Psychometric Assessment of an Item Bank for Adaptive Testing on Patient-Reported Experience of Care Environment for Severe Mental Illness: Validation Study

Sara Fernandes[1], PhD; Yann Brousse[1], MSc; Xavier Zendjidjian[1], MD, PhD; Delphine Cano[1], MSc; Jérémie Riedberger[1], MSc; Pierre-Michel Llorca[2], MD, PhD; Ludovic Samalin[2], MD, PhD; Daniel Dassa[1], MD, PhD; Christian Trichard[3], MD, PhD; Vincent Laprevote[4,5], MD, PhD; Anne Sauvaget[6], MD, PhD; Mocrane Abbar[7], MD, PhD; David Misdrahi[8], MD, PhD; Fabrice Berna[5], MD, PhD; Christophe Lancon[1], MD, PhD; Nathalie Coulon[9], MD, PhD; Wissam El-Hage[10], MD, PhD; Pierre-Emmanuel Rozier[11], MD; Michel Benoit[12], MD, PhD; Bruno Giordana[12], MD, PhD; Alejandra Caqueo-Urízar[13], PhD; Dong Keon Yon[14], MD, PhD; Bach Tran[1,15], PhD; Pascal Auquier[1], MD, PhD; Guillaume Fond[1], MD, PhD; Laurent Boyer[1], MD, PhD

[1]Assistance Publique-Hopitaux de Marseille, Aix-Marseille University, UR3279: Health Service Research and Quality of Life Center - CEReSS, Marseille, France

[2]Department of Psychiatry, Centre Hospitalier Universitaire de Clermont-Ferrand, University of Clermont Auvergne, Centre national de la recherche scientifique, Institut national polytechnique de Clermont Auvergne, Institut Pascal UMR 6602, Clermont-Ferrand, France

[3]Etablissement public de santé Barthélemy Durand, Etampes, France

[4]Department of Addictology and Psychiatry, Centre Psychothérapique de Nancy, Laxou, France

[5]Département de Psychiatrie, Centre Hospitalier Régional Universitaire de Strasbourg, Université de Strasbourg, Institut national de la santé et de la recherche médicale U1114, Fédération de Médecine Translationnelle de Strasbourg, Strasbourg, France

[6]Nantes Université, Centre Hospitalier Régional Universitaire de Nantes, Movement - Interactions - Performance - MIP UR 4334, Nantes, France

[7]Department of Psychiatry, Centre Hospitalier Universitaire de Nîmes, University of Montpellier, Nîmes, France

[8]National Centre for Scientific Research UMR 5287 - Institut de Neurosciences Cognitives et Intégratives d'Aquitaine, University of Bordeaux, Centre Hospitalier Charles Perrens, Bordeaux, France

[9]Centre Expert Schizophrénie, Centre Expert TSA-SDI et Centre Référent de Réhabilitation Psychosociale et de Remédiation Cognitive - C3R, Centre Hospitalier Alpes Isère, Grenoble, France

[10]Centre Hospitalier Régional Universitaire de Tours, Clinique Psychiatrique Universitaire, Tours, France

[11]Department of Psychiatry, Centre Hospitalier des Pyrénées, Pau, France

[12]Department of Psychiatry, Hopital Pasteur, University Hospital of Nice, Nice, France

[13]Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Chile

[14]Center for Digital Health, Medical Science Research Institute, Kyung Hee University College of Medicine, Department of Pediatrics, Kyung Hee University Medical Center, Seoul, Republic of Korea

[15]Institute of Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam

Corresponding Author:
Sara Fernandes, PhD
Assistance Publique-Hopitaux de Marseille, Aix-Marseille University, UR3279: Health Service Research and Quality of Life Center - CEReSS
27, Boulevard Jean-Moulin
Marseille, 13385
France
Phone: 33 660185077
Email: sarah.fernandes@ap-hm.fr

Abstract

Background: The care environment significantly influences the experiences of patients with severe mental illness and the quality of their care. While a welcoming and stimulating environment enhances patient satisfaction and health outcomes, psychiatric facilities often prioritize staff workflow over patient needs. Addressing these challenges is crucial to improving patient experiences and outcomes in mental health care.

Objective: This study is part of the Patient-Reported Experience Measure for Improving Quality of Care in Mental Health (PREMIUM) project and aims to establish an item bank (PREMIUM-CE) and to develop computerized adaptive tests (CATs) to measure the experience of the care environment of adult patients with schizophrenia, bipolar disorder, or major depressive disorder.

Methods: We performed psychometric analyses including assessments of item response theory (IRT) model assumptions, IRT model fit, differential item functioning (DIF), item bank validity, and CAT simulations.

Results: In this multicenter cross-sectional study, 498 patients were recruited from outpatient and inpatient settings. The final PREMIUM-CE 13-item bank was sufficiently unidimensional (root mean square error of approximation=0.082, 95% CI 0.067-0.097; comparative fit index=0.974; Tucker-Lewis index=0.968) and showed an adequate fit to the IRT model (infit mean square statistic ranging between 0.7 and 1.0). DIF analysis revealed no item biases according to gender, health care settings, diagnosis, or mode of study participation. PREMIUM-CE scores correlated strongly with satisfaction measures (r=0.69-0.78; P<.001) and weakly with quality-of-life measures (r=0.11-0.21; P<.001). CAT simulations showed a strong correlation (r=0.98) between CAT scores and those of the full item bank, and around 79.5% (396/498) of the participants obtained a reliable score with the administration of an average of 7 items.

Conclusions: The PREMIUM-CE item bank and its CAT version have shown excellent psychometric properties, making them reliable measures for evaluating the patient experience of the care environment among adults with severe mental illness in both outpatient and inpatient settings. These measures are a valuable addition to the existing landscape of patient experience assessment, capturing what truly matters to patients and enhancing the understanding of their care experiences.

Trial Registration: ClinicalTrials.gov NCT02491866; https://clinicaltrials.gov/study/NCT02491866

Introduction

The health care environment, which encompasses design features (ie, cleanliness, food, privacy, waiting time, basic amenities) and the overall atmosphere (or climate) [1], has been recognized as a significant factor influencing the experiences of patients with severe mental illness (SMI) [2-5]. It is an important factor in the quality of patient care [2,6-8], contributing to improved patient satisfaction [9] and improved health outcomes [10,11]. In a recent study, patients identified a welcoming environment as one of the most important aspects of their care [12]. Indeed, a calm and welcoming environment helps to improve patients' sense of control and empowerment and, consequently, reinforces their willingness to follow recommended treatments. In addition, the care environment is the patients' first impression and can lead to a positive image of the therapeutic process [13]. A supportive environment promotes communication between patients and staff, can help reduce stressful stimuli, and thus prevents relapses and risky behavior. The priority for psychiatric facilities is therefore to provide patients with a warm and safe atmosphere that allows for positive social interactions, with opportunities for stimulating activities, enabling patients to facilitate their recovery and transition to the community. Different theoretical models can shed light on the additional nonpharmacological and biopsychosocial effects of a patient's care experience, including the placebo response effects and the set and setting theory [14,15].

Recommended features to promote patient recovery [16-19] include smaller, home-like units with well-decorated common spaces, open designs, access to nature and daylight, and an environment that is clean, well laid out, and ensures privacy and security for personal effects. However, psychiatric facilities are often criticized for prioritizing staff workflow over patient needs [2], leading in some cases to a perceived "prison-like atmosphere" [16,20,21] characterized by conflicting routines and rules and a lack of stimulation [22,23]. Some patients have reported feelings of boredom, loneliness, and stigmatization in these environments [21-26]. The lack of stimulating activities and positive social interactions is a barrier to patients' successful recovery [24-30]. These negative experiences can contribute to decreased patient satisfaction, increased levels of anxiety and stress among patients, ineffective care, and signs of burnout among staff [27,31,32]. Emphasis should be placed on the design of psychiatric facilities, as a difficult environment is a barrier to care, and patients often perceive such an environment as a lack of attention from staff [30]. In psychiatry, patients cope with an unfamiliar and potentially stressful environment [33], and a better understanding of their experiences is essential to identify and improve current barriers.

Given this growing interest, it is necessary to provide a valid and reliable instrument for measuring patients' experience of the care environment, applicable to both inpatient and outpatient settings, as care pathways for patients with SMI often combine several care modalities. Previous research has demonstrated that patients with SMI can provide reliable and valid responses to self-administered questionnaires; the impact of psychiatric symptoms and cognitive deficits seems to be negligible [34,35]. The French group PREMIUM (Patient-Reported Experience Measure for Improving Quality of Care in Mental Health) is developing item banks and computerized adaptive tests (CATs) to improve the systematic use of patient-reported experience measures in mental health care [36]. The use of CATs

XSL•FO

RenderX

significantly reduces measurement burden by administering a limited number of items targeted to the respondent's experience level, aiming to improve measurement accuracy.

The objective of this study was to calibrate an item bank and develop a CAT to assess the care environment experienced by adult patients with SMI. These measures will contribute to the current landscape of patient experience measures by providing a valuable complement to PREMIUM measures and capturing what really matters to patients.

Methods

Study Population and Procedure

This is a national, multicenter, cross-sectional study conducted between January 2016 and December 2021. Patients were recruited through in- and outpatient psychiatric settings of a French teaching hospital (Assistance Publique-Hôpitaux de Marseille), the FondaMental Foundation's expert centers [37], and through an online survey. In mental health settings, stable patients who met the inclusion criteria were identified and approached by a member of their usual care team to invite them to participate in the study. The link to the web survey was distributed through patient associations.

Inclusion criteria were as follows: age older than 18 years and younger than 65 years with a diagnosis of schizophrenia, bipolar disorder, or major depressive disorder (MDD), receiving inpatient or outpatient psychiatric care, and speaking or reading French. Vulnerable persons (ie, pregnant or nursing women, persons under legal protection) or those unable to complete a self-administered questionnaire were not included in the study.

Current recommendations suggest a sample size of 300-500 observations for multiparameter item response theory (IRT) models [38-40]. Consequently, we estimated that a sample of around 500 patients would be sufficient to obtain reasonably stable estimates.

Data Collection

Data were collected through paper questionnaires in health care settings and online through a web survey. Patients reported the following sociodemographic and clinical characteristics: gender, age, educational level, marital status, occupational status, main diagnosis (schizophrenia, bipolar disorders, or MDD), duration of illness, and quality of life (QoL) as measured using the medical outcome study 12-item Short Form (SF-12) [41], which describes 8 QoL dimensions: physical functioning, social functioning, role physical, role emotional, mental health, vitality, bodily pain, general health, and 2 composite scores for physical and mental QoL (ranging from 0 to 100, with higher scores indicating better QoL). Adequate psychometric properties of the SF-12 have been demonstrated among individuals with SMI [42], and the SF-12 has proven to be a good alternative to the SF-36 for minimizing response burden.

The PREMIUM for Care Environment (PREMIUM-CE) item bank consists of 16 items designed for patients with SMI and measures their experience regarding the care environment over the past 4 weeks. Participants respond to the items on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree"

with a "not applicable" response option. Additionally, an overall satisfaction item ("Overall, are you satisfied with the health care facilities in which you receive care?") and a visual analog scale (VAS; minimum 0 to maximum 10) were collected. PREMIUM-CE items were identified through face-to-face interviews with patients with SMI and a systematic review of existing patient-reported experience measure; then the item pool was refined based on an expert review and cognitive interviews with patients with SMI [4,5,36].

Statistical Analysis

Basic Descriptive Analysis

Descriptive statistics were calculated to describe participants' characteristics, including frequencies and percentages for categorical variables and means and SDs for continuous variables. Response rates, means and SDs, and ceiling and floor effects were also calculated for each item.

IRT Assumptions

Unidimensionality, local independence, and monotonicity are the 3 fundamental assumptions of IRT [43]. Data were randomly divided into 2 data sets (n=249 each), one for exploratory factor analysis and one for confirmatory factor analysis (CFA) with the weighted least squares mean and variance estimator to ensure that the PREMIUM-CE was sufficiently unidimensional [44]. Local independence was examined using residual correlations from the final CFA model. Monotonicity was examined using visual inspection of characteristic item curves.

Calibration and Fitting an IRT Model

Item parameters were estimated using the generalized partial credit model (GPCM) [45] and compared to the partial credit model [46]. IRT handles missing values by using full information maximum likelihood estimation, which uses all available information, and GPCM is recommended when the amount of missing data is high (20% or more) [38]. Item fit was assessed by examining the mean square infit statistics, which reflect the information-weighted mean squared residuals between the observed and expected response patterns. PREMIUM-CE scores (θ) were estimated by the Bayesian Expected a Posteriori estimation method [47], and a linear transformation was performed to obtain PREMIUM-CE scores ranging from 0 to 100 (higher scores indicate better experience with the care environment). The information curve of the final item bank was calculated, and high measurement precision was defined as an information score >10, corresponding to a reliability of >0.90 [39].

Differential Item Functioning Analysis

DIF was examined using ordinal logistic regression models [48,49] by gender (man vs woman), age (median distribution), care setting (outpatient vs inpatient), psychiatric diagnosis (schizophrenia vs bipolar disorder vs MDD), and mode of study participation (online survey vs health care settings).

External Validity

Construct validity was examined through convergent and discriminative validity assessments. For convergent validity, Spearman's rank correlations were computed between

PREMIUM-CE scores and both satisfaction (global satisfaction item and VAS) and QoL (SF-12 subscales and composite scores) scores. Our hypothesis was that PREMIUM-CE scores would have strong correlations with satisfaction scores ($r > 0.60$), which are 2 related measures, and weak correlations with QoL scores ($r < 0.30$). For discriminant validity, relationships between PREMIUM-CE scores and sociodemographic and clinical characteristics of the respondents were examined by using 2-tailed t tests, ANOVA, and Pearson correlations. The Q-Q plot was used to determine that the data are approximately normally distributed. Based on previous studies of the determinants of patient satisfaction with psychiatric services [3,50,51], our hypotheses were that higher levels of patient experience of the care environment were associated with older age, being female, being nonsingle, or being in an outpatient setting.

CAT Simulations

These simulations using participants' actual responses were run from the calibrated item bank and compared to identify the best performing CAT version. The stopping rules were based on standard error of measurement (SEM) values of 0.33, 0.44, and 0.55 (corresponding to a reliability between 0.90 and 0.70 [52]). The item administered at baseline was the one that offered the most information to the population mean ($\theta = 0$), and then items were administered according to the maximum Fisher information criterion [53].

The indicators used at each stage of the psychometric analyses are presented in Table S1 in Multimedia Appendix 1 [44,54-70]. All of the statistical analyses were performed using the following software: SPSS (version 20.0; IBM Corp), MPlus (version 7.0; Muthen & Muthen), and R (version 4.2.0; R Core Team), using packages *mirt* [71], *lordif* [72], *BifactorIndicesCalculator* [73], and *mirtCAT* [74]. A 2-tailed $P < .05$ was considered statistically significant.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the relevant ethics committee (2014-A01152-45). The study was registered in ClinicalTrials.gov (NCT02491866). All participants provided nonopposition, as required by French law. Additionally, all data were anonymized.

Results

Sample Characteristics

Of the 498 participants, 50.2% (250/498) were men, 72.3% (345/477) were unemployed, 73.5% (350/476) were single, and 70.5% (337/478) had an education level of a bachelor's degree or higher. The average age was 40.9 (SD 11.9) years, and the mean duration of illness was 12.9 (SD 9.3) years. In total, 51.8% (253/488) of the participants had a diagnosis of schizophrenia, 24.4% (119/488) had bipolar disorder, or 23.8% (116/488) had MDD, and 77.7% (387/497) of them were outpatients. The characteristics of the sample are presented in Table 1.

Table 1. Sample characteristics.

Characteristics	Values
Study participation, n/N (%)	
Health care setting	271/498 (54.4)
Online survey	227/498 (45.6)
Sociodemographic data	
Gender (man), n/N (%)	250/498 (50.2)
Age (years), mean (SD; n=496)	40.9 (11.9)
Marital status (single), n/N (%)	350/476 (73.5)
Educational level (<bachelor's degree), n/N (%)	141/478 (29.5)
Employment status (unemployed), n/N (%)	345/477 (72.3)
Clinical data	
Care setting, n/N(%)	
Outpatient	387/498 (77.7)
Inpatient	111/498 (22.3)
Inpatient with involuntary commitment	40/111 (36.1)
Main diagnosis (n=488), n (%)	
Schizophrenia	253 (51.8)
Bipolar disorder	119 (24.4)
Major depressive disorder	116 (23.8)
Duration of illness (years; n=469)	
Value, mean (SD)	12.9 (9.3)
<5 years, n (%)	105 (22.4)
≥5 years, n (%)	364 (77.6)
Quality of life (SF-12[a] scores), mean (SD)	
Physical functioning (n=490)	46.5 (11.4)
Social functioning (n=491)	34.3 (11.8)
Role physical (n=491)	40.5 (11.1)
Role emotional (n=491)	33.3 (12.4)
Mental health (n=493)	45.0 (11.1)
Vitality (n=491)	51.2 (10.3)
Bodily pain (n=493)	44.1 (12.8)
General health (n=492)	34.8 (10.5)
Physical composite (n=484)	43.8 (10.3)
Mental composite (n=484)	39.3 (11.5)

[a]SF-12: 12-item Short Form.

Basic Descriptive Statistics

The mean item scores ranged from 2.07 (SD 1.32) to 3.24 (SD 0.89), and most items had a missing data rate <10% (except items CE10, CE12, and CE15). The floor and ceiling effects ranged from 1.8% to 10.6% and from 10% to 45.2%, respectively. The interitem correlation values ranged from 0.01 to 0.79, and 3 pairs of items showed too high interitem correlations (>0.70): items CE3-CE4 (r=0.73), items CE3-CE5 (r=0.78), and items CE4-CE5 (r=0.79). Items CE3 and CE5 were excluded because their content was considered less relevant than the remaining items. The lowest scores were for item CE15 ("food was of good quality"), item CE12 ("you had access to media (telephone, computer, internet or Wi-Fi connection, etc)," and item CE10 ("the health care facilities were well equipped"). Table 2 summarizes the distribution of responses for each item.

Table 2. Descriptive statistics of PREMIUM-CE item bank.

| Item number	Content item	Score, mean (SD)	Floor effect (%)	Ceiling effect (%)	Missing values (%)	Skewness coefficient	Interitem correlations (Range)
CE1	The health care facilities were easily accessible (distance from home, parking, etc)	3.13 (1.09)	4.6	45.2	2	–1.42	0.01-0.66
CE2	The health care facilities were easy to find (eg, signage present and adapted)	3.13 (1.04)	3.4	43.4	1.8	–1.34	0.10-0.51
CE3	The health care facilities were welcoming	2.86 (1.17)	6.2	35.1	0.2	–0.96	0.27-0.78
CE4	The health care facilities were well-laid-out	2.91 (1.09)	4.8	33.5	0.4	–1.04	0.26-0.79
CE5	The health care facilities were pleasant	2.68 (1.19)	6.8	28.7	0	–0.72	0.27-0.79
CE6	The health care facilities were quiet enough	2.83 (1.15)	7.0	30.9	0.4	–1.06	0.21-0.50
CE7	The health care facilities were comfortable (chairs, armchairs, beds, etc)	2.89 (1.07)	4.8	30.1	0.4	–1.09	0.31-0.64
CE8	The health care facilities were clean	3.24 (0.89)	1.8	44.8	0.2	–1.46	0.21-0.63
CE9	The health care facilities were adapted to your needs	3.01 (1.03)	3.4	36.3	1.2	–1.13	0.23-0.68
CE10	The health care facilities were well equipped (materials for activities, group rooms, etc)	2.63 (1.22)	5.8	21.3	20.5	–0.68	0.25-0.66
CE11	The waiting time was acceptable	2.76 (1.21)	7.2	30.7	1.8	–0.90	0.30-0.49
CE12	You had access to media (telephone, computer, internet or Wi-Fi connection, etc)	2.24 (1.39)	10.0	17.5	26.7	–0.21	0.18-0.52
CE13	The sanitary facilities (toilets, bathroom, etc) were clean	3.08 (1.05)	3.2	39.2	7.4	–1.24	0.19-0.63
CE14	The health care facilities guarantee the respect for your privacy	3.12 (1.06)	4.4	41.8	4.6	–1.41	0.25-0.59
CE15	The food was of good quality, if you had to eat	2.07 (1.32)	10.6	10.0	39.4	–0.15	0.01-0.31
CE16	The smoking ban was respected	3.16 (1.02)	3.2	42.6	6.8	–1.37	0.21-0.50

IRT Assumptions

In EFA, 2 factors had eigenvalue greater than 1, and the scree plot and parallel analysis indicated 2 factors. The eigenvalue of the first factor was 6.46 and explained 46.11% of the total variance; the second factor was 1.33, and the ratio was 4.86. Evaluations indicated that the 2 spatial accessibility items (CE1 and CE2) may form a separate factor, and after a content review, only item CE1 was kept as it was deemed the most relevant. The 1-factor CFA model provided evidence to support the unidimensionality of the remaining 13 items (root mean square error of approximation [RMSEA]=0.082; 95% CI 0.067-0.097; comparative fit index=0.974; Tucker-Lewis index=0.968) and no items showed local dependence (all residual correlations were above |0.20|). Of the 13 items in the bank, 10 were recoded to meet the monotonicity assumption (Table S2 in Multimedia Appendix 1), which improved the model fit (Akaike information criterion=–3343.78 and Bayes information criterion=–3428). Cronbach α was .91.

Calibration and Fitting an IRT Model

The GPCM was used to calibrate the item bank and showed superior fit compared to the partial credit model (10,192.60 and 10,367.43 for Akaike information criterion and 10,382.07 and 10,506.38 for Bayes information criterion; and χ^2=198.84; P<.001); item fit was good (infit values ranging between 0.74 and 1.00). IRT parameter estimates for the 13 items showed slopes ranging from 0.55 to 2.85 and thresholds ranging from –2.07 to 2.29. Item parameters and item fit are provided in Table S2 in Multimedia Appendix 1. As shown in Figure 1, PREMIUM-CE provided the most information in the scale range between –2.6 and 1.4 and had a high measurement accuracy (reliability >0.90) in a shorter range between –2.1 and 0.7 (which corresponds to 88.6% of total information). Item CE7 was the most informative of the bank—"the health care facilities were comfortable," whereas item CE15 was the least informative—"the food was of good quality."

Figure 1. The test information for the Patient-Reported Experience Measure for Improving Quality of Care in Mental Health for Care Environment (PREMIUM-CE) item bank.
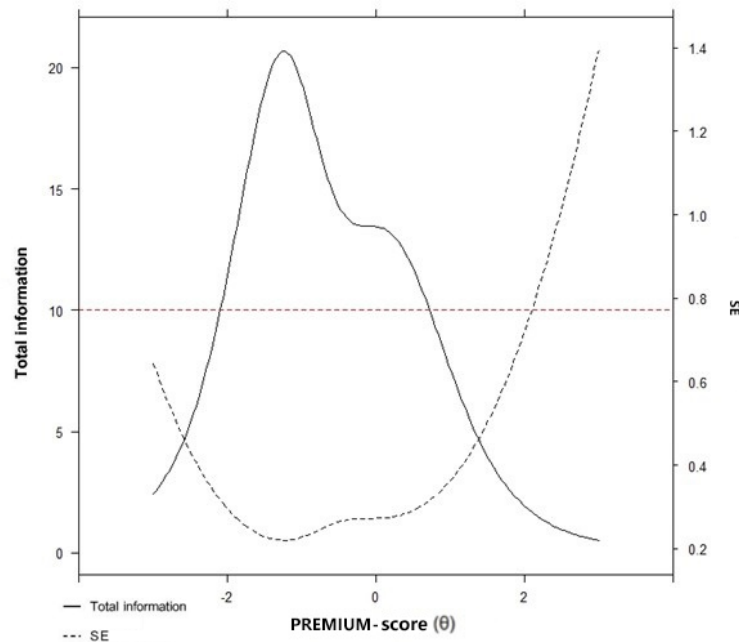
Differential Item Functioning Analysis

Responses to items CE6 (quiet) and CE13 (sanitary) were flagged for overall DIF but with negligible magnitude according to health care settings. Likewise, the DIF magnitude was negligible for item CE16 (smoking ban) according to mode of study participation and for item CE15 (food) according to gender, mode of study participation, and diagnostic after pooling bipolar disorder and MDD (mood disorders vs schizophrenia; P=.02; ΔR^2=.013). None of the items showed significant DIF for age. DIF results are provided in Table S3 in Multimedia Appendix 1.

External Validity

As expected, there were strong correlations between the PREMIUM-CE item bank and overall satisfaction and VAS, supporting convergent validity. Similarly, all SF-12 dimensions were weakly correlated with the PREMIUM-CE item bank, except for bodily pain and vitality. Associations were found between better experience of the care environment (ie, higher PREMIUM-CE scores) and older age, being a woman, being voluntarily admitted to a hospital, and being recruited through health care facilities. There was no significant effect of educational level, marital status, employment status, diagnosis, or duration of illness. These results are presented in Table 3.

Table 3. Comparison of PREMIUM-CE scores with sociodemographic and clinical data and proxy measures of quality of care.

Characteristics	Correlation coefficient (r)	Mean (SD)	P value
Study participation			<.001
Health care setting	N/A[a]	63.12 (17.58)	
Online survey	N/A	52.53 (20.34)	
Sociodemographic data			
Age	0.19	N/A	<.001
Gender			.04
Man	N/A	56.52 (18.84)	
Woman	N/A	60.18 (20.19)	
Marital status			.51
Single	N/A	58.67 (19.12)	
Nonsingle	N/A	57.33 (21.37)	
Educational level			.07
<Bachelor's degree	N/A	60.88 (18.86)	
≥Bachelor's degree	N/A	57.34 (20.04)	
Employment status			.60
Employed	N/A	57.58 (19.36)	
Unemployed	N/A	58.64 (19.95)	
Clinical data			
Care setting			.01
Outpatient	N/A	57.34 (20.11)	
Inpatient voluntarily admitted	N/A	65.54 (17.61)	
Inpatient involuntarily admitted	N/A	54.33 (14.52)	
Main diagnosis			.58
Schizophrenia	N/A	57.46 (18.48)	
Bipolar disorder	N/A	59.73 (20.72)	
Major depressive disorder	N/A	58.22 (20.97)	
Duration of illness			.14
<5 years	N/A	60.86 (18.79)	
≥5 years	N/A	57.64 (20.08)	
Proxy measures			
Item of overall satisfaction	0.78	N/A	<.001
Visual analog scale	0.69	N/A	<.001
Quality of life (SF-12[b] scores)			
Physical functioning	0.14	N/A	.003
Social functioning	0.19	N/A	<.001
Role physical	0.21	N/A	<.001
Role emotional	0.19	N/A	<.001
Mental health	0.12	N/A	.01
Vitality	–0.04	N/A	.41
Bodily pain	0.09	N/A	.05
General health	0.13	N/A	.004

Characteristics	Correlation coefficient (r)	Mean (SD)	P value
Physical composite	0.14	N/A	.001
Mental composite	0.11	N/A	.01

[a]N/A: not applicable.

[b]SF-12: 12- items short form.

CAT Simulations

As reported in Table 4, the results of the CAT simulations based on SEM <.33 and <.44 were both acceptable in terms of accuracy and precision, but the scenario based on SEM <.33 (corresponding to a reliability of 0.90) was the most efficient. Of the 498 participants included in the simulation, 79.5% (396) achieved a reliable score with an average of 7 items administered.

Table 4. Mean scores and precision indicators for each computerized adaptive test simulation.

Precision level and indicators	Values
SEM[a]<0.33	
Mean (SD)	58.30 (19.39)
Correlation coefficient (r)	0.98
RMSE[b]	0.17
Mean number of items	6.95
SEM<0.44	
Mean (SD)	58.35 (18.75)
Correlation coefficient (r)	0.95
RMSE	0.29
Mean number of items	4.46
SEM<0.55	
Mean (SD)	50.57 (21.27)
Correlation coefficient (r)	0.92
RMSE	0.37
Mean number of items	3.10

[a]SEM: standard error of measurement.

[b]RMSE: root mean square error.

Discussion

Principal Findings

In this study, we report the calibration and initial evaluation of a new PREMIUM-CE item bank measuring patients' experience of the care environment that can be used for CATs. The PREMIUM-CE questionnaire is the first available questionnaire thus far to assess the quality of the care environment, applicable in outpatient and inpatient settings, for adults with SMI. This new measure covers different facets of the care environment, including ease of access in time and space, facility layout and basic amenities, food quality, comfort and cleanliness, respect for privacy, and smoking ban. PREMIUM-CE items address both concerns common to all patients (eg, cleanliness or food) and those more specific to psychiatric patients (eg, therapeutic workshops). Existing instruments measure more objective aspects (eg, checklists fulfilled by direct observation), and patients with SMI were not involved in the development and validation process [75].

PREMIUM-CE has undergone rigorous psychometric evaluation, consistent with previous studies conducted as part of the French PREMIUM initiative [36]. Although the RMSEA was slightly above the criterion of <.08, our results provide evidence of sufficient unidimensionality, and the item pool meets the assumptions for IRT modeling. Research has shown that the RMSEA statistic is problematic for assessing the unidimensionality of item banks measuring health concepts [76], as RMSEA is sensitive to model complexity (number of estimated parameters) and skewed data distributions [77]. These results are comparable to other calibration studies of item banks of patient-reported measures [78-83]. Overall, our results demonstrate that PREMIUM-CE has strong psychometric properties for patients with SMI, with negligible measurement bias by gender, health care settings, and mode of study participation. Items CE10, CE12, and CE15 had a higher rate of missing data than the other items, but this rate was below 40%, which remains acceptable by psychometrics standards [84]. In addition, these items had lower scores compared to

others, meaning that efforts should be targeted on these aspects to improve the experience of patients with SMI. Future studies should examine whether changes to these items are required. The absence of a large DIF magnitude according to health care settings will make it possible to study changes in the experience of psychiatric patients over time, for whom care pathways often combine inpatient and outpatient care modalities. The 13 items in the final version of the PREMIUM-CE are listed in Multimedia Appendix 1, Table S4. In addition, the CAT version showed comparable measurement accuracy to the full item bank with high correlations between scores with an average of only 7 items administered.

External validity, explored using validated questionnaires and sociodemographic and clinical data, generally supported our initial hypotheses. Previous research has demonstrated that some factors, such as age, gender, marital status, and physical and mental health status, can influence individuals' experiences within a specific environment [2,3]. It is important to note, however, that the literature has not consistently established clear associations for age, gender, and marital status [3]. According to our results, older age, being female, being voluntarily admitted, and reporting a good physical and mental quality of life are associated with higher levels of patient experience of the care environment. As previously described [85], women reported higher levels of experience than men. Likewise, older people tend to be more accommodating, perhaps because they have fewer expectations than younger people [86]. Also, contrary to what might be expected, voluntarily admitted patients reported higher levels of experience than outpatients, although some patients reported a preference for community mental health treatment, which they considered less stigmatizing [87] and compatible with professional and social functioning. Furthermore, the literature has shown that hospitalization, particularly in the context of involuntary admission, can have a negative impact on patient experience [3], because it can be experienced as traumatic or particularly stressful for patients [88].

However, our results suggest that patients voluntarily admitted to the hospital may have a more holistic and structured experience compared to outpatients, conducive to positive therapeutic relationships with staff, whereas constraint has a negative effect on therapeutic relationships in the case of involuntarily admitted patients [88-90]. Finally, a positive but weak association was found between higher levels of patient experience and better QoL, as previously reported in other studies [51]. A calm and welcoming care environment contributes to patients feeling more comfortable and safer [50], which can reduce stress and anxiety and enhance relationships with staff, thereby promoting patients' recovery. Participants completing the online survey reported a poorer experience of the care environment than participants in health care settings because the latter may be more favorable due to fear of a negative effect on their relationships with staff, or this difference may be due to a possible recall bias.

The most poorly rated items by patients were related to accessing equipment (CE10), media (CE12), and food (CE15). Difficulties with access to equipment (eg, for art therapy) and media (eg, televisions or computers) are related to boredom,

isolation, frustration, and higher levels of distress in patients [25]. A variety of individual or group activities could be offered to patients, such as therapeutic workshops in self-expression (ie, writing), art (ie, photography or painting), psychosocial rehabilitation (ie, cooking, which may also improve diet habits), or body awareness (ie, sophrology), to help patients develop social skills and promote social reintegration, improve confidence and self-esteem, build emotional resilience, and enjoy themselves. Facilities should have basic amenities such as affordable Wi-Fi and a working television in a common room accessible to all patients. Likewise, rooms should be equipped with a minimum package of free channels; however, not all facilities are equally equipped, and the cost of access to Wi-Fi and pay television channels can vary by as much as 2-fold. The content of what is broadcast on television should also be a therapeutic consideration. For example, it seems logical to avoid broadcasting distressing news or uninspiring programs and to favor the broadcasting of cultural works that could be the object of an exchange after viewing, such as a film club. The use of cell phones in health care settings presents challenges in terms of the potential risk of theft or breakage, as well as concerns about maintaining confidentiality. Additionally, it can be a source of tension with staff (eg, if the telephone credit is exceeded). There is no law that prohibiting the use of cell phones because communicating is a fundamental individual freedom, but the internal rules of the facilities can regulate their use by specifying the times and places of use and prohibit taking pictures of patients and staff. Furthermore, psychiatrists may occasionally prohibit a patient from keeping a cell phone, computer, or tablet as part of a medical decision, particularly in the case of placement in a seclusion room or for medical conditions. Previous studies have shown that a healthy diet is essential for good mental health and can prevent the worsening of symptoms [91,92], and that patients' satisfaction with hospital food services strongly influences their overall satisfaction with hospital care [93]. Diets such as the Mediterranean diet have been shown to improve patient outcomes [91]. Providing a menu tailored to patient preferences while focusing on food quality (taste, presentation, flavor, preparation, and variety), as well as the hospital environment, will help improve inpatient appetite and satisfaction [93]. In summary, the current challenges of hospital food service are to transition to a diet that is lower in meat, closer to the Mediterranean diet, without plastic packaging, and low in processed products while increasing the attractiveness of local and seasonal products, all while maintaining costs [91]. By contrast, the most highly rated items by patients were related to spatial accessibility (CE1), cleanliness (CE8), and smoking ban (CE16). Although health care facilities are under a total smoking ban throughout their whole facilities (including in specifically dedicated "smoking areas" or outside), the reality is often more flexible to accommodate patients who cannot leave the health care facilities, even temporarily (eg, patients under constraint). Proposals for smoking cessation assistance should be systematically offered to patients.

Limitations

Some limitations of this study are worth noting. Our sample size, while relatively modest, was sufficient to obtain accurate

estimates. Current recommendations suggest that at least 300 observations are sufficient when using multiparameter models like the GPCM [38-40]. However, our results showed that the assumptions required for IRT calibration were met and that the model fit was adequate. In addition, some DIF analyses comparing subgroups with sample sizes smaller than those recommended for DIF analyses (at least 200 observations per group [94]) may have lacked the statistical power to detect a statistically significant DIF. These DIF findings should be regarded as preliminary, and future work with a larger sample will allow us to confirm these results. Although participants from the online survey and those from health care settings may have reported different levels of experience, this mixed survey design was chosen to ensure inclusivity across various subgroups, as supported by previous research on the equivalence of administration methods [95]. DIF analysis revealed that none of the items was flagged with a large DIF magnitude according to the patient's mode of study participation, suggesting that the data can be pooled without substantial bias. It was not possible to calculate a participation rate or to compare the characteristics of respondents and nonrespondents. This study was widely disseminated nationally, and our sample included inpatients and outpatients with diverse characteristics from different geographic regions of the country. Patients self-reported their diagnosis, and some data (488/498, 2.5%) were missing. However, the risk of misdiagnosis is considered minimal because all participants were fully informed about the study scope and diagnostic criteria. Additionally, this approach closely mirrors the real-world conditions of PREMIUM use. The title of the study mentioned general experience of care to limit the self-selection bias of patients with extreme care environment experiences. Future work will confirm the generalizability of our results. PREMIUM-CE has greater measurement accuracy for patients with scores between –2.1 and +0.7 (ie, reporting low to moderate levels of experience), and thus more items are needed to estimate scores for patients at both ends of the latent continuum. Future work should also reevaluate the precision and accuracy of the CAT in an independent sample and under real-world conditions. Finally, criterion validity could not be assessed because, to our knowledge, no gold standard was available and evidence for construct validity was limited. Future validation studies should examine the relationship between this new measure and objective assessments of the care environment (eg, evaluation by architects or other professionals).

Conclusion

The PREMIUM-CE item bank and its CAT version have demonstrated strong psychometric properties, making them robust measures for assessing patient experience of the care environment, applicable in both outpatient and inpatient settings, for adults with SMI. These measures contribute to the current landscape of patient experience measures by providing a valuable complement to PREMIUM measures of what really matters to patients.

Authors' Contributions

LB was responsible for the conceptualization, supervision, project administration, and funding acquisition. SF was responsible for the methodology, formal analysis, and writing the original draft. LB, SF, and GF are responsible for review and editing. All the authors read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Indicators of psychometric performance, parameter estimates (discrimination and thresholds) and fit statistics, differential item functioning results, and list of the 13-item of the Patient-Reported Experience Measure for Improving Quality of Care in Mental Health for Care Environment (PREMIUM-CE) item bank (English and French versions).
[DOCX File , 27 KB-Multimedia Appendix 1]

References

1. Nicholls D, Kidd K, Threader J, Hungerford C. The value of purpose built mental health facilities: use of the Ward Atmosphere Scale to gauge the link between milieu and physical environment. Int J Ment Health Nurs. 2015;24(4):286-294. [doi: 10.1111/inm.12138] [Medline: 25975388]
2. Daigle K, Frankel L. Collaboration to improve experience in hospital environments. In: Di Bucchianico G, Shin CS, Shim S, Fukuda S, Montagna G, Carvalho C, editors. Advances in Industrial Design. Cham, Switzerland. Springer; 2020:403-409.
3. Woodward S, Berry K, Bucci S. A systematic review of factors associated with service user satisfaction with psychiatric inpatient services. J Psychiatr Res. 2017;92:81-93. [doi: 10.1016/j.jpsychires.2017.03.020] [Medline: 28412601]

4.   Fernandes S, Fond G, Zendjidjian XY, Baumstarck K, Lançon C, Berna F, et al. Measuring the patient experience of mental health care: a systematic and critical review of patient-reported experience measures. Patient Prefer Adherence. 2020;14:2147-2161. [FREE Full text] [doi: 10.2147/PPA.S255264] [Medline: 33192054]

5.   Fernandes S, Fond G, Zendjidjian X, Michel P, Lançon C, Berna F, et al. A conceptual framework to develop a patient-reported experience measure of the quality of mental health care: a qualitative study of the PREMIUM project in France. J Mark Access Health Policy. 2021;9(1):1885789. [FREE Full text] [doi: 10.1080/20016689.2021.1885789] [Medline: 33680364]

6.   Henriksen K, Isaacson S, Sadler BL, Zimring CM. The role of the physical environment in crossing the quality chasm. Jt Comm J Qual Patient Saf. 2007;33(Suppl 11):68-80. [doi: 10.1016/s1553-7250(07)33114-0] [Medline: 18173167]

7.   Zhao Y, Mourshed M. Patients' perspectives on the design of hospital outpatient areas. Buildings. 2017;7(4):117. [FREE Full text] [doi: 10.3390/buildings7040117]

8.   Connellan K, Gaardboe M, Riggs D, Due C, Reinschmidt A, Mustillo L. Stressed spaces: mental health and architecture. HERD. 2013;6(4):127-168. [doi: 10.1177/193758671300600408] [Medline: 24089185]

9.   Jovanović N, Miglietta E, Podlesek A, Malekzadeh A, Lasalvia A, Campbell J, et al. Impact of the hospital built environment on treatment satisfaction of psychiatric in-patients. Psychol Med. 2022;52(10):1969-1980. [doi: 10.1017/S0033291720003815] [Medline: 33087185]

10.   Ulrich RS, Zimring C, Zhu X, DuBose J, Seo HB, Choi YS, et al. A review of the research literature on evidence-based healthcare design. HERD. 2008;1(3):61-125. [doi: 10.1177/193758670800100306] [Medline: 21161908]

11.   Jamshidi S, Parker JS, Hashemi S. The effects of environmental factors on the patient outcomes in hospital environments: a review of literature. Front Archit Res. 2020;9(2):249-263. [FREE Full text] [doi: 10.1016/j.foar.2019.10.001]

12.   Kelly EL, Davis L, Mendon S, Kiger H, Murch L, Pancake L, et al. Provider and consumer perspectives of community mental health services: implications for consumer-driven care. Psychol Serv. 2019;16(4):572-584. [doi: 10.1037/ser0000244] [Medline: 29722997]

13.   Ghaffari F, Shabak M, Norouzi N, Fallah SN. Hospital salutogenic public spaces: a conceptual framework of effective perceptional environment quality components on patients' satisfaction. Int J Build Pathol Adapt. 2021;41(5):965-987. [doi: 10.1108/ijbpa-05-2021-0071]

14.   Colloca L, Barsky AJ. Placebo and nocebo effects. N Engl J Med. 2020;382(6):554-561. [doi: 10.1056/NEJMra1907805] [Medline: 32023375]

15.   Hartogsohn I. Set and setting, psychedelics and the placebo response: an extra-pharmacological perspective on psychopharmacology. J Psychopharmacol. 2016;30(12):1259-1267. [doi: 10.1177/0269881116677852] [Medline: 27852960]

16.   Jovanović N, Campbell J, Priebe S. How to design psychiatric facilities to foster positive social interaction—a systematic review. Eur Psychiatry. 2019;60:49-62. [FREE Full text] [doi: 10.1016/j.eurpsy.2019.04.005] [Medline: 31112827]

17.   Ulrich RS, Bogren L, Gardiner SK, Lundin S. Psychiatric ward design can reduce aggressive behavior. J Environ Psychol. 2018;57:53-66. [doi: 10.1016/j.jenvp.2018.05.002]

18.   Shepley MM, Watson A, Pitts F, Garrity A, Spelman E, Kelkar J, et al. Mental and behavioral health environments: critical considerations for facility design. Gen Hosp Psychiatry. 2016;42:15-21. [doi: 10.1016/j.genhosppsych.2016.06.003] [Medline: 27638966]

19.   Borge L, Fagermoen MS. Patients' core experiences of hospital treatment: wholeness and self-worth in time and space. J Ment Health. 2009;17(2):193-205. [doi: 10.1080/09638230701505996]

20.   Staniszewska S, Mockford C, Chadburn G, Fenton SJ, Bhui K, Larkin M, et al. Experiences of in-patient mental health services: systematic review. Br J Psychiatry. 2019;214(6):329-338. [FREE Full text] [doi: 10.1192/bjp.2019.22] [Medline: 30894243]

21.   Shattell MM, Andes M, Thomas SP. How patients and nurses experience the acute care psychiatric environment. Nurs Inq. 2008;15(3):242-250. [doi: 10.1111/j.1440-1800.2008.00397.x] [Medline: 18786217]

22.   Molin J, Graneheim UH, Lindgren BM. Quality of interactions influences everyday life in psychiatric inpatient care--patients' perspectives. Int J Qual Stud Health Well-being. 2016;11:29897. [FREE Full text] [doi: 10.3402/qhw.v11.29897] [Medline: 26806313]

23.   Lindgren BM, Aminoff C, Graneheim UH. Features of everyday life in psychiatric inpatient care for self-harming: an observational study of six women. Issues Ment Health Nurs. 2015;36(2):82-88. [doi: 10.3109/01612840.2014.941077] [Medline: 25625707]

24.   Lilja L, Hellzén O. Former patients' experience of psychiatric care: a qualitative investigation. Int J Ment Health Nurs. 2008;17(4):279-286. [doi: 10.1111/j.1447-0349.2008.00544.x] [Medline: 18666911]

25.   Folke F, Hursti T, Kanter JW, Arinell H, Tungström S, Söderberg P, et al. Exploring the relationship between activities and emotional experience using a diary in a mental health inpatient setting. Int J Ment Health Nurs. 2018;27(1):276-286. [doi: 10.1111/inm.12318] [Medline: 28220616]

26.   Lindgren BM, Ringnér A, Molin J, Graneheim UH. Patients' experiences of isolation in psychiatric inpatient care: insights from a meta-ethnographic study. Int J Ment Health Nurs. 2019;28(1):7-21. [doi: 10.1111/inm.12519] [Medline: 29975446]

27. Molin J, Graneheim UH, Ringnér A, Lindgren BM. From ideals to resignation—interprofessional teams perspectives on everyday life processes in psychiatric inpatient care. J Psychiatr Ment Health Nurs. 2016;23(9-10):595-604. [doi: 10.1111/jpm.12349] [Medline: 27785861]

28. Donald F, Duff C, Lee S, Kroschel J, Kulkarni J. Consumer perspectives on the therapeutic value of a psychiatric environment. J Ment Health. 2015;24(2):63-67. [doi: 10.3109/09638237.2014.954692] [Medline: 25915815]

29. Stewart D, Burrow H, Duckworth A, Dhillon J, Fife S, Kelly S, et al. Thematic analysis of psychiatric patients' perceptions of nursing staff. Int J Ment Health Nurs. 2015;24(1):82-90. [doi: 10.1111/inm.12107] [Medline: 25382159]

30. Molin J, Strömbäck M, Lundström M, Lindgren BM. It's not just in the walls: patient and staff experiences of a new spatial design for psychiatric inpatient care. Issues Ment Health Nurs. 2021;42(12):1114-1122. [FREE Full text] [doi: 10.1080/01612840.2021.1931585] [Medline: 34142934]

31. Lindgren BM, Molin J, Lundström M, Strömbäck M, Renberg ES, Ringnér A. Does a new spatial design in psychiatric inpatient care influence patients' and staff's perception of their care/working environment? A study protocol of a pilot study using a single-system experimental design. Pilot Feasibility Stud. 2018;4:191. [FREE Full text] [doi: 10.1186/s40814-018-0383-4] [Medline: 30607254]

32. Laursen J, Danielsen A, Rosenberg J. Effects of environmental design on patient outcome: a systematic review. HERD. 2014;7(4):108-119. [doi: 10.1177/193758671400700410] [Medline: 25303431]

33. Johansson IM, Skärsäter I, Danielson E. The meaning of care on a locked acute psychiatric ward: patients' experiences. Nord J Psychiatry. 2009;63(6):501-507. [doi: 10.3109/08039480903118208] [Medline: 19688635]

34. Baumstarck K, Boyer L, Boucekine M, Aghababian V, Parola N, Lançon C, et al. Self-reported quality of life measure is reliable and valid in adult patients suffering from schizophrenia with executive impairment. Schizophr Res. 2013;147(1):58-67. [doi: 10.1016/j.schres.2013.03.008] [Medline: 23566495]

35. Reininghaus U, Priebe S. Measuring patient-reported outcomes in psychosis: conceptual and methodological review. Br J Psychiatry. 2012;201(4):262-267. [FREE Full text] [doi: 10.1192/bjp.bp.111.107615] [Medline: 23028084]

36. Fernandes S, Fond G, Zendjidjian X, Michel P, Baumstarck K, Lancon C, et al. The Patient-Reported Experience Measure for Improving qUality of care in Mental health (PREMIUM) project in France: study protocol for the development and implementation strategy. Patient Prefer Adherence. 2019;13:165-177. [FREE Full text] [doi: 10.2147/PPA.S172100] [Medline: 30718945]

37. Schürhoff F, Fond G, Berna F, Bulzacka E, Vilain J, Capdevielle D, et al. A national network of schizophrenia expert centres: an innovative tool to bridge the research-practice gap. Eur Psychiatry. 2015;30(6):728-735. [doi: 10.1016/j.eurpsy.2015.05.004] [Medline: 26072427]

38. Dai S, Vo TT, Kehinde OJ, He H, Xue Y, Demir C, et al. Performance of polytomous IRT models with rating scale data: an investigation over sample size, instrument length, and missing data. Front Educ. 2021;6:721963. [FREE Full text] [doi: 10.3389/feduc.2021.721963]

39. Cappelleri JC, Lundy JJ, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. Clin Ther. 2014;36(5):648-662. [FREE Full text] [doi: 10.1016/j.clinthera.2014.04.006] [Medline: 24811753]

40. Nguyen TH, Han HR, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. Patient. 2014;7(1):23-35. [FREE Full text] [doi: 10.1007/s40271-013-0041-0] [Medline: 24403095]

41. Ware J, Kosinski M, Keller S. How to score the SF-12 physical and mental health summary scales, 2nd edition. The Health Institute, New England Medical Center. Boston, MA.; 1995. URL: https://www.researchgate.net/profile/John-Ware-6/publication/291994160_How_to_score_SF-12_items/links/58dfc42f92851c369548e04e/How-to-score-SF-12-items.pdf [accessed 2023-02-07]

42. Salyers MP, Bosworth HB, Swanson JW, Lamb-Pagone J, Osher FC. Reliability and validity of the SF-12 health survey among people with severe mental illness. Med Care. 2000;38(11):1141-1150. [doi: 10.1097/00005650-200011000-00008] [Medline: 11078054]

43. Embretson SE, Reise SP. Item Response Theory for Psychologists. Mahwah, NJ. Lawrence Erlbaum Associates; 2000.

44. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care. 2007;45(5 Suppl 1):S22-S31. [FREE Full text] [doi: 10.1097/01.mlr.0000250483.85507.04] [Medline: 17443115]

45. Muraki E. A generalized partial credit model: application of an EM algorithm. Appl Psychol Meas. 1992;16(2):159-176. [doi: 10.1177/014662169201600206]

46. Masters GN. A rasch model for partial credit scoring. Psychometrika. 1982;47(2):149-174. [doi: 10.1007/BF02296272]

47. Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. Psychometrika. 1981;46(4):443-459. [doi: 10.1007/bf02293801]

48. Zieky M. Practical questions in the use of DIF statistics in test development. In: Wainer H, Holland PW, editors. Differential Item Functioning. Hillsdale, NJ. Lawrence Erlbaum Associates; 1993:337-347.

49. Rogers HJ. Differential item functioning. In: Everitt BS, Howell DC, editors. Encyclopedia of Statistics in Behavioral Science. Chichester, UK. John Wiley & Sons; 2005:485-490.

50. Chen H, Li M, Wang J, Xue C, Ding T, Nong X, et al. Factors influencing inpatients' satisfaction with hospitalization service in public hospitals in Shanghai, People's Republic of China. Patient Prefer Adherence. 2016;10:469-477. [FREE Full text] [doi: 10.2147/PPA.S98095] [Medline: 27110101]

51. Priebe S, Miglietta E. Assessment and determinants of patient satisfaction with mental health care. World Psychiatry. 2019;18(1):30-31. [FREE Full text] [doi: 10.1002/wps.20586] [Medline: 30600613]

52. Harvill LM. Standard error of measurement. Educ Meas Issues Pract. 1991;10(2):33-41. [doi: 10.1111/j.1745-3992.1991.tb00195.x]

53. Choi SW, Swartz RJ. Comparison of CAT item selection criteria for polytomous items. Appl Psychol Meas. 2009;33(6):419-440. [FREE Full text] [doi: 10.1177/0146621608327801] [Medline: 20011456]

54. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334. [doi: 10.1007/BF02310555]

55. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res. 1992;21(2):230-258. [doi: 10.1177/0049124192021002005]

56. Kline RB. Principles and Practice of Structural Equation Modeling, 2nd Edition. New York, NY. Guilford Press; 2005.

57. Bjorner JB, Kosinski M, Ware JE. Calibration of an item pool for assessing the burden of headaches: an application of item response theory to the Headache Impact Test (HIT). Qual Life Res. 2003;12(8):913-933. [doi: 10.1023/a:1026163113446] [Medline: 14651412]

58. Fliege H, Becker J, Walter OB, Bjorner JB, Klapp BF, Rose M. Development of a Computer-Adaptive Test for depression (D-CAT). Qual Life Res. 2005;14(10):2277-2291. [doi: 10.1007/s11136-005-6651-9] [Medline: 16328907]

59. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974;19(6):716-723. [doi: 10.1109/TAC.1974.1100705]

60. Schwarz G. Estimating the dimension of a model. Ann Statist. 1978;6(2):461-464. [FREE Full text] [doi: 10.1214/aos/1176344136]

61. Ware JE, Bjorner JB, Kosinski M. Practical implications of item response theory and computerized adaptive testing: a brief summary of ongoing studies of widely used headache impact scales. Med Care. 2000;38(Suppl 9):II73-I182. [Medline: 10982092]

62. Baker FB. The Basics of Item Response Theory, 2nd Edition. Washington, DC. ERIC Clearinghouse on Assessment and Evaluation; 2001.

63. Chang HH, Ying Z. A global information approach to computerized adaptive testing. Appl Psychol Meas. 1996;20(3):213-229. [doi: 10.1177/014662169602000303]

64. Bond TG, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences, 3rd Edition. New York, NY. Lawrence Erlbaum Associates; 2015.

65. Wright B, Linacre JM. Reasonable mean-square fit values. Rasch Meas Trans. 1994;8:370-371. [FREE Full text]

66. Zumbo BD. A handbook on the theory and methods of Differential Item Functioning (DIF): logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores. Directorate of Human Resources Research and Evaluation, Department of National Defense. Ottawa, ON.; 1999. URL: https://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf [accessed 2023-02-20]

67. Choi SW, Reise SP, Pilkonis PA, Hays RD, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Qual Life Res. 2010;19(1):125-136. [FREE Full text] [doi: 10.1007/s11136-009-9560-5] [Medline: 19941077]

68. Michel P, Auquier P, Baumstarck K, Pelletier J, Loundou A, Ghattas B, et al. Development of a cross-cultural item bank for measuring quality of life related to mental health in multiple sclerosis patients. Qual Life Res. 2015;24(9):2261-2271. [doi: 10.1007/s11136-015-0948-0] [Medline: 25712324]

69. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). J Clin Epidemiol. 2008;61(1):17-33. [doi: 10.1016/j.jclinepi.2006.06.025] [Medline: 18083459]

70. Ferketich S. Focus on psychometrics. Aspects of item analysis. Res Nurs Health. 1991;14(2):165-168. [doi: 10.1002/nur.4770140211] [Medline: 2047538]

71. Chalmers RP. Mirt: a Multidimensional Item Response Theory package for the R Environment. J Stat Softw. 2012;48(6):1-29. [FREE Full text] [doi: 10.18637/jss.v048.i06]

72. Choi SW, Gibbons LE, Crane PK. Lordif: an R Package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. J Stat Softw. 2011;39(8):1-30. [FREE Full text] [doi: 10.18637/jss.v039.i08] [Medline: 21572908]

73. Dueber D. Package 'BifactorIndicesCalculator': a package for computing statistical indices relevant to bifactor measurement models. CRAN. 2022. URL: https://cran.r-project.org/web/packages/BifactorIndicesCalculator/BifactorIndicesCalculator.pdf [accessed 2023-03-16]

74. Chalmers RP. Generating adaptive and non-adaptive test interfaces for multidimensional item response theory applications. J Stat Softw. 2016;71(5):1-38. [FREE Full text] [doi: 10.18637/jss.v071.i05]

75. Elf M, Nordin S, Wijk H, Mckee KJ. A systematic review of the psychometric properties of instruments for assessing the quality of the physical environment in healthcare. J Adv Nurs. 2017;73(12):2796-2816. [FREE Full text] [doi: 10.1111/jan.13281] [Medline: 28207946]

76. Reise SP, Scheines R, Widaman KF, Haviland MG. Multidimensionality and structural coefficient bias in structural equation modeling: a bifactor perspective. Educ Psychol Meas. 2012;73(1):5-26. [doi: 10.1177/0013164412449831]

77. Cook KF, Kallen MA, Amtmann D. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. Qual Life Res. 2009;18(4):447-460. [FREE Full text] [doi: 10.1007/s11136-009-9464-4] [Medline: 19294529]

78. Kwan YH, Uy EJ, Bautista DC, Xin X, Xiao Y, Lee GL, et al. Development and calibration of a novel positive mindset item bank to measure Health-Related Quality of Life (HRQoL) in Singapore. PLoS One. 2019;14(7):e0220293. [FREE Full text] [doi: 10.1371/journal.pone.0220293] [Medline: 31365554]

79. Haley SM, Fragala-Pinkham MA, Dumas HM, Ni P, Gorton GE, Watson K, et al. Evaluation of an item bank for a computerized adaptive test of activity in children with cerebral palsy. Phys Ther. 2009;89(6):589-600. [FREE Full text] [doi: 10.2522/ptj.20090007] [Medline: 19423642]

80. Crins MHP, Terwee CB, Klausch T, Smits N, de Vet HCW, Westhovens R, et al. The Dutch-Flemish PROMIS physical function item bank exhibited strong psychometric properties in patients with chronic pain. J Clin Epidemiol. 2017;87:47-58. [doi: 10.1016/j.jclinepi.2017.03.011] [Medline: 28363734]

81. Pilkonis PA, Yu L, Dodds NE, Johnston KL, Lawrence SM, Hilton TF, et al. An item bank for abuse of prescription pain medication from the Patient-Reported Outcomes Measurement Information System (PROMIS®). Pain Med. 2017;18(8):1516-1527. [FREE Full text] [doi: 10.1093/pm/pnw233] [Medline: 28339555]

82. Rindestig FC, Wiberg M, Chaplin JE, Henje E, Dennhag I. Psychometrics of three Swedish physical pediatric item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS)®: pain interference, fatigue, and physical activity. J Patient Rep Outcomes. 2021;5(1):105. [FREE Full text] [doi: 10.1186/s41687-021-00382-2] [Medline: 34637029]

83. Terwee CB, Crins MHP, Boers M, de Vet HCW, Roorda LD. Validation of two PROMIS item banks for measuring social participation in the Dutch general population. Qual Life Res. 2019;28(1):211-220. [FREE Full text] [doi: 10.1007/s11136-018-1995-0] [Medline: 30203302]

84. Kandel H, Khadka J, Goggin M, Pesudovs K. Patient-reported outcomes for assessment of quality of life in refractive error: a systematic review. Optom Vis Sci. 2017;94(12):1102-1119. [doi: 10.1097/OPX.0000000000001143] [Medline: 29095758]

85. Olusina AK, Ohaeri JU, Olatawura MO. Patient and staff satisfaction with the quality of in-patient psychiatric care in a Nigerian general hospital. Soc Psychiatry Psychiatr Epidemiol. 2002;37(6):283-288. [doi: 10.1007/s00127-002-0548-5] [Medline: 12111034]

86. Lundqvist LO, Ahlström G, Wilde-Larsson B, Schröder A. The patient's view of quality in psychiatric outpatient care: patients' view of psychiatric care. J Psychiatr Ment Health Nurs. 2012;19(7):629-637. [doi: 10.1111/j.1365-2850.2012.01899.x] [Medline: 22385022]

87. Beckers T, Koekkoek B, Tiemens B, Jaeqx-van Tienen L, Hutschemaekers G. Substituting specialist care for patients with severe mental illness with primary healthcare. Experiences in a mixed methods study. J Psychiatr Ment Health Nurs. 2019;26(1-2):1-10. [doi: 10.1111/jpm.12499] [Medline: 30270481]

88. Zendjidjian XY, Auquier P, Lançon C, Loundou A, Parola N, Faugère M, et al. Determinants of patient satisfaction with hospital health care in psychiatry: results based on the SATISPSY-22 questionnaire. Patient Prefer Adherence. 2014;8:1457-1464. [FREE Full text] [doi: 10.2147/PPA.S67641] [Medline: 25368515]

89. Roche E, Madigan K, Lyne JP, Feeney L, O'Donoghue B. The therapeutic relationship after psychiatric admission. J Nerv Ment Dis. 2014;202(3):186-192. [doi: 10.1097/NMD.0000000000000102] [Medline: 24566503]

90. Wyder M, Bland R, Blythe A, Matarasso B, Crompton D. Therapeutic relationships and involuntary treatment orders: service users' interactions with health-care professionals on the ward. Int J Ment Health Nurs. 2015;24(2):181-189. [doi: 10.1111/inm.12121] [Medline: 25628260]

91. Fond G, Young AH, Godin O, Messiaen M, Lançon C, Auquier P, et al. Improving diet for psychiatric patients: high potential benefits and evidence for safety. J Affect Disord. 2020;265:567-569. [doi: 10.1016/j.jad.2019.11.092] [Medline: 31757621]

92. Gill R, Tyndall SF, Vora D, Hasan R, Megna JL, Leontieva L. Diet quality and mental health amongst acute inpatient psychiatric patients. Cureus. 2021;13(1):e12434. [FREE Full text] [doi: 10.7759/cureus.12434] [Medline: 33552754]

93. Messina G, Fenucci R, Vencia F, Niccolini F, Quercioli C, Nante N. Patients' evaluation of hospital foodservice quality in Italy: what do patients really value? Public Health Nutr. 2013;16(4):730-737. [FREE Full text] [doi: 10.1017/S1368980012003333] [Medline: 22874795]

94. Scott NW, Fayers PM, Aaronson NK, Bottomley A, de Graeff A, Groenvold M, et al. Differential Item Functioning (DIF) analyses of health-related quality of life instruments using logistic regression. Health Qual Life Outcomes. 2010;8:81. [FREE Full text] [doi: 10.1186/1477-7525-8-81] [Medline: 20684767]

95. Hagan TL, Belcher SM, Donovan HS. Mind the mode: differences in paper vs. web-based survey modes among women with cancer. J Pain Symptom Manage. 2017;54(3):368-375. [FREE Full text] [doi: 10.1016/j.jpainsymman.2017.07.005] [Medline: 28711752]

Abbreviations

CAT: computerized adaptive test
CFA: confirmatory factor analysis
DIF: differential item functioning
GPCM: generalized partial credit model
IRT: item response theory
MDD: major depressive disorder
PREMIUM: Patient-Reported Experience Measure for Improving Quality of Care in Mental Health
PREMIUM-CE: Patient-Reported Experience Measure for Improving Quality of Care in Mental Health for Care Environment
QoL: quality of life
RMSEA: root mean square error of approximation
SEM: standard error of measurement
SF-12: 12-item Short Form
SMI: severe mental illness
VAS: visual analog scale

XSL•FO
RenderX