments for breast cancer treatments.

Author Disclosures

Drs Purba, Mukuria, and Rencz are members of the EuroQol Group. Views expressed in the article are those of the authors and are not necessarily those of the EuroQol Research Foundation. Author disclosure forms can be accessed below in the Supplemental Material section.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2024.12.003.

Article and Author Information

Accepted for Publication: December 6, 2024

Published Online: January 24, 2025

doi: https://doi.org/10.1016/j.jval.2024.12.003

Author Affiliations: Department of Health Policy (Pangestu, Rencz) and Doctoral School of Business and Management (Pangestu), Corvinus University of Budapest, Budapest, Hungary; Faculty of Economics and Business, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia (Pangestu); Department of Psychology, Faculty of Psychology (Purba, Setyowibowo), and Department of Surgery, Oncology Division, Hasan Sadikin General Hospital (Azhar), Universitas Padjadjaran, Bandung, Indonesia; School of Medicine and Population Health, University of Sheffield, Sheffield, England, UK (Mukuria).

Correspondence: Fredrick Dermawan Purba, PhD, Department of Psychology, Faculty of Psychology, Universitas Padjadjaran, Bandung, Indonesia. Email: fredrick.purba@unpad.ac.id

Authorship Confirmation: All authors certify that they meet the ICMJE criteria for authorship.

Funding/Support: This study was funded by the EuroQol Research Foundation (1644-RA).

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 express sincere gratitude to all the patients who participated in this study. They also extend their thanks to the Oncology Division of Hasan Sadikin General Hospital for granting them the opportunity to conduct the research. Furthermore, they acknowledge the invaluable support provided by the team of research assistants from Universitas Padjadjaran.

REFERENCES

1. Brazier J, Peasgood T, Mukuria C, et al. The EQ-HWB: overview of the development of a measure of health and wellbeing and key results. *Value Health*. 2022;25(4):482–491.
2. Peasgood T, Mukuria C, Carlton J, et al. What is the best approach to adopt for identifying the domains for a new measure of health, social care and carer-related quality of life to measure quality-adjusted life years? Application to the development of the EQ-HWB? *Eur J Health Econ*. 2021;22(7):1067–1081.
3. Mukuria C, Connell J, Carlton J, et al. 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*. 2022;25(4):492–511.
4. Mukuria C, Peasgood T, McDool E, Norman R, Rowen D, Brazier J. Valuing the EQ health and wellbeing short using time trade-off and a discrete choice experiment: a feasibility study. *Value Health*. 2023;26(7):1073–1084.
5. Peasgood T, Mukuria C, Brazier J, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ-HWB. *Value Health*. 2022;25(4):525–533.
6. Engel L, Kosowicz L, Bogatyreva E, et al. Face Validity of four preference-weighted quality-of-life measures in residential aged care: a think-aloud study. *Patient*. 2023;16(6):655–666.
7. Gibson AEJ, Longworth L, Bennett B, Pickard AS, Shaw JW. Assessing the content validity of preference-based measures in cancer. *Value Health*. 2024;27(1):70–78.
8. Bailey C, Dalziel K, Constable L, et al. The performance of the EQ-HWB-S as a measure of quality-of-life of caregivers in families that have experienced adverse events. Published online *Eur J Health Econ*. April 5, 2024. https://doi.org/10.1007/s10198-024-01688-w.
9. Kuharic M, Mulhern B, Sharp LK, Turpin RS, Pickard AS. Comparison of EQ health and well-being long and short with other preference-based measures among United States informal caregivers. Published online *Value Health*. March 14, 2024. https://doi.org/10.1016/j.jval.2024.03.003.
10. Kuharić M, Pickard AS, Mukuria C, Finch AP. The measurement properties of the EQ health and well-being and EQ health and well-being short form in

italian population: a comparative study with EQ-5D-5L. Published online *Value Health*. March 14, 2024. https://doi.org/10.1016/j.jval.2024.03.002.
11. Lee P, Engel L, Lubetkin E, Gao L. Exploring the comparability between EQ-5D and the EQ health and wellbeing in the general australian population. *Value Health*. 2024;27(4):508–517.
12. Long C, Mao Z, Yang Z. A head-to-head comparison of EQ health and well-being and EQ-5D-5L in patients, carers, and general public in China. Published online *Value Health*. March 5, 2024. https://doi.org/10.1016/j.jval.2024.02.012.
13. Thai TM, Engel L. *Head-to-head comparison of the EQ-HWB-S, ReQoL, EQ-5D-5Lmeasure in Australian general population and mental health wellbeing*. Presented at:. Copenhagen, Denmark: The 4th EuroQol Early Career Researchers Meeting; March 5, 2024.
14. Bailey C, Dalziel K, Jones R, Hiscock H, Devlin NJ, Peasgood T. The Validity of the EuroQol health and wellbeing short version (EQ-HWB-S) instrument in parents of children with and without health conditions. Published online *Pharmacoeconomics*. January 18, 2024. https://doi.org/10.1007/s40273-024-01351-5.
15. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249.
16. Landrigan PJ, Fuller R, Acosta NJR, et al. The Lancet Commission on pollution and health. *Lancet*. 2018;391(10119):462–512.
17. Pangestu S, Purba FD, Setyowibowo H, Mukuria C, Rencz F. *Examining the associations between objective and subjective financial toxicity and EQ-5D-5L Outcomes in Patients With Breast Cancer*. presented at:. Copenhagen, Denmark: The 4th EuroQol Early Career Researchers Meeting 2024; March 5, 2024 2024.
18. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. Dec 2011;20(10):1727–1736.
19. Purba FD, Hunfeld JAM, Iskandarsyah A, et al. The Indonesian EQ-5D-5L value set. *Pharmacoeconomics*. 2017;35(11):1153–1165.
20. Cella DF, Tulsky DS, Gray G, et al. The functional assessment of cancer therapy scale: development and validation of the general measure. *J Clin Oncol*. 1993;11(3):570–579.
21. Fairclough DL, Cella DF. Functional Assessment of Cancer Therapy (FACT-G): non-response to individual questions. *Qual Life Res*. 1996;5(3):321–329.
22. Cheung YB, Daniel R, Ng GY. Response and non-response to a quality-of-life question on sexual life: a case study of the simple mean imputation method. *Qual Life Res*. 2006;15(9):1493–1501.
23. King MT, Norman R, Mercieca-Bebber R, et al. The Functional Assessment of Cancer Therapy Eight Dimension (FACT-8D), a Multi-Attribute Utility Instrument Derived From the Cancer-Specific FACT-General (FACT-G) Quality of Life Questionnaire: Development and Australian Value Set. *Value Health*. Jun 2021;24(6):862–873.
24. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes*. 2007;5:63.
25. Yiu HHE, Buckell J, Petrou S, Stewart-Brown S, Madan J. Derivation of a UK preference-based value set for the Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS) to allow estimation of Mental Well-being Adjusted Life Years (MWALYs). *Soc Sci Med*. 2023;327:115928. https://doi.org/10.1016/j.socscimed.2023.115928.
26. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther*. 2009;17(3):163–170.
27. Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717–1727.
28. McDool E, Mukuria C, Brazier J. A comparison of the SF-6Dv2 and SF-6D UK utility values in a mixed patient and healthy population. *Pharmacoeconomics*. 2021;39(8):929–940.
29. Rencz F, Brodszky V, Janssen MF. A direct comparison of the measurement properties of EQ-5D-5L, PROMIS-29+2 and PROMIS Global Health Instruments and EQ-5D-5L and PROPr Utilities in a General Population Sample. *Value Health*. 2023;26(7):1045–1056.
30. Nikl A, Janssen MF, Brodszky V, Rencz F. A head-to-head comparison of the EQ-5D-5L and 15D descriptive systems and index values in a general population sample. *Health Qual Life Outcomes*. 2023;21(1):17.
31. McDool E, Mukuria C, Peasgood T. Psychometric performance of the EQ Health and wellbeing short in a United Kingdom population sample. Published *Value Health*. May 23, 2024. https://doi.org/10.1016/j.jval.2024.05.007.
32. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.
33. Cohen J. Quantitative methods in psychology: a power primer. *Psychol Bull*. 1992;112:1155–1159.
34. Richardson JT. Eta squared and partial eta squared as measures of effect size in educational research. *Educ Res Rev*. 2011;6(2):135–147.
35. Gwet KL. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*. Advanced Analytics, LLC; 2014.
36. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. Mar 1977;33(1):159–174.
37. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
38. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*. 1994;6(4):284.
39. Davies A, Waylen A, Leary S, et al. Assessing the validity of EQ-5D-5L in people with head & neck cancer: does a generic quality of life measure perform as well as a disease-specific measure in a patient population? *Oral Oncol*. 2020;101:104504.
40. Lin FJ, Longworth L, Pickard AS. Evaluation of content on EQ-5D as compared to disease-specific utility measures. *Qual Life Res*. May 2013;22(4):853–874.
41. van Dongen-Leunis A, Redekop WK, Uyl-de Groot CA. Which questionnaire should be used to measure quality-of-life utilities in patients with acute leukemia? An evaluation of the validity and interpretability of the EQ-5D-5L and preference-based questionnaires derived from the EORTC QLQ-C30. *Value Health*. 2016;19(6):834–843.
42. Bottomley A, Flechtner H, Efficace F, et al. Health related quality of life outcomes in cancer clinical trials. *Eur J Cancer*. 2005;41(12):1697–1709.
43. Heidary Z, Ghaemi M, Hossein Rashidi B, Kohandel Gargari O, Montazeri A. Quality of life in breast cancer patients: a systematic review of the qualitative studies. *Cancer Control*. 2023;30:10732748231168318.
44. Hernández Alava M, Pudney SE, Wailoo AJ. Does EQ-5D Tell the whole story? statistical methods for comparing the thematic coverage of clinical and generic outcome measures, with application to breast cancer. *Value Health*. 2023;26(9):1398–1404.
45. Giesinger JM, Efficace F, Aaronson N, et al. Past and current practice of patient-reported outcome measurement in randomized cancer clinical trials: a systematic review. *Value Health*. 2021;24(4):585–591.
46. Hanandita W, Tampubolon G. Does reporting behaviour bias the measurement of social inequalities in self-rated health in Indonesia? An anchoring vignette analysis. *Quality of Life Research*. 2016;25(5):1137–1149.
47. Chen CH, Tang ST, Chen CH. Meta-analysis of cultural differences in Western and Asian patient-perceived barriers to managing cancer pain. *Palliat Med*. 2012;26(3):206–221.
48. Birtel MD, Mitchell BL. Cross-cultural differences in depression between White British and South Asians: Causal attributions, stigma by association, discriminatory potential. *Psychol Psychother*. 2023;96(1):101–116.
49. Cheung YB, Wong LC, Tay MH, et al. Order effects in the assessment of quality of life in cancer patients. *Quality of Life Research*. 2004;13(7):1217–1223.
50. Childs AL. Effect of order of administration of health-related quality of life interview instruments on responses. *Quality of Life Research*. 2005;14(2):493–500.