BMC Psychiatry

RESEARCH

Open Access

Psychometric properties of Japanese version of the Recovering Quality of Life (ReQoL)

Ayako Hada[1,2], Takayuki Kawaguchi[1], Kaori Usui[1], Mai Iwanaga[1], Sayaka Sato[1], Sosei Yamaguchi[1*], Naoaki Kuroda[3,4] and Chiyo Fujii[1]

Abstract

Purpose This study aims to examine the psychometric properties of the Recovering Quality of Life– Japanese version (ReQoL-J) for use in the community setting.

Methods This cross-sectional study was conducted among service users of community mental health service facilities. A series of confirmatory factor analyses (CFAs) was performed. After the best-fitting model was identified, omega indices, and intraclass correlation (ICC) were calculated to evaluate dimensionality and reliability. The correlation coefficients between the ReQoL-J scores and the other scales were calculated for testing construct validity. We also estimated the correlation coefficients between the ReQoL-J-10 and −20.

Results A total of 395 service users from 23 facilities in Japan participated in this study. Of these, 66 responded to the test-retest reliability. The bifactor model showed the best fit to the data: $\chi^2/df = 1.74$, CFI = 0.993, RMSEA = 0.031, and SRMR = 0.038. ECV and *OmegaH* indicated unidimensionality. ICC for the Re-QoL-J-10 and −20 were 0.700 and 0.766, respectively. The ReQoL-J-10 showed a high correlation with the ReQoL-J-20 ($r = .95$, $p < .001$). The EQ-5D-5 L, a new brief scale for subjective personal agency (SPA-5), well-being and the quality of life scale (QLS) scores were positively correlated with the ReQoL-J score. On the other hand, the patient health questionnaire-9 (PHQ-9) and the WHODAS scores were negatively correlated with the ReQoL-J score.

Conclusion Factor structure, internal consistency, test-retest reliability, and construct validity of the ReQoL-J were confirmed. The ReQoL-J is a promising tool as a patient-reported outcome measure in mental health services for the Japanese community setting.

Keywords Patient-reported outcome, Factor structure, Psychometric properties, Recovering quality of life

*Correspondence:
Sosei Yamaguchi
sosei.yama@ncnp.go.jp
[1]Department of Community Mental Health and Law, National Institute of Mental Health, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Kodaira, Tokyo 187-8553, Japan
[2]Department of Mental Health and Psychiatric Nursing, Institute of Science Tokyo, Bunkyo-ku, Tokyo 113-8510, Japan
[3]Department of Public Mental Health Research, National Institute of Mental Health, National Center of Neurology and Psychiatry, Kodaira, Tokyo 187-8553, Japan
[4]Health Services Research and Development Center, University of Tsukuba, 1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8575, Japan

Introduction

Patient-reported outcome measures (PROMs) have become essential to assess the outcome of mental health services over the past two decades. When objective outcomes are difficult to obtain, PROMs are useful [1]. Indeed, PROMs are often poorly related to physiological and other objective outcomes such as performance-based outcomes, clinician-reported outcomes, or biomarkers [2–5]. Also, patients' feedback improves professional services [6–7]. Thus, PROMs are necessary outcome measures to evaluate services rather than an alternative to objective outcomes. Psychometrically valid and reliable PROMs are needed particularly in the mental health field [8–9].

Historically, one of the central interests in PROMs has been how to measure health-related quality of life in people with physical or psychiatric illnesses. The concept of quality of life (QoL) itself is defined as individuals' perceptions of their position in life in the context of the culture and value systems in which they live and concerning their goals, expectations, standards, and concerns [10]. The similarities between the concepts of QoL and well-being have been discussed for a decade [11–12]. On the other hand, depressed mood or disabilities are concepts that oppose the QoL concept [13–14]. In addition, previous studies reported that several variables are associated with QoL among people with mental illness such as psychiatric symptoms, factional impairment, and personal agency [15–18]. In short, the concept of QoL and relevant variables have been gradually shaped in the challenging process of measuring QoL and developing its scales.

There are both self-reported and observed reported QoL scales. In people with mental illness, they have been not strongly correlated with each other [19–22]. In this context, while numerous researchers have developed self-reported QoL scales [23–27], their usefulness has been often insufficient. For example, the EuroQol 5 Dimension (EQ-5D) and Short-form 6 dimension (SF-6D) are some of the most popular QoL scales in health research; however, these scales have contents too generic to measure QoL among people with mental illness [28–29]. Brazier et al. also examined the validity and responsiveness of several QoL measures and found limited evidence, particularly in people with schizophrenia and bipolar disorders. These results may potentially be attributed to the item development process through the literature reviews and experts' views [28]. In other words, a QoL measure reflecting people with mental illness is required to accurately assess their subjective views on their own life.

In this context, the Recovering Quality of Life questionnaire (ReQoL) with 20 items and 10 items as a short version was developed to measure "recovery-focused quality of life" for mental health service users who were experiencing mental health difficulties [30–32]. The ReQoL has two clear strengths. First, the items of the ReQoL were developed through collaboration with multiple stakeholders including people with mental illness and detailed interviews with over 70 people with mental illness [31]. Second, this scale reflects important contents of personal recovery which originated in the service users' grassroots movement [33–35], since the concepts of QoL and personal recovery had theoretically overlapped each other [36–37]. For example, the framework for personal recovery consists of several facets: connectedness, hope and optimism about the future, identity, meaning in life, and empowerment [38]. Although personal recovery is likely to highlight positive aspects, the relevant studies also pointed out several negative aspects (e.g., absence of depression or anxiety) [15, 39–40]. Therefore, the ReQoL includes both domains in terms of enhancing QoL and reducing QoL [41].

Several studies have measured the validity and reliability of ReQoL. For example, three studies examined the factor structure of the ReQoL through confirmatory factor analysis (CFA) (see Table 1). A bifactor model consisting of two specific group factors (Negative and Positive affects) and the Global factor was identified as the best-fitted model to the data [30–31, 41–42]. The factor loadings on the global (general) factor were substantially larger than the loadings on the negative and positive ones (specific group factors). This finding indicated that the ReQoL was an essentially unidimensional model. In addition to the factorial validity, the internal consistency, test-retest reliability, and construct validity were also confirmed in several countries. Cronbach's alphas using a sample of service users for the ReQoL-10 and ReQoL-20 items were reported to be 0.92 and 0.96, respectively [31]. ICCs for patient and general population samples were 0.90 and 0.87, respectively [31]. In terms of convergent validity, both ReQoL-10 and − 20 showed high correlations with measures such as the

Table 1 ReQoL-20 CFA studies

References		*N*	Country	Model	Goodness of Fit	
					CFI	RMSEA
Keetharuth et al. (2019)	Study 1	2262	UK	bifactor model: global, negative affects, and positive affects	0.954	0.061
	Study 2	4266	UK	bifactor model: global, negative affects, and positive affects	0.971	0.066
Xu et al. (2021)		500	HK	bifactor model: global, positively worded, and negatively worded	0.991	0.029

Note. CFI, Comparative Fit Index; RMSEA, Root Mean Square Error of Approximation

Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS), Clinical Outcomes in Routine Evaluation (CORE)-10 [31], and EQ-5D [42]. The Recovering Quality of Life-Utility Index (ReQoL-UI) which was developed based on the ReQoL-20, showed moderate correlation with the EQ-5D-5 L in the general population of Hong Kong [43]. In addition, recovery from severe mental illness involves developing a sense of control over one's illness through personal agency (i.e., what people can do on their own) [44], is facilitated by personal agency [45–46]. Therefore, the Five-item Subjective and Personal Agency scale (SPA-5: [47]) may be highly correlated with ReQoL Based on global evidence of its psychometric properties, ReQoL has consequently become a recommended scale as one of the standard outcome measures in the mental health field [48].

Despite the worldwide spread of the ReQoL, no studies examined the validity and reliability of the ReQoL in a Japanese setting. For example, while the factor structure found in the United Kingdom (UK) and Hong Kong may be acceptable to Japanese cultural backgrounds, and the bifactor model can be a candidate for the best-fitting model to the Japanese sample, this hypothesis has not yet been addressed. Once the dimensional structure of a set of items has been validated, the resulting sum scores are often considered to have a robust psychometric basis [49]. However, the original instrument, developed and validated in its original cultural context, may not necessarily exhibit the same degree of robustness when used in a different cultural setting. For the use of the ReQoL in the Japanese population, its psychometric properties should be assessed. This study aims to examine the psychometric properties of the ReQoL, specifically including factorial validity, dimensionality and reliability, concurrent validity, and convergent validity among people with mental illness in the Japanese community setting.

Methods

Overall design

This study employed a cross-sectional design. A convenient sampling method was used to collect the data via community mental health service facilities such as day-care centres, outpatient clinics community, and employment support agencies in Japan. Among 27 facilities we recruited, a total of 23 facilities participated in this study (Fig. 1). All potential participants were recruited by the primary staff of each facility, and those meeting inclusion criteria were enrolled in the study. All responses were anonymized and managed with a research ID at each facility. Researchers merged responses of T1 and T2 with research IDs for test-retest reliability. The data were collected from August 1 to November 30, 2023. This study was approved by the Research Ethics Committee of the National Centre for Neurology and Psychiatry (no. A2023-065).

Participants and recruitment

We set the following inclusion criteria of this study: service users (a) who used services from August 1 to November 30, 2023, (b) who were diagnosed with mental disorders as their primary diagnosis, (c) who were under 65 years of age, and (d) who understand the explanation from staff members and voluntarily consented to participate in this research. Exclusion criteria were those (a) who were diagnosed with dementia or an intellectual disability, (b) who were not fluent in Japanese, or (c) who used the service briefly (less than one month). All potential participants were informed by a primary staff member about the aims of the study, the affiliation of the principal investigator, the security of personal information, and ethical considerations for participation in the study. The participants who voluntarily agreed to participate in this study by ticking the consent box on the research document completed their questionnaire. The primary care staff (non-family members who provide and manage care directly) assessed the patient's sociodemographic characteristics and health conditions such as day function, using the WHODAS and QLS include staff-reported versions, simultaneously (T1). The participants in two facilities completed the ReQoL twice with an interval of two weeks (T2) for test-retest reliability.

Main measurement

The ReQoL was originally developed by Keetharuth et al. in the UK (Use of the ReQoL can only be made under licence and that a licence can be discussed by contactig healthoutcomes@innovation.ox.uk.) [31]. The Clinical Outcomes team at Oxford University Innovation (OUI) translated it into Japanese (ReQoL-J). Detailed information on the translation process is available on the OUI website (https://innovation.ox.ac.uk/clinical-outcomes/services/translation-linguistic-validation/). Its t ranslation process included the following 9 steps: (a) concept elaboration, (b) two forward translations, (c) forward translation reconciliation, (d) two back-translations, (e) back-translation review, (f) developer review, (g) independent proofreading, (h) cognitive debriefing interviews, and (i) professional formatting, which was consistent with the process recommended by Eremenco et al. [50]. The ReQoL-J-20 consists of 20 items and one additional physical health item. Of the 20 items, 11 items are positively worded and 9 items are negatively worded. All questions are scored with a five-point Likert scale ranging from "None of the time" to "Most or all of the time." The positively worded questions are scored from 0 to 4. The scores for negatively worded questions are reversed, scored from 4 to 0. The ReQoL-J-20 score is

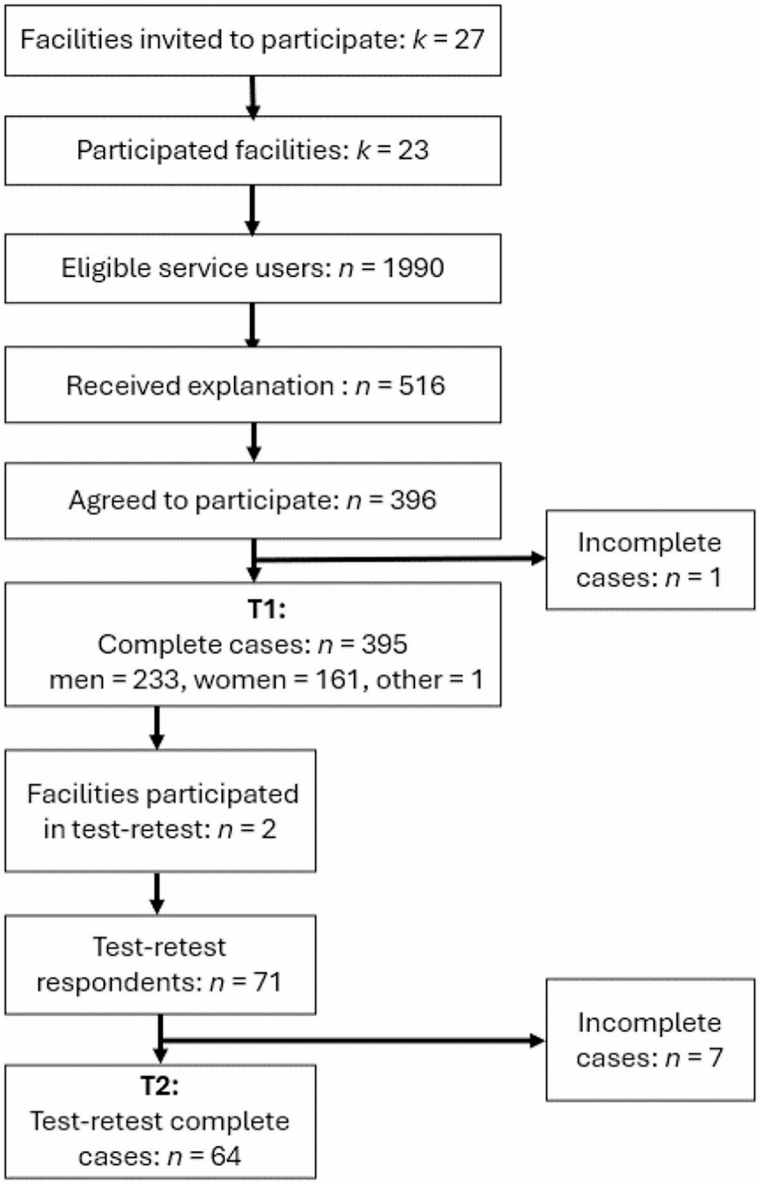

Fig. 1 Participant recruitment flowchart

calculated by summing the scores of all items except for the physical health item, and ranges from 0 to 80. The ReQoL-J-10 (11 items) comprises the first 10 items of the ReQoL-J-20 and the physical health item. The ReQoL-J-10 score is calculated by summing the scores of the first 10 questions of the ReQoL-J-20. The ReQoL-J-10 ranges from 0 to 40. Higher scores of both ReQoL-J-20 and 10 indicate a better quality of life. Following the ReQoL-20 scoring guide, missing data were imputed using mean imputation.

Other measurements for concurrent validity and convergent validity

We used the EuroQol 5 dimension 5 level (EQ-5D-5 L) to examine the concurrent validity and for comparison to a gold standard measure. To test the convergent validity which compares against validated measures of the proximate concepts, we also used the four PROMs and two observer-rated scales including the Five-item Subjective and Personal Agency scale (SPA-5), Single-item well-being scale (SiWB), the Patient Health Questionnaire 9-item version (PHQ-9), the 12-item version of World Health Organization Disability Assessment Schedule 2.0 (WHO-DAS), and Quality of Life Scale (QLS).

With regards to the PROMs, the EQ-5D-5 L consists of five items for measuring quality of life (QOL): mobility,

self-care, usual activities, pain and discomfort, and anxiety and depression. We calculated the scores according to the scoring method of the Japanese version of the EQ-5D-5 L [51]. A higher score indicates a better quality of life.

The SPA-5 was developed to measure subjective personal agency based on the view of service users and in collaboration with people with schizophrenia [47]. Each of the five items of SPA-5 is rated on a five-point Likert scale using a range of 1 (strongly disagree) to 5 (strongly agree). A higher score indicates higher subjective personal agency.

The SiWB has a simple question for own life "Overall, how satisfied are you with life as a whole these days?" with a 0–10 response scale where 0 is "Not at all" and 10 is "Completely" [52]. This item was recommended through an interdisciplinary workshop on the measurement of well-being.

The PHQ-9 is a self-reported nine-item depression scale based on the criteria of a major depressive episode of DSM-IV developed by Spitzer et al. [53]. Each item checks the frequency of depressive symptoms over the previous two weeks with a 4-point Likert scale from 0 to 3. The psychometric properties' validity and reliability of the Japanese version of PHQ-9 were reported [54–55]. A higher score indicates severe depression.

The WHO-DAS is a widely used tool for assessing disability and functional impairment [56–57]. The WHO-DAS 2.0 is administered in the form of a 12-item or 36-item questionnaire covering six domains of disability in the 30-day period prior to the assessment. This scale is based on the International Classification of Functioning, Disability and Health conceptual framework [58]. Total scores were calculated as the sum of scores of all items. Total and domain scores for both versions were converted to a range of 0 to 100. A higher score indicates severe disability. Since a previous study reported a strong positive correlation between the scores of WHODAS-12 and the WHODAS-36 among mental health service users [59], we used the WHODAS-12, which is easier to use, in this study. The WHODAS was measured by not only users (subjective measure) but also primary care staff (objective measure).

The QLS is an observer-rated scale to measure QOL based on clinicians' views. The QLS is a 21-item scale rated from a semi-structured interview providing information on symptoms and functioning during the preceding 4 weeks [60]. The QLS consists of the following 5 scales: interpersonal relationships, instrumental roles, intrapsychic foundations, common objects and activities, and total score. A higher score indicates better objective QOL [61]. The QLS was assessed by the primary care staff in this study.

Statistical analysis

Descriptive statistics and factorial validity

First, we calculated mean, SD, skewness, and kurtosis, to examine distribution of data. Frequencies for the ReQoL-20 and ReQoL-10 were also plotted. Then, a series of CFAs using all samples was conducted. A single-factor and theoretical 2-factor model were fitted to the data. Finally, the theoretical bifactor model which was the same as the factor structure in the literature [30–31, 41–42] was fitted to the data. These models were compared in terms of goodness-of-fit indices. We used weighted least squares means and variance adjusted (WLSMV) as an estimator. A good fit was defined as chi-squared divided by degrees of freedom (χ^2/df) < 2, comparative fit index (CFI) > 0.97, and root mean square error appropriation (RMSEA) < 0.05. An acceptable fit was defined as χ^2/df < 3, CFI > 0.95, and RMSEA < 0.08 [62 – 63]. Standardised root mean square residual (SRMR) is the square root of the difference between the residuals of the sample covariance matrix and the hypothesised model. A value less than 0.08 is considered a good fit [64]. In the terms of goodness of fit, chi-square, RMSEA, CFI and SRMR of a model should be reported [65].

Dimensionality and reliability

After CFA, the internal consistency of the model was calculated using the omega coefficient. The omega coefficient is a preferred index of the internal consistency of a psychological measure when the scale is made up of more than one factor [66–68]. The assumptions for the alpha coefficient are (a) one-factor structure (congeneric model), (b) factor loadings that are approximately the same for all observed variables (tau-equivalent model), and (c) no correlations among the error variables for the observed variables (parallel model). Alpha coefficients for psychological scales are biased when these assumptions are not met, which is commonly the case in practice [66]. Furthermore, omega indices are useful for evaluation of unidimensionality. Therefore, we adopted Omega as reliability indices. Omega indicates a proportion of variance of all the items explained by all the factors. Omega subscale (*OmegaS*) indicates the proportion of the variance among the items of each specific group factor that is explained by both the global factor and the specific group factor. Omega hierarchical (*OmegaH*) indicates the percentage of the variance of the whole measurement explained by the general factor and all the specific group factors explained only by the general factor. Similarly, the percentage of the variance among items of each specific group factor explained by the group factor is also called omega hierarchical subscale (*OmegaHS*) [69–70]. Explained common variance (ECV) and a percentage of uncontaminated correlations (PUC) were also calculated to examine for unidimensionality. When PUC is greater

than 0.80, or when ECV is greater than 0.60 and *OmegaH* is greater than 0.70, the relative bias from using a unidimensional measurement model rather than a bifactor measurement model is likely to be small [71]. Stucky & Edelen suggested that an item set is sufficiently indicative of unidimensionality when ECV values are 0.85 or higher [72].

To examine test-retest reliability, we calculated intraclass correlation (ICC [1, 2]) with a two-way random effects model absolute agreement between T1 and T2. McGraw and Wong defined 10 types of ICC [73]. Of those, the ICC (2,1) indicates variance between patients divided by variance between patients plus variance between measurements plus error variance. The ICC ranges from 0 to 1, where ICC = 1 represents perfect reliability of the measurement and between 0.5 and 0.7, and between 0.75 and 0.9 indicate moderate and good reliability, respectively [74].

Table 2 Demographics of the participants

	Mean	SD
Age	43.3	11.9
Number of hospitalisations	0.25	0.70
	n	%
Gender		
Male	233	59.0
Female	161	40.8
Other	1	0.3
Marital status		
Married/with partner	326	82.5
Unmarried/divorced/widowed	69	17.5
Education		
Secondary school	41	10.4
High school	172	43.5
Vocational school	50	12.7
Junior college	17	4.3
University (undergraduate degree)	111	28.1
University (graduate degree)	4	1.0
Main diagnosis		
F0: Organic mental disorder	5	1.3
F1: Substance abuse	22	5.6
F2: Schizophrenia	163	41.4
F3: Major depression	65	16.4
F3: Bipolar disorder	44	11.1
F4: Neurotic, stress-related and somatoform disorder	42	10.6
F5: Behavioural syndrome associated with physiological disturbances and physical factors	3	0.8
F6: Disorder of adult personality and behaviour	2	0.5
F7: Mental retardation	3	0.8
F8: Disorder of psychological development	37	9.3
F9: Behavioural and emotional disorder with onset usually occurring in childhood and adolescence	9	2.3
Others	0	0.0
Living situation		
Living with family members/partner	231	58.5
Living alone	127	32.2
Residential facility	37	9.4

Concurrent validity and convergent validity

To test performance of the ReQoL-10, Pearson's correlation coefficients between the ReQoL-10 and ReQoL-20 scores were computed. The ReQoL-10 and ReQoL-20 scores were correlated with the EQ-5D-5 L to examine the concurrent validity, and with the SPA-5, the SiWB, the PHQ-9, the WHO-DAS, and the QLS to examine convergent validity of the ReQoL-J. It was hypothesised that the ReQoL-10 and 20 would show a positive correlation with the EQ-5D-5 L, the SPA-5, the SiWB, and the QLS, and a negative correlation with the PHQ-9 and the WHO-DAS. Statistical significance for these analyses was established at p-value < 0.001 due to multiple comparisons.

Analyses were conducted using R statistical software. We used the "lavaan" package (v0.6-17; [75]) for CFA and "irr" package (v0.84.1; [76]) for ICC calculation. Omega indices were calculated with the "Bifactor Indices Calculator" [77].

Results

Sample characteristics

Of 1990 eligible service users from 23 facilities, 516 were informed about this study by research team members or research collaborators of each facility. Of these, 396 agreed to participate in this study and 395 participants completed the questionnaire (Fig. 1). Two hundred and thirty-three participants were male, 161 were female, and mean age was 43.3 years (SD = 11.9). The most frequent diagnoses were schizophrenia [F2] (n = 163), major depression [F3] (n = 65), and bipolar disorder [F3] (n = 44). Other demographics of the participants were shown in Table 2.

Factorial validity

Mean, SD, skewness, and kurtosis for the ReQoL-J are shown in Table 3. All skewness and kurtosis ranged from -1.15 to 0.68. Histograms for the ReQoL-J-20 and ReQoL-J-10 are shown in Figs. 2 and 3. The mean scores (SD) for the ReQoL-J-20 and the ReQoL-J-10 were 43.99 (14.88) and 22.16 (8.09), respectively, and no sum score of scales showed ceiling or floor effect. Table 4 shows the results of CFAs and model comparisons. The goodness of fit of the model was significantly better for the theoretical 2-factor model than the single-factor model. Improvement was also noted from the theoretical 2-factor to theoretical bifactor model, the absolute value of goodness of fit was goodfit: $\chi^2/df = 1.76$, CFI = 0.993, RMSEA = 0.031, and SRMR = 0.038. Factor loadings of the bifactor model are shown in Table 5. Factor loadings of item-2, -4, -5, -6, -7, -8, -10, -11, -15 and – 19 on the General factor were

Table 3 Mean, SD, skewness, kurtosis ($N = 395$)

Item No.	*n*	Mean	SD	Skewness	Kurtosis
1	395	2.50	1.13	-0.42	-0.59
2	395	2.37	1.21	-0.25	-0.97
3	395	2.21	1.16	-0.20	-0.81
4	395	2.27	1.15	-0.13	-0.95
5	395	2.11	1.24	-0.02	-1.01
6	395	2.71	1.33	-0.78	-0.61
7	395	2.33	1.19	-0.15	-0.96
8	395	1.74	1.30	0.28	-1.08
9	395	2.26	1.28	-0.32	-0.98
10	395	1.66	1.19	0.33	-0.81
11	395	2.00	1.24	0.04	-1.02
12	395	2.38	1.18	-0.30	-0.89
13	395	2.28	1.18	-0.29	-0.88
14	395	2.19	1.36	-0.30	-1.14
15	395	1.37	1.22	0.68	-0.50
16	395	2.78	1.27	-0.75	-0.58
17	395	2.06	1.31	-0.15	-1.15
18	395	2.21	1.35	-0.27	-1.12
19	395	2.28	1.15	-0.02	-0.98
20	395	2.27	1.16	-0.29	-0.80
21	392	2.76	1.04	-0.58	-0.28
ReQOL-20 total	395	43.99	14.88	-0.14	-0.38
ReQOL-10 total	395	22.16	8.09	-0.14	-0.38

Note. Use of the ReQoL can only be made under licence and that a licence can be discussed by contacting healthoutcomes@innovation.ox.ac.uk. ReQoL™ Version 1.1 © Copyright, The University of Sheffield 2016, 2018. All Rights Reserved. The authors have asserted their moral rights. Oxford University Innovation Limited is exclusively licensed to grant permissions to use the ReQoL™

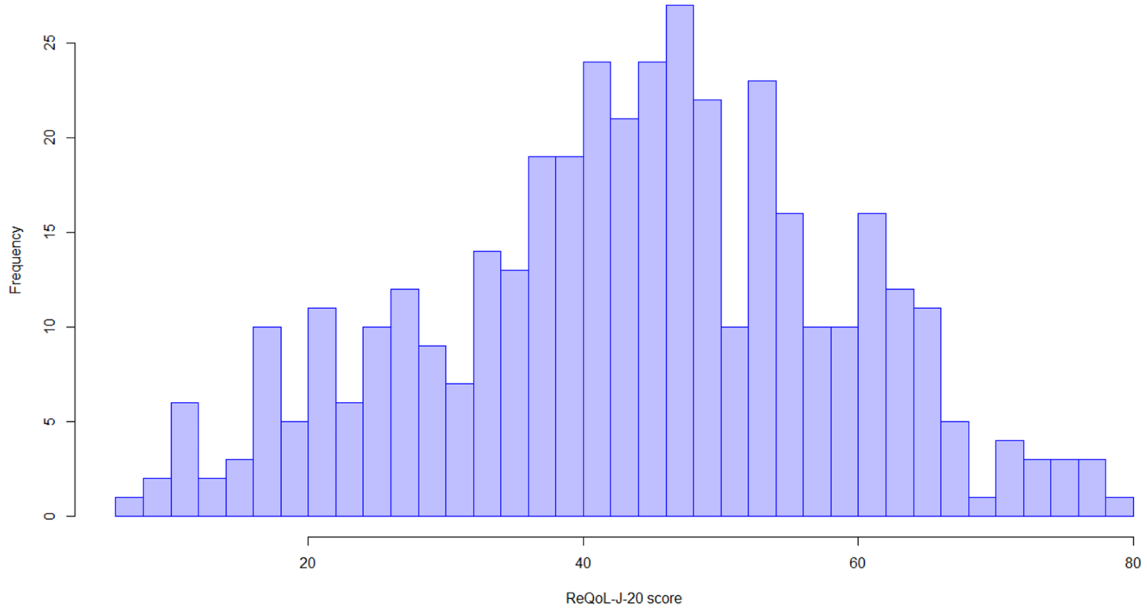

Fig. 2 Histogram of the ReQoL-J-20 score. X-axis indicates total score for Re-QoL-J-20. Y-axis indicates frequency

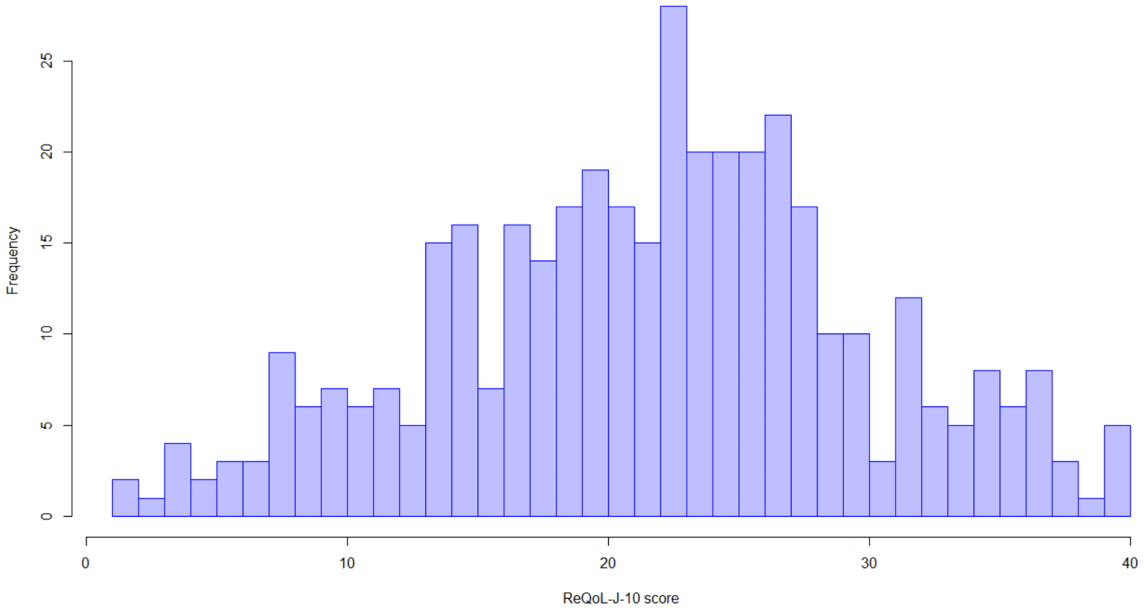

Fig. 3 Histogram of the ReQoL-J-10 score. X-axis indicates total score for Re-QoL-J-10. Y-axis indicates frequency

Table 4 Model comparison of a single factor, theoretical 2-factor, and theoritical bifactor model (*N* = 395)

Model	χ^2	df	χ^2/df	$\Delta\chi^2$ (df)	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR	ΔSRMR
Single-factor model	946.516	170	5.56	Ref	0.920	Ref	0.100	Ref	0.100	Ref
Theoretical 2-factor model	332.188	169	1.97	614.328 (1)*	0.985	0.290	0.043	0.057	0.052	0.052
Theoretical bifactor model	261.944	149	1.76	70.244 (20)*	0.993	0.008	0.031	0.012	0.038	0.014

Note. *, $p < .001$; CFI, comparative fit index; ΔCFI, difference of comparative fit index; RMSEA, root mean square of error approximation; ΔRMSEA, difference of root mean square of error approximation, SRMR, standardised root mean square residual; ΔSRMR, difference of standardised root mean square residual

Table 5 Factor loadings for the ttheoritical bifactor model (*N* = 395)

Item No.	Contents	Bifactor model		
		Positively worded	Negatively worded	General
1	I found it difficult to get started with everyday tasks		**0.512**	**0.392**
2	I felt able to trust others	0.094		**0.584**
3	I felt unable to cope		**0.505**	**0.430**
4	I could do the things I wanted to do	0.210		**0.761**
5	I felt happy	0.231		**0.770**
6	I thought my life was not worth living		**0.401**	**0.514**
7	I enjoyed what I did	0.156		**0.776**
8	I felt hopeful about my future	0.014		**0.710**
9	I felt lonely		**0.452**	0.292
10	I felt confident in myself	-0.068		**0.764**
11	I did things I found rewarding	0.246		**0.662**
12	I avoided things I needed to do		**0.459**	0.220
13	I felt irritated		**0.470**	0.279
14	I felt like a failure		**0.516**	**0.508**
15	I felt in control of my life	-0.241		**0.586**
16	I felt terrified		**0.565**	**0.377**
17	I felt anxious		**0.647**	**0.404**
18	I had problems with my sleep		**0.417**	0.284
19	I felt calm	-0.265		**0.695**
20	I found it hard to concentrate		**0.554**	**0.300**

Note. Factor loadings > 0.3 are in bold. ReQoL™ Version 1.1 © Copyright, University of Sheffield 2016, 2018. All Rights Reserved. The authors have asserted their moral rights. Oxford University Innovation Limited is exclusively licensed to grant permissions to use the ReQoL™

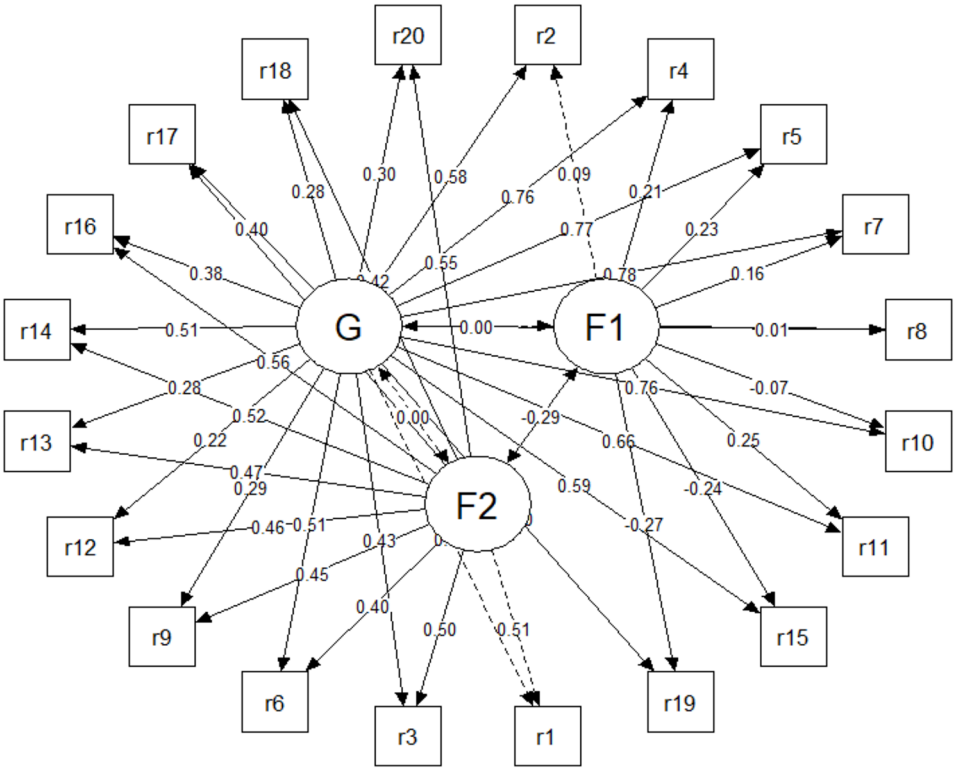

Fig. 4 Theoretical bifactor model of the ReQoL-J. Theoretical bifactor model of the ReQoL-J. Abbreviations: G, General factor; F1, Factor 1; F2, Factor 2. r1 to r20 denote item numbers

Table 6 Omega indices

	Omega/OmegaS	OmegaH/OmegaHS
General	0.926	**0.720**
Positively worded factor	0.905	0.003
Negatively worded factor	0.874	0.572

Note. Explained Common Variance (ECV) was **0.658** and Percentage of Uncontaminated Correlations (PUC) was 0.521. When PUC is greater than 0.80, or when ECV is greater than 0.60 and *OmegaH* is greater than 0.7, the relative bias from using a unidimensional measurement model rather than a bifactor measurement model is likely to be small (Reise et al., 2013) [71]

greater than loadings on specific group factors (Table 5; Fig. 4).

Dimensionality and reliability

Omega indices are shown in Table 6. *Omega* was 0.926, indicating high reliability. ECV, PUC, and *OmegaH* were 0.658 (>0.6), 0.521 (<0.8), and 0.720 (>0.7), respectively. Since ECV is greater than 0.60 and *OmegaH* is greater than 0.7, the relative bias from using a *unidimensional measurement model* is likely to be small for these data (Reise et al., 2013). ICCs for Re-QoL-J-10 and –20 were 0.700 (95% CI = 0.550 to 0.806) and 0.766 (95% CI = 0.641 to 0.851), respectively, suggesting moderate test-retest reliability.

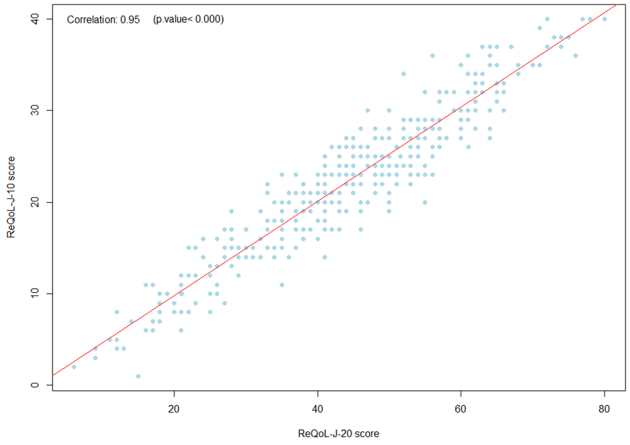

Fig. 5 Scatter plot of the ReQoL-J-20 total score and the ReQoL-J-10 total score. X-axis indicates the ReQoL-J-10 total score. Y-axis indicates the ReQoL-J-20 total score. Plots are shown in blue dots. Red solid line indicates regression line

Concurrent validity and convergent validity

The ReQoL-J-10 showed a significant and strong correlation with the ReQoL-J-20 (r = .95, p < .001). The scatter plot of the ReQoL-20 and the ReQoL-10 is shown in Fig. 5. With regard to the concurrent validity, the ReQoL-J-20 was significantly and positively correlated with the EQ-5D-5 L (r = .597, p < .001) (Table 7). Similarly, regarding convergent validity, the ReQoL-J-20 scores showed

Table 7 Correlations of the ReQoL-J with other variables

	ReQoL-J-10			ReQoL-J-20		
Measurements	***r***	***p***	**95% CI**	***r***	***p***	**95% CI**
Patient-reported measures						
EQ-5D-5 L	0.544	< 0.001	0.470, 0.610	0.597	< 0.001	0.530, 0.657
SPA-5	0.567	< 0.001	0.496, 0.630	0.579	< 0.001	0.509, 641
SiWB	0.664	< 0.001	0.606, 0.717	0.686	< 0.001	0.629, 0.735
PHQ-9	-0.675	< 0.001	-0.726, -0.618	-0.765	< 0.001	-0.803, -0.721
WHODAS	-0.528	< 0.001	-0.596, -0.453	-0.602	< 0.001	-0.661, -0.534
Staff-rated measures						
QLS						
Interpersonal relationships	0.199	< 0.001	0.103, 0.292	0.162	0.001	0.064, 0.256
Instrumental roles	0.183	< 0.001	0.085, 0.276	0.173	< 0.001	0.075, 267
Intrapsychic foundations	0.185	< 0.001	0.088, 0.279	0.162	0.001	0.064, 0.256
Common objective activities	0.127	0.011	0.029, 0.223	0.114	0.023	0.016, 210
Total	0.210	< 0.001	0.113, 0.302	0.181	< 0.001	0.084, 0.274
WHODAS	-0.203	< 0.001	-0.296, -0.107	-0.194	< 0.001	-0.288, -0.098

Note. EQ-5D-5 L, EuroQol 5-dimension 5-level; SPA-5, Five-item Subjective and Personal Agency scale; SiWB, Single-item well-being scale; PHQ-9, Patient Health Questionnaire 9-item version; WHODAS, 12 item-version of World Health Organization Disability Assessment Schedule 2.0; QLS, Quality of Life Scale

Table 8 Correlations of subscale scores of the ReQoL-J with other variables

	Positively worded			Negatively worded		
Measurements	***r***	***p***	**95% CI**	***r***	***p***	**95% CI**
Patient-reported measures						
EQ-5D-5 L	0.412	< 0.001	0.327, 0.491	0.614	< 0.001	0.549, 0.672
SPA-5	0.589	< 0.001	0.521, 0.650	0.428	< 0.001	0.343, 0.505
SiWB	0.631	< 0.001	0.568, 0.687	0.567	< 0.001	0.495, 0.630
PHQ-9	-0.527	< 0.001	-0.594, 0.451	-0.789	< 0.001	-0.823,. -749
WHODAS	-0.424	< 0.001	-0.502, 0.340	-0.612	< 0.001	-0.258, -0.066.
Staff-rated measures						
QLS						
Interpersonal relationships	0.219	< 0.001	0.123, 0.311	0.072	0.156	-0.027, 0.169
Instrumental roles	0.166	< 0.001	0.069, 0.261	0.136	0.007	0.037, 0.232
Intrapsychic foundations	0.157	0.002	0.060, 0.252	0.126	0.012	0.027, 0.222
Common objective activities	0.111	0.027	0.013, 208	0.009	0.079	-0.010, 0.185
Total	0.204	< 0.001	0.108, 0.207	0.088	0.116	0.017, 0.212
WHODAS	-0.175	< 0.001	-0.269, 0.008	-0.116	0.001	0.017, 0.212

Note. EQ-5D-5 L, EuroQol 5-dimension 5-level; SPA-5, Five-item Subjective and Personal Agency scale; SiWB, Single-item well-being scale; PHQ-9, Patient Health Questionnaire 9-item version; WHODAS, 12 item-version of World Health Organization Disability Assessment Schedule 2,0; QLS, Quality of Life Scale

significant correlation with other PROMs: a positive correlation with the SPA-5 (r =.579, p <.001) and SiWB scores (r =.686, p <.001), and a negative correlation with the PHQ-9 (r = -.765, p <.001) and self-rated WHO-DAS scores (r = -0.602, p <.001). There were significant but small correlations between the ReQoL-J-20 and staff-rated assessments (QLS's subscales [r <.173] and WHO-DAS [r = -0.194, p <.001]) apart from the QLS subscale score of common objective activities (r =.114, p =.023). Almost the same tendency was found in the correlation between the ReQoL-J scores (Re-QoL-10, positively worded, and negatively worded subscale scores) and other scales (Table 8).

Discussion

This study evaluated the psychometric properties of the ReQoL-J including factorial validity, dimensionality and reliability, concurrent validity, and convergent validity. In CFA, the bifactor model consisting of two specific group factors (i.e., positively and negatively worded factors) and a general factor was determined as the best-fit model for this study sample. This model showed good internal consistency (omega = 0.927), and ECV and *omegaH* suggested unidimensionality. In terms of test-retest reliability, the ReQoL-J-20 and – 10 showed acceptable values of ICCs (0.766, and 0.700, respectively). The ReQoL-J-10 was strongly correlated with the ReQoL-J-20, and likely to have measuring accuracy as well as the ReQoL-J-20. Furthermore, the ReQoL-J-10 and 20 showed moderate

correlation with other measures indicating goal concurrent and convergent validity.

In this study, the mean scores (SD) for the ReQoL-J-20 and the ReQoL-J-10 were almost identical to those obtained from the original version of the ReQoL using the UK sample (mean [SD]: 43.27 [19.93] and 21.99 [10.56], respectively; Keetharuth et al., 2024 [40], Table 13 in p.69). Although the mean scores of the ReQoL in Japan and the UK were similar, the measurement invariance or equivalence of the ReQoL-J is still unclear. In a comparison of the ReQoL-J and the ReQoL-HK, the mean scores for the ReQoL-J-20 and the ReQoL-J-10 were lower than those for the ReQoL-HK (mean: 60.33.27 and 28.54, respectively; Xu et al., 2021 [42], Table 4 on p. 6). However, as Xu et al. pointed out, direct comparisons are inappropriate without measurement invariance or equivalence. Further analysis is needed to test measurement invariance or equivalence across different countries.

While the factor analyses tested three models, the bifactor model of the ReQoL-J-20 demonstrated the highest model fit index values to the Japanese sample. This finding was consistent with previous studies in the UK [30–31] and Hong Kong [42]. Although ECV for the Traditional Chinese ReQoL (ReQoL-TC) was 51% in a Hong Kong study [42], and ECV for the original ReQoL in the UK were 78.9% in study 1, and 84.5% in study 2 [30]. Our research showed that ECV was 65.8% and *OmegaH* was 0.720, which was sufficient evidence to suggest unidimensionality. Thus, the factor structure of the ReQoL-J-20 is likely to be consistent with the original ReQoL of the UK version. The rigorous translation process by OUI may produce these findings. Unbiased measurement between two (or more) translated instruments is necessary to detect the result of true differences across comparable different languages [50, 78].

In this study, the bifactor model consisting of two specific group factors (i.e., positively and negatively worded factors) and a general factor was evaluated as the best model in terms of the fitness to the data. This finding is consistent with the original ReQoL of the UK version. As we noted previously, the ReQoL was originally developed to measure not only a positive aspect but also a negative aspect of quality of life. It may be subject to a method bias [79]. A mixture of positive (e.g., "I could do the things I wanted to do") and negative valence items (e.g., "I avoided things I needed to do") is likely to lead to careless responses from participants [80–81]. In such cases, reverse-worded items can influence model fit enough that researchers are likely to reject a single-factor model for a unidimensional scale [82–83]. Instead, the addition of a general factor as method factors (i.e., bifactor model) is recommended. This may be the reason why the bifactor model of the ReQoL, including its Japanese version, showed good fit to the data in several studies.

Our study confirms the reliability of the ReQoL-J. Koo et al. [74] suggested that the criteria of ICC between 0.5 and 0.7, and between 0.75 and 0.9 indicate moderate, and good reliability, respectively. The ReQoL met these criteria and is thus considered to have an acceptable level of reliability.

The scores of the ReQoL-J were consistently and moderately correlated with other relevant patient-reported outcome measures indicating good concurrent and convergent validity. Regarding concurrent validity, the ReQoL-J scores (i.e., ReQoL-10 and – 20 scores) showed significant positive correlation with the EQ-5D-5 L. This finding is consistent with the literature [31, 41–43, 84–87]. In terms of convergent validity, the ReQoL-J scores showed significant positive correlation with the SPA-5 and the SiWB. As pointed out in past studies, QoL correlates with well-being and personal agency [11–12, 88]. Therefore, findings being consistent with conceptual associations between variables indicated convergent validity. In particular, personal agency plays an important role in recovery from severe mental illness, as it is one of the facilitators of recovery [45, 46]. Our findings support this claim, showed relationship QoL in recovery and personal agency. On the other hand, ReQoL-J scores showed significant negative correlations with the PHQ-9 and the WHODAS ($p < .001$). Additionally, the ReQoL-J-10 and 20 showed moderate to strong correlation with the PHQ-9 (-0.675 and -0.765, respectively). As the severity of mental health problems or dysfunction showing negative correlations with QoL has been indicated in previous literature [18, 89], our findings replicated these relationships and indicated convergent validity of the ReQoL-J.

On the other hand, the ReQoL-J scores showed lower correlations with the staff-reported WHODAS than the patient-reported one. Likewise, correlations between the ReQoL-J scores and the QLS scores were overall low ($r < .199$). The correlation between scoring on the *Common objective activities* subscale of the QLS and ReQoL was not significant. This finding supports previous studies' results [19, 20, 21, 22, 90–91]. The ReQoL is likely to reflect a subjective QoL which contains mental health aspects, reasonable to understand the process of recovery from mental disorders and quality of life. Also, we should be conscious of subjective QoL and recognise interventions tailored to meet individual needs.

The original ReQoL was developed by service users and researchers, combining qualitative and quantitative evidence through a co-production process [30–32, 41]. This development process could be the strength of this scale. Our findings showed the ReQoL-J has similar psychometric properties to the original ReQoL and can be a promising tool to measure the subjective QoL of people living with mental health problems in Japanese cultural settings. Since quality of life is complex, multiple

approaches from different theoretical views are needed [92]. It is important to consider who is assessing and determining quality of life. Every measurement tool has pros and cons for psychometric properties. We should pay more attention to selecting measurement tools when an assessment of QoL is needed.

Limitations

This study has several limitations. Firstly, our study used a convenience sampling method. Considering the generality, further research should be conducted to support our findings. Secondly, our study did not analyse multigroup CFA or differential item functioning (DIF) using item response theory (IRT). Such strategies are essential to ensure measurement invariance or equivalence, which makes scale scores comparable across different populations (e.g., gender differences, different diagnostic groups, and different countries). Further research is needed to elucidate individual differences in the ReQoL-J. Thirdly, we used a clinical sample in this study. Comparing the ReQoL-J scores between clinical and healthy control samples would enhance the existing evidence on its psychometric properties, thereby contributing to the broader utility of the ReQoL-J.

Conclusions

Our findings confirm the factor validity of the ReQoL-J, which indicates the same factor structure in previous studies in other countries. Both the ReQoL-J-20 and ReQoL-10 showed good internal consistency, test-retest reliability, concurrent validity and construct validity, affirming the usefulness of using the ReQoL-J as an outcome measure for assessing subjective QoL for the Japanese clinical sample. Future studies need to examine the measurement invariance or equivalence of ReQoL-J and to compare the scale score with a healthy sample.

Acknowledgements

We are grateful all the staff members of the participating agencies. We are very grateful to all the participants in the study.

Author contributions

AH, SY, SS, KU, MI, and TK designed the study protocol. SY and NK secured the funding. KU, MI, TK, SS, and SY coordinated data collection and collected data. AH and SY performed statistical analysis. CF managed the research team members. AH, KU, MI, TK, SS, and SY drafted the manuscript. All the authors have approved the final version of the manuscript.

Funding

Primary funding for this work was through a grant from Japan Society for the Promotion of Science KAKENHI (grant no. JP23H00905 to SY); secondary funding came from the Japan Ministry of Health, Labour and Welfare (grant no. 23GC1011 to NK).

Data availability

The datasets generated and/or analysed during the current study are not publicly available, in accordance with the Japanese government's Ethical Guidelines for Medical and Health Research Involving Human Subjects and the ethics committee approval for this study, but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Research Ethics Committee of the National Centre for Neurology and Psychiatry (no. A2023-065). All participants were informed about the aims of the study, the affiliation of the principal investigator, the security of personal information and ethical considerations for participation in the study. Anonymity and voluntary participation were assured. Informed consent was obtained from all participants involved in this study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 3 December 2024 / Accepted: 17 February 2025
Published online: 20 March 2025

References

1. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook
2. Hannan MT, Felson DT, Pincus THEODORE. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol. 2000;27(6):1513–7.
3. Jones P, W. Health status measurement in chronic obstructive pulmonary disease. Thorax. 2001;56:880–7.
4. Müller-Bühl U, Engeser P, Klimm HD, Wiesemann A. Quality of life and objective disease criteria in patients with intermittent claudication in general practice. Fam Pract. 2003;20(1):36–40.
5. Yamaguchi S, Ojio Y, Koike J, Matsunaga A, Ogawa M, Kikuchi A, Kawashima T, Tachimori H, Bernick P, Kimura H, Inagaki A, Watanabe H, Kishi Y, Yoshida K, Hirooka T, Oishi S, Matsuda Y, Fujii C. Associations between readmission and patient-reported measures in acute psychiatric inpatients: a multicenter prospective longitudinal study. Soc Psychiatry Psychiatr Epidemiol. 2024. https://doi.org/10.1007/s00127-024-02710-5.
6. Weldring T, Smith SMS. Article commentary: patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Serv Insights. 2013;6. https://doi.org/10.4137/hsi.s11093. HSI.S11093.
7. Knaup C, Koesters M, Schoefer D, Becker T, Puschner B. Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. Br J Psychiatry. 2009;195(1):15–22.
8. Baandrup L, Rasmussen JØ, Mainz J, Videbech P, Kristensen S. Patient-reported outcome measures in mental health clinical research: a descriptive review in comparison with clinician-rated outcome measures. Int J Qual Health Care. 2022;34(Supplement1):ii70–97.
9. Roe D, Mazor Y, Gelkopf M. Patient-reported outcome measurements (PROMs) and provider assessment in mental health: a systematic review of the context of implementation. Int J Qual Health Care. 2022;34(Supplement1):ii28–39.
10. WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med. 1998;28(3):551–8.
11. Pinto S, Fumincelli L, Mazzo A, Caldeira S, Martins JC. Comfort, well-being and quality of life: discussion of the differences and similarities among the concepts. Porto Biomedical J. 2017;2(1):6–12.
12. Skevington SM, Böhnke JR. How is subjective well-being related to quality of life? Do we need two concepts and both measures? Soc Sci Med. 2018;206:22–30.
13. Edwards TC, Patrick DL, Topolski TD. Quality of life of adolescents with perceived disabilities. J Pediatr Psychol. 2003;28(4):233–41.
14. Strine TW, Kroenke K, Dhingra S, Balluz LS. The associations between depression, Health-Related Quality of Life, social support, life satisfaction, and disability in community Dwelling US adults. J Nerv Mental Disease. 2009;197(1):61–4.

15. Connell J, Brazier J, O'Cathain A, Lloyd-Jones M, Paisley S. Quality of life of people with mental health problems: a synthesis of qualitative research. Health Qual Life Outcomes. 2012;10:1–16.
16. Eack SM, Newhill CE. Psychiatric symptoms and quality of life in schizophrenia: a meta-analysis. Schizophr Bull. 2007;33(5):1225–37. https://doi.org/10.1093/schbul/sbl071.
17. Fleury M-J, Grenier G, Bamvita J-M. Associated and mediating variables related to quality of life among service users with mental disorders. Qual Life Res. 2018;27(2):491–502. https://doi.org/10.1007/s11136-017-1717-z.
18. Nevarez-Flores AG, Sanderson K, Breslin M, Carr VJ, Morgan VA, Neil AL. Systematic review of global functioning and quality of life in people with psychotic disorders. Epidemiol Psychiatric Sci. 2019;28(1):31–44. https://doi.org/10.1017/S2045796018000549.
19. Bengtsson-Tops A, Hansson L, Sandlund M, Bjarnason O, Korkeila J, Merinder L, Middelboe T. Subjective versus interviewer assessment of global quality of life among persons with schizophrenia living in the community: a nordic multicentre study. Qual Life Res. 2005;14:221–9.
20. Hayhurst KP, Massie JA, Dunn G, Lewis SW, Drake RJ. Validity of subjective versus objective quality of life assessment in people with schizophrenia. BMC Psychiatry. 2014;14:1–8.
21. Kusel Y, Laugharne R, Perrington S, McKendrick J, Stephenson D, Stockton-Henderson J, Burns T. Measurement of quality of life in schizophrenia: a comparison of two scales. Soc Psychiatry Psychiatr Epidemiol. 2007;42:819–23.
22. Sainfort F, Becker M, Diamond R. Judgments of quality of life of individuals with severe mental disorders: patient self-report versus provider perspectives. Am J Psychiatry. 1996;153(4):497–502. https://doi.org/10.1176/ajp.153.4.497.
23. Bobes J, Garcia-Portilla P, Saiz PA, Bascaran T, Bousono M. Quality of life measures in schizophrenia. Eur Psychiatry. 2005;20(S3):S313–7.
24. Citrome L, Mychaskiw MA, Cortez A, Opler M, Sopina L, Kotak S. Clinical outcome assessment instruments in schizophrenia: a scoping literature review with a focus on the potential of patient-reported outcomes. Innovations Clin Neurosci. 2023;20(4–6):14–33.
25. Cooke PJ, Melchert TP, Connor K. Measuring well-being: a review of instruments. Couns Psychol. 2016;44(5):730–57.
26. Schrank B, Bird V, Tylee A, Coggins T, Rashid T, Slade M. Conceptualising and measuring the well-being of people with psychosis: systematic review and narrative synthesis. Soc Sci Med. 2013;92:9–21.
27. van Krugten FCW, Feskens K, Busschbach JJV, Hakkaart-Van Roijen L, Brouwer WBF. Instruments to assess quality of life in people with mental health problems: a systematic review and dimension analysis of generic, domain- and disease-specific instruments. Health Qual Life Outcomes. 2021;19(1). https://doi.org/10.1186/s12955-021-01883-w.
28. Brazier J. Is the EQ–5D fit for purpose in mental health? Br J Psychiatry. 2010;197(5):348–9. https://doi.org/10.1192/bjp.bp.110.082453.
29. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Soeteman D, Brazier J. Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. Br J Psychiatry. 2014;205(3):236–43. https://doi.org/10.1192/bjp.bp.112.122283.
30. Keetharuth AD, Bjorner JB, Barkham M, Browne J, Croudace T, Brazier J. Exploring the item sets of the recovering quality of life (ReQoL) measures using factor analysis. Qual Life Res. 2019;28(4):1005–15. https://doi.org/10.1007/s11136-018-2091-1.
31. Keetharuth AD, Brazier J, Connell J, Bjorner JB, Carlton J, Taylor Buck E, Ricketts T, McKendrick K, Browne J, Croudace T, Barkham M. Recovering quality of life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. Br J Psychiatry: J Mental Sci. 2018;212(1):42–9. https://doi.org/10.1192/bjp.2017.10.
32. Keetharuth AD, Taylor BE, Acquadro C, Conway K, Connell J, Barkham M, Carlton J, Ricketts T, Barber R, Brazier J. Integrating qualitative and quantitative data in the development of outcome measures: the case of the recovering quality of life (ReQoL) measures in mental health populations. Int J Environ Res Public Health. 2018;15:1342.
33. Boodoo R, Saunders EFH, Thompson KS, Salzer M, Tan T-L, Passley-Clarke J, Fooks AR, Torrey WC, Kunkel EJ. Recovery at 30: perspectives from psychiatry clinicians and senior faculty. Commun Ment Health J. 2024. https://doi.org/10.1007/s10597-024-01308-0.
34. Cohen CI. Neorecovery: a critical analysis of the relationship between neoliberalism and the recovery movement. Commun Ment Health J. 2024. https://doi.org/10.1007/s10597-024-01275-6.
35. Ragins M. 30 years of recovery: perspectives from an aging revolutionary. Commun Ment Health J. 2024. https://doi.org/10.1007/s10597-024-01276-5.
36. Grundy A, Keetharuth AD, Barber R, Carlton J, Connell J, Taylor Buck E, Barkham M, Ricketts T, Robotham D, Rose D, Kay J, Hanlon R, Brazier J. Public involvement in health outcomes research: lessons learnt from the development of the recovering quality of life (ReQoL) measures. Health Qual Life Outcomes. 2019;17:60. https://doi.org/10.1186/s12955-019-1123-z.
37. Amering M, Schmolke M. Recovery in mental health: reshaping scientific and clinical responsibilities. Chichester: Wiley-Blackwell; 2009.
38. Leamy M, Bird V, Boutillier CL, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry. 2011;199(6):445–52. https://doi.org/10.1192/bjp.bp.110.083733.
39. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, Jones ML, Paisley S, O'Cathain A, Barkham M, Knapp M, Byford S, Gilbody S, Parry G. 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 (Winchester Eng). 2014;18(34):vii–188. https://doi.org/10.3310/hta18340.
40. Connell J, O'Cathain A, Brazier J. Measuring quality of life in mental health: are we asking the right questions? Soc Sci Med. 2014;120:12–20.
41. Keetharuth A, Brazier J, Connell J, Carlton J, Taylor Buck E, Ricketts T, Barkham M. (2024). Development and validation of the recovering quality of life (ReQoL) outcome measures (Version 1). The University of Sheffield. https://doi.org/10.15131/shef.data.25218911.v1
42. Xu RH, Keetharuth AD, Wang LL, Cheung AWL, Wong ELY. Psychometric evaluation of the Chinese recovering quality of life (ReQoL) outcome measure and assessment of health-related quality of life during the COVID-19 pandemic. Front Psychol. 2021;12:663035.
43. Xu RH, Keetharuth AD, Wang LL, Cheung AWL, Wong ELY. (2022). Measuring health-related quality of life and well-being: a head-to-head psychometric comparison of the EQ-5D-5L, ReQoL-UI and ICECAP-A. *European*. J Health Econ, 1–12.
44. Bellack AS, Drapalski A. Issues and developments on the consumer recovery construct. World Psychiatry. 2012;11:156–60. https://doi.org/10.1002/j.2051-5545.2012.tb00117.x.
45. Lysaker PH, Leonhardt BL. Agency: its nature and role in recovery from severe mental illness. World Psychiatry. 2012;11:165–6. https://doi.org/10.1002/j.2051-5545.2012.tb00121.x.
46. van Weeghel J, van Zelst C, Boertien D, Hasson-Ohayon I. Conceptualizations, assessments, and implications of personal recovery in mental illness: a scoping review of systematic reviews and meta-analyses. Psychiatr Rehabil J. 2019;42(2):169–81. https://doi.org/10.1037/prj0000356.
47. Yamaguchi S, Shiozawa T, Matsunaga A, Bernick P, Sawada U, Taneda A, Osumi T, Fujii C. Development and psychometric properties of a new brief scale for subjective personal agency (SPA-5) in people with schizophrenia. Epidemiol Psychiatric Sci. 2020;29:e111. https://doi.org/10.1017/S2045796020000256.
48. Mckenzie E, Matkin L, Fialho S, Emelurumonye L, Gintner IN, Ilesanmi T, Jagger C, Quinney B, Anderson S, Baandrup E, Bakhshy L, Brabban AK, Coombs A, Correll T, Cupitt CU, Keetharuth C, Lima AD, Mccrone DN, Moller P, Addington M, D. Developing an international standard set of patient-reported outcome measures for psychotic disorders. Psychiatric Serv. 2022;73(3):249–58. https://doi.org/10.1176/appi.ps.202000888.
49. Widaman KF, Revelle W. Thinking thrice about sum scores, and then some more about measurement and analysis. Behav Res Methods. 2023;55(2):788–806. https://doi.org/10.3758/s13428-022-01849-w.
50. Eremenco S, Pease S, Mann S, Berry P. Patient-reported outcome (PRO) consortium translation process: Consensus development of updated best practices. J Patient-reported Outcomes. 2018;2(12). https://doi.org/10.1186/s41687-018-0037-6.
51. Ikeda S, Shirowa T, Igarashi A, Noto S, Fukuda T, Saito S, Shimozuma K. Developing a Japanese version of the EQ-5D-5L value set. J Natl Inst Public Health. 2015;64(1):47–55. (in Japanese).
52. VanderWeele TJ, Trudel-Fitzgerald C, Allin P, Farrelly C, Fletcher G, Frederick DE, Kubzansky LD. Current recommendations on the selection of measures for well-being. Prev Med. 2020;133:106004.
53. Spitzer RL, Kroenke K, Williams JBW, the Patient Health Questionnaire Primary Care Study Group. &. Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. JAMA, 282(18), 1737–44. https://doi.org/10.1001/jama.282.18.1737
54. Inagaki M, Ohtsuki T, Yonemoto N, Kawashima Y, Saitoh A, Oikawa Y, Yamada M. Validity of the Patient Health Questionnaire (PHQ)-9 and PHQ-2 in general

internal medicine primary care at a Japanese rural hospital: a cross-sectional study. Gen Hosp Psychiatry. 2013;35(6):592–7.
55. Muramatsu K, Kamijima K. (2009). Puraimarikea shinnryou to utubyou sukuriningu tsuru: Patient Health Questionnaire-9 nihongoban 'Kokoroto Karadano Shitsumonhyou' (Primary care and depression screening tool: The Japanese version of the Patient health Questionnaire-9 'Questionnaire of Mind and Body'). *Shindan Chiryou 2009*, 97, 1465–1473. (In Japanese).
56. World Health Organization. (2010). Measuring health and disability: manual for WHO Disability Assessment Schedule (WHODAS 2.0). Available from: https://iris.who.int/bitstream/handle/10665/43974/9789241547598_eng.pdf?sequence=1
57. World Health Organization. (2015). *Measuring health and disability: manual for WHO Disability Assessment Schedule (WHODAS 2.0)*. Nihon Hyouronsha,(in Japanese).
58. World Health Organization. International Classification of Functioning, disability, and Health: Children & Youth Version. ICF-CY. World Health Organization; 2007.
59. Iwanaga M, Yamaguchi S, Sato S, Usui K, Nakanishi K, Nishiuchi E, Shimodaira M, So Y, Fujii C. Comparison of the 12-item and 36-item versions of the World Health Organization Disability Assessment Schedule (WHODAS) 2.0 using longitudinal data from community mental health outreach service users. Neuropsychopharmacol Rep. 2024;44(2):457–63.
60. Heinrichs DW, Hanlon TE, Carpenter WT Jr. The quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. Schizophr Bull. 1984;10(3):388–98.
61. Heinrichs DW, Halton TE, Carpenter WT. (2001). *The Quality of Life Scale.* Seiwa, Tokyo; 2001 (in Japanese).
62. Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990;107(2):238–46.
63. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods Psychol Res Online. 2003;8(2):23–74.
64. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model. 1999;6(1):1–55.
65. Kline RB. (2016). *Principles and practice of structural equation modeling. 5th edition.* Guilford publications.
66. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol. 2014;105:399–412.
67. Peters GJY. The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. Eur Health Psychol. 2014;16(2):56–69.
68. Zinberg RE, Revelle W, Yovel I, Li W. Cronbach's α, Revelle's β, and McDonald's ωH: their relations with each other and two alternative conceptualizations of reliability. Psychomterika. 2005;70(1):123–33.
69. Reise SP, Bonifay WE, Haviland MG. Scoring and modeling psychological measures in the presence of multidimensionality. J Pers Assess. 2013;95(2):129–40. https://doi.org/10.1080/00223891.2012.725437.
70. Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: calculating and interpreting statistical indices. Psychol Methods. 2016;21(2):137–50. https://doi.org/10.1037/met0000045.
71. Reise SP, Scheines R, Widaman KF, Haviland MG. Multidimensionality and structural coefficient bias in structural equation modeling: a bifactor perspective. Educ Psychol Meas. 2013;73(1):5–26.
72. Stucky BD, Edelen MO. Using hierarchical IRT models to create unidimensional measures from multidimensional data. In: Reise SP, Revicki DA, editors. Handbook of item response theory modeling: applications to typical performance Assessment. New York, NY: Routledge; 2015.
73. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1(1):30.
74. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
75. Rosseel Y. Lavaan: an R Package for Structural equation modeling. J Stat Softw. 2012;48(2):1–36. https://doi.org/10.18637/jss.v048.i02.
76. Gamer M, Lemon J, Gamer MM, Robinson A, Kendall's W. Package 'irr'. Var Coefficients Interrater Reliab Agreem. 2012;22:1–32.
77. Dueber DM. (2017). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. https://doi.org/10.13023/edp.tool.01 [Also available at http://sites.education.uky.edu/apslab/resources/]
78. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. Eval Health Prof. 2005;28(2):212–32.
79. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol. 2003;88(5):879.
80. Steinmann I, Strietholt R, Braeken J. A constrained factor mixture analysis model for consistent and inconsistent respondents to mixed-worded scales. Psychol Methods. 2022;27(4):667–702. https://doi.org/10.1037/met0000392.
81. Xin Z, Chi L. Wording effect leads to a controversy over the construct of the social dominance orientation scale. J Psychol. 2010;144(5):473–88.
82. Schmitt N, Stuits DM. Factors defined by negatively keyed items: the result of careless respondents? Appl Psychol Meas. 1985;9(4):367–73.
83. Woods CM. Careless responding to reverse-worded items: implications for confirmatory factor analysis. J Psychopathol Behav Assess. 2006;28:186–91.
84. Franklin M, Enrique A, Palacios J, Richards D. Psychometric assessment of EQ-5D-5L and ReQoL measures in patients with anxiety and depression: construct validity and responsiveness. Qual Life Res. 2021;30:2633–47.
85. Grochtdreis T, König HH, Gallinat J, Konnopka A, Schulz H, Lambert M, Karow A, Dams J. Validation of the recovering quality of life (ReQoL) questionnaires for patients with anxiety, obsessive-compulsive, stress-related, somatoform and personality disorders in Germany. J Psychiatr Res. 2023;157:202–11. https://doi.org/10.1016/j.jpsychires.2022.11.032.
86. Grochtdreis, T., König, H. H., Keetharuth, A. D., Gallinat, J., Konnopka, A., Schulz,H.,… Dams, J. (2023). Psychometric evaluation of the German version of the Recovering Quality of Life (ReQoL) measures in patients with affective disorders. European Journal of Health Economics, 24(4), 499–512.
87. van Aken BC, de Beurs E, Mulder CL, Van Der Feltz-Cornelis CM. The Dutch recovering quality of life questionnaire (ReQoL) and its psychometric qualities. Eur J Psychiatry. 2020;34(2):99–107.
88. Ho WW, Chiu MY, Lo WT, Yiu MG. Recovery components as determinants of the health-related quality of life among patients with schizophrenia: structural equation modelling analysis. Australian New Z J Psychiatry. 2010;44(1):71–84.
89. Tang AL, Thomas SJ. Relationships between depressive symptoms, other psychological symptoms, and quality of life. Psychiatry Res. 2020;289:113049.
90. Lasalvia A, Bonetto C, Malchiodi F, Salvi G, Parabiaghi A, Tansella M, Ruggeri M. Listening to patients' needs to improve their subjective quality of life. Psychol Med. 2005;35(11):1655–65.
91. Renwick, L., Drennan, J., Sheridan, A., Owens, L., Lyne, J., O'Donoghue, B.,… Clarke,M. (2015). Subjective and objective quality of life at first presentation with psychosis.Early intervention in psychiatry, 11(5), 401–410.
92. Diener E, Suh E. Measuring quality of life: Economic, social, and subjective indicators. Soc Indic Res. 1997;40:189–216.

Publisher's note