

An Adaptation of the Profile of Mood States for Use in Adults With Phenylketonuria

Journal of Inborn Errors of Metabolism
& Screening
2016, Volume 4: 1–11
© The Author(s) 2016
DOI: 10.1177/2326409816669373
iem.sagepub.com



Elizabeth D. Bacci, PhD¹, Kathleen W. Wyrwich, PhD²,
Katharine S. Gries, PhD¹, Yinpu Chen, PhD³, Rishabh Jain, MS³,
Laurel Konkol, MS³, Markus J. Merilainen, MD³,
and Haoling H. Weng, MD³

Abstract

Adults with phenylketonuria (PKU) experience disturbances in mood. This study used qualitative and quantitative techniques to adapt the 65-item Profile of Mood States (POMS) for the assessment of key mood domains in adults with PKU. First, cognitive interviews on 58 POMS items (excluding 7 Friendliness domain items) among 15 adults and adolescents (age ≥ 16 years) with PKU were conducted to eliminate items poorly understood or considered irrelevant to PKU; 17 items were removed. Next, the remaining POMS items were quantitatively examined (Mokken scaling and Rasch analysis) in 115 adult patients with PKU. An additional 21 items were removed iteratively, resulting in the 20-item draft PKU-POMS. Finally, the psychometric properties of the draft PKU-POMS were examined. The instrument displayed strong psychometric properties (reliability, validity, and responsiveness) over 6 domains (Anxiety, Depression, Anger, Activity, Tiredness, and Confusion) and all items were well understood in the final cognitive interviews with 10 adults with PKU.

Keywords

phenylketonuria, profile of mood states, patient-reported outcome, psychometric, item reduction, cognitive interview

Introduction

Phenylketonuria (PKU; Online Mendelian Inheritance in Man 261600 and 261630) is a rare autosomal recessive inborn error of metabolism in which the body is deficient in the phenylalanine (Phe) hydroxylase enzyme and as a result is unable to process the amino acid Phe. If untreated, Phe will accumulate in the body and cause mental retardation (in children), microcephaly, delayed speech, delayed social skills, psychiatric symptoms, and behavioral abnormalities.¹ Uncontrolled blood Phe levels in adulthood are also associated with executive dysfunction; lack of concentration; anxiety; depression; and a variety of behavioral, psychiatric, and mood problems.^{2–5} ten Hoedt and colleagues⁶ conducted a randomized double-blind, placebo-controlled cross-over study of Dutch patients with PKU to determine the effect of high Phe levels on mood using a revised version of the Profile of Mood States (POMSr) instrument designed to provide data on 5 categories of mood states: tension, depression, anger, vigor, and fatigue.⁷ Results indicated that when patients were in the high Phe intake period, their overall mood was worse ($p = .017$), and they were more fatigued ($p = .021$) and less vigorous ($p = .006$) than when in

the placebo period, providing evidence of the potential importance of the POMS and mood symptoms in understanding the effects of Phe intake in adults with PKU. However, no patient-reported outcome (PRO) measure of mood symptoms has been developed specifically for patients with PKU.

The original 65-item POMSr PRO, developed in 1971, has been used across various therapeutic areas for the assessment of a respondent's transient and variable moods; however, it has never been used with patients with PKU.⁸ The 65-item POMS is a self-administered evaluation that uses a 5-point Likert scale ranging from 0 (*Not at all*), 1 (*A little*) 2 (*Moderately*), 3 (*Quite*

¹ Evidera, Inc, Seattle, WA, USA

² Evidera Inc, Bethesda, MD, USA

³ BioMarin Pharmaceuticals Inc, Novato, CA, USA

Received May 19, 2016, and in revised form July 26, 2016. Accepted for publication August 5, 2016.

Corresponding Author:

Elizabeth D. Bacci, Evidera, Inc, 1417 4th Avenue, Suite 510, Seattle, WA 98101, USA.

Email: elizabeth.bacci@evidera.com



a bit), to 4 (*Extremely*). Mood states are interpreted through 6 mood domains: tension-anxiety (9 items), depression-dejection (15 items), anger-hostility (12 items), vigor-activity (8 items), fatigue-inertia (7 items), and confusion-bewilderment (7 items). The POMS user manual contains the full list of these items.⁸ The remaining 7 items relate to the friendliness domain; however, subsequent analyses by the POMS developers showed evidence that the friendliness domain is considered too weak to be scored.⁹ Although widely used, the POMS instrument has been subjected to various revisions in order to reduce the number of items or to better adapt the instrument for the assessment of mood symptoms in a specific target patient population¹⁰⁻¹⁴ or specific culture.^{7,15-18} These studies used either qualitative or quantitative research techniques to modify the instrument, resulting in a reduced number of items, the addition of items, or the reconfiguration of the domain structure. Limited validation research has been conducted to assess the psychometric properties of some of the reduced or adapted POMS instruments; no extensive validation research has been conducted on any one version.¹⁹⁻²¹

Aims

The aim of the current study was to use qualitative and quantitative analytic techniques to adapt the 65-item POMS for the assessment of key mood states in adults with PKU and to assess the psychometric properties of this revised instrument, as this instrument was not developed or validated specifically for use in patients with PKU.

Methods

A multistep process that incorporated both qualitative and quantitative analyses was used to develop the PKU-POMS. This process consisted of 3 phases: (1) qualitative assessment of item comprehensibility, acceptability, and relevance to adults with PKU; (2) quantitative assessment of item and domain performance and item reduction; and (3) quantitative and qualitative assessment of the draft items.

Phase 1: Qualitative Assessment of Item Comprehensibility, Acceptability, and Relevance to Adults With PKU

In phase 1, one-on-one qualitative (mixture of in person and telephone) interviews were conducted in the United States among individuals with PKU in order to determine if any of the 58 POMS items (after removal of the seven friendliness items) should be removed due to the lack of comprehension, acceptability, and relevance to patients with PKU.

Patient population. Participants were eligible to participate if they were English speaking, 16 years of age or older, diagnosed with PKU, and were not currently taking or had not taken any medication intended to treat PKU (eg, levodopa, pegvaliase, Kuvan, BH4, and neutral amino acid) in the past 4 weeks.

Participants were recruited through the distribution of study fliers; in person or via e-mail to patients attending the 2014 National PKU Alliance Conference and posted to a study Web site. This study was approved by the Ethical and Independent Review Services Institutional Review Board (<http://www.ean-direview.com/>).

Procedures and analysis. Study staff trained on the conduct of patient interviews conducted one-on-one cognitive debriefing interviews with participating patients using a semi-structured interview guide, which involved asking the participants questions about the 58 POMS items to address the comprehensiveness, acceptability, and relevance to patients with PKU. Following completion of the interviews, 2 study staff members independently reviewed the cognitive debriefing portion of the interview transcripts to determine if any of the POMS items should be removed based on established thresholds. Items were considered for removal if >25% of respondents did not understand the item, did not think the item was acceptable because it was not a “good word to use,” or did not find the item relevant to PKU. Only the items retained through this analysis were included in the analyses in phase 2 (quantitative assessment).

Phase 2: Quantitative Assessment of Item and Domain Performance and Item Reduction

In phase 2, a series of iterative quantitative analyses were conducted in order to assess item and domain performance of the retained POMS items from phase 1, and to remove poorly performing items, as needed, to achieve acceptable measurement properties in patients with PKU. This phase consisted of 3 steps including (1) examination of item-level descriptive and domain structure, (2) item reduction using Mokken scale assessment (MSA), and (3) finally item evaluation using Rasch analysis.

Data source. The pegvaliase (BMN 165) clinical development program was designed to demonstrate the safety and efficacy of pegvaliase in the treatment of adult patients with PKU. Data were used from PRISM-1 study (formerly referred to as BMN 165-301)—a phase 3, open-label, randomized study designed to further characterize the safety of pegvaliase of individuals being treated in induction, titration, and maintenance dose regimens in adults with PKU who have not had previous exposure to pegvaliase (ClinicalTrials.gov: NCT01819727). Study participants were 18 to 70 years old and recruited from 32 study sites in the United States. Individuals were required to have a diagnosis of PKU with blood Phe concentration >600 $\mu\text{mol/L}$ at screening and average blood Phe concentration of >600 $\mu\text{mol/L}$ over the past 6 months per medical history. The analytic population for the current analysis included all enrolled patients who completed a POMS assessment at day 1.

Step 1: Examination of item-level descriptive and domain structure. The frequency and percentage of each response category was

assessed, as well as the mean, percentages of minimum and maximum responses for floor and ceiling effects²² (ie, >50% of participants selecting a response of 0 or 4), and percentage missing. An item was flagged for further investigation and possible removal if it exhibited floor/ceiling effects or excessive missingness (>5% missing), as this is indicative that the range of available response options are not appropriate for the patient population.

To understand the domain structure of the reduced POMS prior to further quantitative analyses to assess the instrument, confirmatory factor analysis (CFA) was conducted. The CFA was conducted to model the fit of the items to each of the 6 POMS domains separately. The following statistics and threshold values were used to evaluate model fit: comparative fit index (CFI) ≥ 0.90 , root mean square error of approximation (RMSEA) ≤ 0.10 , and weighted root mean square residual (WRMR) ≤ 0.05 . Acceptable item fit was demonstrated if items had a factor loading ≥ 0.40 .^{23,24}

Step 2: Item reduction using MSA. The MSA is a nonparametric method of data reduction that uses a probabilistic, hierarchical search procedure to identify the best subset of similar items within a measurement scale or domain of items.²⁵ The model assumes that the data are defined by unidimensionality (ie, 1 latent trait is being measured), local monotonicity (ie, as a person's mood decreases, the chance of giving a response indicating improved mood should never increase), and local independence (ie, responses to any 2 items should be independent and not depend on any other aspects of the respondent or items). Loevinger's coefficient H^{26} was used to assess the scalability of the POMS items within each domain, and the 6 domains within the overall instrument. Mokken²⁷ proposed an interpretation of this scalability coefficient: $H > 0.5$ indicates a strong scale, $0.4 < H < 0.5$ indicates a medium scale, and $0.3 < H < 0.4$ indicates a weak scale. Items with loadings < 0.4 were considered candidates for item removal. An iterative process of item removal was used by removing items within a domain one by one and checking the scalability coefficients for each iterative revision. This process continued until no items had loadings below 0.4 and the scalability coefficient for each domain was considered medium or strong.

Step 3: Final item evaluation using Rasch analysis. Following the Mokken scaling analysis, separate Rasch analyses were conducted using the reduced sets of items within each of the 6 POMS domains. Rasch analysis was selected to determine if each item and domain fit the Rasch measurement model,²⁸ which has scaling properties of linear and interval measurement.²⁹ With the Rasch model, the probability of a positive response is modeled as a logistic function of the simultaneous difference between a patient's PKU severity and the severity that is measured by that item response. The following parameters were examined to determine acceptable Rasch model fit: (1) threshold ordering (ie, determine if each item's response category probability distributions indicate the proper ordering of response category shifts across all categories); (2) residual fit

statistics to assess item redundancy and multidimensionality (ie, evidence that an item is being affected by some other dimension other than the latent trait the scale is measuring); and (3) χ^2 item fit of the observed data to the model. Individual item test of fit showed the χ^2 probability for each item, where items with $P < .05$ were considered ill fitting. Items with high negative residuals (< -3.0) indicated an overfitting item, wherein the information provided by this item did not add additional value to the measure. High positive residuals (> 3.0) indicated that the item was underfitting, indicating that the item had a poor model fit and the response categories were underdiscriminating or not discriminating differences in severity.³⁰ Failure to meet these standard thresholds can be indicative of a flawed item in need of revision or removal.

An iterative series of Rasch models were fit and examined in relation to the thresholds above, such that within each iteration, the previous iteration's most problematic item(s) was removed, continuing as long as there was poor fit of the Rasch model. However, to increase domain reliability, all efforts were made to retain at least 3 items per domain.³¹ The remaining items were deemed the final items to compose the draft POMS for use in patients with PKU (ie, PKU-POMS) subjected to a final quantitative and qualitative confirmatory assessment in phase 3.

Phase 3: Quantitative and Qualitative Assessment of the Draft Items

In phase 3, the psychometric properties of the draft PKU-POMS were assessed using data from PRISM-1 study. These assessments included internal consistency reliability using Cronbach's α (day 1); convergent validity using Pearson's correlations between scores on each PKU-POMS domain and scores on the Attention-Deficit Hyperactivity Disorder Rating Scale-IV (ADHD RS-IV) and plasma Phe values (day 1); and responsiveness using Pearson's correlations between change from day 1 to end of study (EOS) on each PKU-POMS domain and change in the ADHD RS-IV and plasma Phe. Plasma Phe was selected for use in the validation analyses as exploratory, because although it is an objective measure of patients' health in this patient population, it was unknown if changes in mood states would correspond to changes in Phe.

Finally, one-on-one cognitive interviews were conducted via telephone to investigate the acceptability of the draft PKU-POMS in adults with PKU, including the relevance of all retained items as well as the understandability of the instructions, recall period, response options, item content, and domains. The patient population was a separate, but highly similar, patient population of the initial qualitative interviews conducted in phase 1 (with the same inclusion/exclusion criteria required for participation), in order to ensure that the final draft instrument is understood in the same population in which it was created and validated.

Results

Phase 1: Qualitative Assessment of Item Comprehensibility, Acceptability, and Relevance to Patients With PKU

Interviews were conducted with 9 women (60%) and 6 men (40%) with PKU (mean age: 30.4 ± 12.9 years). All participants were diagnosed with PKU during the first year of life, and the majority were currently following a Phe-restricted diet (93.3%). Only 2 participants (13.3%) reported a formal diagnosis of anxiety, and 2 (13.3%) reported a formal diagnosis of depression.

Seventeen POMS items were identified that were not well understood (eg, Muddled, Shaky), not acceptable as a good word to use (eg, Bushed, Bewildered), or not relevant to patients with PKU (eg, Unworthy, Guilty). These 17 items were removed from the proposed PKU-POMS, as detailed in Table 1, resulting in 41 remaining items.

Phase 2: Quantitative Assessment of Item and Domain Performance and Item Reduction

Data from a total of 115 patients with PKU were available for the analysis from PRISM-1 study. Patient demographic characteristics are presented in Table 2.

Step 1. Examination of item-level descriptive and domain structure. Descriptive statistics for the 41 remaining POMS items are shown in Table 3. At day 1, 14 of the 41 items displayed floor effects (ie, item where $\geq 50\%$ of sample responded “not at all”), while no items displayed ceiling effects (ie, item where $\geq 50\%$ of sample responded “extremely”). No items exhibited missingness $>2\%$.

The CFA was conducted using the retained 41 items, yielding estimates of global model fit for each individual POMS domain. Using thresholds for acceptable fit, global model fit was found to be unacceptable for all 6 POMS domains (Table 4). In addition, the items weary, alert, and efficient were found to have low item loadings (ie, <0.40). The findings of lack of domain and item-level fit indicated that these 41 remaining POMS items did not conform to the 6-domain structure of the original POMS in this patient population, and informed the item analyses that followed.

Step 2: Item reduction using MSA. The MSA was conducted to assess item fit to each domain as well as the scalability of each of the 6 domains to the overall model. Using the H coefficient as an indicator of scalability, the depression-dejection and fatigue-inertia domains were found to be strong, tension-anxiety and anger-hostility were found to be medium, and vigor-activity and confusion-bewilderment were found to be weak (column 1 of Table 5).

In addition, each domain included at least 1 item that had a low item loading (ie, $H < 0.40$). Items with low item scalability coefficients were removed from each domain, one at a time,

Table 1. POMS Items Removed From Revised POMS Following PKU Patient Interviews.^a

POMS Domain (Number of Items)	Items Removed
Vigor-activity (8)	Q56. Full of pep Q60. Carefree Q63. Vigorous
Confusion-bewilderment (7)	Q37. Muddled Q50. Bewildered
Tension-anxiety (9)	Q2. Tense Q10. Shaky Q16. On edge Q27. Restless
Fatigue-inertia (7)	Q11. Listless Q65. Bushed
Anger-hostility (12)	Q12. Peeved Q24. Spiteful
Depression-dejection (15)	Q9. Sorry for things done Q18. Blue Q23. Unworthy Q62. Guilty

Abbreviations: PKU, phenylketonuria; POMS, profile of mood states.

^aN = 15

Table 2. Sociodemographic and Clinical Characteristics at Day 1.^a

	Day 1
Age, years	
n	115
Mean (SD)	29.1 (8.8)
Median (range)	28.0 (16–51)
Sex, n (%)	
Male	59 (51.3%)
Female	56 (48.7%)
Race, n (%)	
American Indian or Alaska Native	1 (0.9%)
Black or African American	3 (2.6%)
White or Caucasian	109 (94.8%)
Other	2 (1.7%)
Plasma Phe level at day 1 (μmol/L)	
n	114
Mean (SD)	1217 (389.0)
Median (range)	1211 (510–2330)
ADHD RS-IV at day 1	
n	114
Mean (SD)	9.9 (5.7)
Median (range)	9.5 (0–26)

Abbreviations: ADHD RS-IV, Attention Deficit Hyperactivity Disorder Rating Scale IV; Phe, phenylalanine; SD, standard deviation.

^aN = 115.

and the fit of the domain reassessed. This iterative process continued within each domain separately until all items demonstrated a loading above 0.4 and each domain had an H value indicative of a strong scale (column 2 of Table 5). Specifically, in the first iteration of the MSA analysis, the following items were removed by domain: alert (vigor-activity), relaxed (tension-anxiety), efficient and uncertain about things (confusion-bewilderment), weary (fatigue-inertia), rebellious (anger-hostility), and terrified

Table 3. Descriptive Statistics of POMS at Day 1.^{a,b}

Item	n	Mean (SD)	Frequency, n (%)					Missing
			Not At All (Floor)	A Little	Moderately	Quite a Bit	Extremely (Ceiling)	
Angry	115	1.1 (1.1)	40 (34.8%)	38 (33.0%)	27 (23.5%)	6 (5.2%)	4 (3.5%)	
Worn out	115	2.1 (1.2)	11 (9.6%)	28 (24.3%)	29 (25.2%)	31 (27.0%)	16 (13.9%)	
Unhappy	114	1.0 (1.0)	42 (36.5%)	47 (40.9%)	15 (13.0%)	7 (6.1%)	3 (2.6%)	1 (0.9%)
Lively	115	2.2 (1.0)	6 (5.2%)	21 (18.3%)	42 (36.5%)	33 (28.7%)	13 (11.3%)	
Confused	115	1.0 (1.0)	47 (40.9%)	36 (31.3%)	22 (19.1%)	7 (6.1%)	3 (2.6%)	
Sad	115	1.0 (1.1)	46 (40.0%)	39 (33.9%)	17 (14.8%)	8 (7.0%)	5 (4.3%)	
Active	115	2.4 (1.1)	4 (3.5%)	22 (19.1%)	33 (28.7%)	40 (34.8%)	16 (13.9%)	
Grouchy	115	1.3 (1.2)	33 (28.7%)	40 (34.8%)	26 (22.6%)	8 (7.0%)	8 (7.0%)	
Energetic	115	2.1 (1.1)	10 (8.7%)	23 (20.0%)	41 (35.7%)	28 (24.3%)	13 (11.3%)	
Panicky	115	1.0 (1.2)	50 (43.5%)	33 (28.7%)	18 (15.7%)	8 (7.0%)	6 (5.2%)	
Hopeless	115	0.5 (0.9)	80 (69.6%)	18 (15.7%)	13 (11.3%)	2 (1.7%)	2 (1.7%)	
Relaxed	115	2.1 (1.1)	7 (6.1%)	28 (24.3%)	36 (31.3%)	34 (29.6%)	10 (8.7%)	
Uneasy	115	1.2 (1.1)	37 (32.2%)	34 (29.6%)	29 (25.2%)	10 (8.7%)	5 (4.3%)	
Unable to concentrate	115	1.5 (1.2)	28 (24.3%)	33 (28.7%)	26 (22.6%)	19 (16.5%)	9 (7.8%)	
Fatigued	115	1.9 (1.2)	15 (13.0%)	37 (32.2%)	21 (18.3%)	31 (27.0%)	11 (9.6%)	
Annoyed	115	1.8 (1.2)	17 (14.8%)	29 (25.2%)	36 (31.3%)	23 (20.0%)	10 (8.7%)	
Discouraged	115	1.2 (1.2)	45 (39.1%)	33 (28.7%)	20 (17.4%)	8 (7.0%)	9 (7.8%)	
Resentful	115	0.5 (0.9)	75 (65.2%)	24 (20.9%)	12 (10.4%)	3 (2.6%)	1 (0.9%)	
Nervous	115	1.6 (1.2)	26 (22.6%)	36 (31.3%)	23 (20.0%)	22 (19.1%)	8 (7.0%)	
Lonely	114	0.9 (1.2)	62 (53.9%)	26 (22.6%)	12 (10.4%)	8 (7.0%)	6 (5.2%)	1 (0.9%)
Miserable	115	0.6 (1.1)	79 (68.7%)	18 (15.7%)	7 (6.1%)	8 (7.0%)	3 (2.6%)	
Cheerful	115	2.3 (1.0)	3 (2.6%)	20 (17.4%)	40 (34.8%)	39 (33.9%)	13 (11.3%)	
Bitter	115	0.5 (0.8)	69 (60.0%)	31 (27.0%)	13 (11.3%)	2 (1.7%)		
Exhausted	115	1.8 (1.3)	19 (16.5%)	33 (28.7%)	27 (23.5%)	22 (19.1%)	14 (12.2%)	
Anxious	115	1.8 (1.2)	18 (15.7%)	33 (28.7%)	28 (24.3%)	23 (20.0%)	13 (11.3%)	
Ready to fight	115	0.6 (1.0)	77 (67.0%)	20 (17.4%)	10 (8.7%)	6 (5.2%)	2 (1.7%)	
Gloomy	115	0.8 (1.1)	63 (54.8%)	30 (26.1%)	10 (8.7%)	8 (7.0%)	4 (3.5%)	
Desperate	115	0.3 (0.8)	91 (79.1%)	13 (11.3%)	8 (7.0%)	2 (1.7%)	1 (0.9%)	
Sluggish	115	1.5 (1.3)	33 (28.7%)	30 (26.1%)	24 (20.9%)	20 (17.4%)	8 (7.0%)	
Rebellious	115	0.6 (0.9)	69 (60.0%)	31 (27.0%)	8 (7.0%)	6 (5.2%)	1 (0.9%)	
Helpless	115	0.4 (0.8)	86 (74.8%)	17 (14.8%)	7 (6.1%)	4 (3.5%)	1 (0.9%)	
Wearry	115	0.8 (0.8)	50 (43.5%)	43 (37.4%)	20 (17.4%)	1 (0.9%)	1 (0.9%)	
Alert	114	2.4 (1.0)	3 (2.6%)	19 (16.5%)	37 (32.2%)	42 (36.5%)	13 (11.3%)	1 (0.9%)
Deceived	114	0.5 (0.9)	83 (72.2%)	15 (13.0%)	9 (7.8%)	5 (4.3%)	2 (1.7%)	1 (0.9%)
Furious	114	0.7 (1.2)	77 (67.0%)	16 (13.9%)	11 (9.6%)	2 (1.7%)	8 (7.0%)	1 (0.9%)
Efficient	113	2.1 (1.1)	9 (7.8%)	23 (20.0%)	35 (30.4%)	38 (33.0%)	8 (7.0%)	2 (1.7%)
Bad-tempered	114	0.9 (1.1)	58 (50.4%)	27 (23.5%)	15 (13.0%)	10 (8.7%)	4 (3.5%)	1 (0.9%)
Worthless	113	0.5 (0.9)	84 (73.0%)	12 (10.4%)	11 (9.6%)	6 (5.2%)		2 (1.7%)
Forgetful	114	1.6 (1.3)	24 (20.9%)	35 (30.4%)	30 (26.1%)	12 (10.4%)	13 (11.3%)	1 (0.9%)
Terrified	114	0.4 (0.7)	80 (69.6%)	24 (20.9%)	7 (6.1%)	3 (2.6%)		1 (0.9%)
Uncertain about things	114	1.7 (1.3)	25 (21.7%)	25 (21.7%)	30 (26.1%)	22 (19.1%)	12 (10.4%)	1 (0.9%)

Abbreviations: POMS, profile of mood states; SD, standard deviation.

^aBoldface values indicate items with floor effects (>50%).

^bN = 115.

(depression-dejection). In the second iteration of the MSA analysis, cheerful (vigor-activity) and resentful (anger-hostility) were removed. Finally, in the third iteration of the MSA analysis, only bitter (anger-hostility) was removed. A total of 10 items were removed in step 2, resulting in 31 remaining items.

Step 3: Final item evaluation using Rasch analysis. The fit of the remaining 31 items to the Rasch measurement model was then assessed, separately within each POMS domain (Table 6). This analysis demonstrated that the fit of each item within the vigor-activity and tension-anxiety domains met the established

thresholds for the Rasch measurement model. For the confusion-bewilderment domain, the fit of each item was acceptable according to the fit residuals and χ^2 values; however, the response thresholds for forgetful were found to be disordered. Specifically, moving from a response of 0 (Not at all) to a 1 (A little) and from 1 (A little) to 2 (Moderately) corresponded to a linear increase in the latent trait of experiencing a confused mood state, but moving from a 3 (Quite a bit) to 4 (Extremely) was actually associated with a lesser experience of a confused mood state. By collapsing response options of Extremely with Quite a bit, the thresholds for this item were ordered, and the fit

Table 4. Confirmatory Factor Analysis of POMS at Day 1.

PKU-POMS Domain, Item	Standardized Coefficient or Model Fits (RMSEA, CFI, WRMR) ^a
Vigor-activity	0.14, 0.97, 0.53
Lively	0.72
Active	0.72
Energetic	0.82
Cheerful	0.57
Alert	0.35
Tension-anxiety	0.17, 0.96, 0.66
Panicky	0.79
Relaxed	-0.35
Uneasy	0.69
Nervous	0.82
Anxious	0.77
Confusion-bewilderment	0.00, 1.00, 0.26
Confused	0.62
Unable to concentrate	0.80
Efficient	-0.28
Forgetful	0.87
Uncertain about things	0.59
Fatigue-inertia	0.09, 1.00, 0.39
Worn out	0.84
Fatigued	0.81
Exhausted	0.90
Sluggish	0.73
Weary	0.30
Anger-hostility	0.12, 0.95, 0.86
Angry	0.70
Grouchy	0.71
Annoyed	0.79
Resentful	0.70
Bitter	0.70
Ready to fight	0.76
Rebellious	0.62
Deceived	0.78
Furious	0.84
Bad-tempered	0.80
Depression-dejection	0.08, 0.99, 0.68
Unhappy	0.78
Sad	0.86
Hopeless	0.85
Discouraged	0.86
Lonely	0.84
Miserable	0.85
Gloomy	0.85
Desperate	0.77
Helpless	0.75
Worthless	0.93
Terrified	0.56

Abbreviations: CFI, comparative fit index; EOS, end of study; PKU, phenylketonuria; POMS, profile of mood states; RMSEA, root mean square error of approximation; WRMR, weighted root mean square residual.

^aModel fit thresholds for acceptable fit included CFI \geq 0.90, RMSEA \leq 0.10, and WRMR \leq 0.05.

of all items in this domain found acceptable. For the fatigue-inertia domain, fatigued was also found to have disordered thresholds, but the removal of this item rather than collapsing item response thresholds increased the performance of this domain overall.

For the remaining 2 domains (anger-hostility and depression-dejection), a more iterative series of Rasch analyses were conducted to arrive at best fit due to the large number of problematic items within these domains in fitting the Rasch measurement model (Table 6). The selection of items to remove was based on the performance of each item in relation to established thresholds, the possibility of collapsing response options, and considering the distribution of each item (eg, presence of high floor effects [$>50\%$]; Table 3). In the anger-hostility domain, 4 items demonstrated disordered thresholds in the initial Rasch analysis. To improve fit, all items were removed but the 3 without floor effects (angry, grouchy, and annoyed), but in the resulting model, grouchy remained disordered. After collapsing the response options of quite a bit and extremely in grouchy, an acceptable fit was found for this domain. For depression-dejection, 6 items displayed threshold disordering in the initial Rasch analysis, and 7 displayed high floor effects. The set of 3 items without floor effects (unhappy, sad, and discouraged) fit the Rasch model well with no threshold disordering. In order to retain at least 4 items, all removed items were reintroduced one by one with the 3 remaining items, and the fit of the items and domain reassessed. Only the reincorporation of lonely performed adequately and was included in the final depression-dejection domain.

Overall, 11 items were removed using Rasch analysis; the Rasch model parameters and item fit statistics of the remaining 20 items are provided in Table 6. A final MSA within each domain of the remaining 20 items was also conducted (column 3 of Table 5); all item loadings were acceptable and the scalability coefficient indicated that all items and scales were medium to strong. Given acceptable scalability and fit to the Rasch model, these remaining 20 items were used to compose the draft PKU-POMS. The 6 domains were renamed to reflect the remaining items after removal of items from phases 1 and 2 as follows: anxiety, depression, anger, activity, tiredness, and confusion (formerly tension-anxiety, depression-dejection, anger-hostility, vigor-activity, fatigue-inertia, and confusion-bewilderment, respectively).

Phase 3: Quantitative and Qualitative Assessment of Draft Items

Psychometric assessment of the draft 20-item PKU-POMS indicated that each domain demonstrated high internal consistency (Cronbach's $\alpha = 0.75-0.87$). Pearson's correlations between each domain and the ADHD RS-IV (except activity and depression) were statistically significant. Correlations with plasma Phe levels were not statistically significant (Table 7). Finally, when using correlations with the ADHD RS-IV inattentiveness scores as the primary analysis of responsiveness, all PKU-POMS domains (except depression) exhibited responsiveness to change from day 1 to EOS (Table 7). Correlations with plasma Phe were treated as exploratory and demonstrated support for the responsiveness of anxiety, depression, and confusion domains.

Finally, cognitive interviews were conducted with 5 women and 5 men with PKU (mean age: 27.2 ± 4.6 years).

Table 5. Mokken Scaling Assessment Results.^a

PKU-POMS Domain, Item	Initial MSA H Scaling Coefficient ^b	Final MSA (Pre-Rasch Analysis) H Scaling Coefficient ^b	Final MSA (Post-Rasch Analysis) H Scaling Coefficient ^b
Tension-anxiety	Fit of item to domain	Fit of item to domain	Fit of item to domain
Full-scale H	0.46 (0.06)	0.57 (0.06)	0.57 (0.06)
Panicky	0.53 (0.06)	0.56 (0.07)	0.56 (0.07)
Relaxed	0.28 ^c (0.08)	— ^c	— ^c
Uneasy	0.47 (0.07)	0.51 (0.07)	0.51 (0.07)
Nervous	0.52 (0.06)	0.60 (0.06)	0.60 (0.06)
Anxious	0.49 (0.06)	0.60 (0.06)	0.59 (0.06)
Vigor-activity			
Full-scale H	0.39 (0.06)	0.55 (0.07)	0.54 (0.07)
Lively	0.43 (0.07)	0.49 (0.09)	0.48 (0.09)
Active	0.41 (0.06)	0.55 (0.07)	0.54 (0.07)
Energetic	0.47 (0.06)	0.61 (0.06)	0.60 (0.06)
Cheerful	0.38 (0.07)	— ^d	— ^d
Alert	0.24 ^c (0.08)	— ^c	— ^c
Confusion-bewilderment			
Full-scale H	0.39 (0.06)	0.55 (0.07)	0.55 (0.07)
Confused	0.40 (0.07)	0.49 (0.1)	0.49 (0.10)
Unable to concentrate	0.47 (0.06)	0.58 (0.07)	0.58 (0.07)
Efficient	0.21 ^c (0.08)	— ^c	— ^c
Forgetful	0.50 (0.05)	0.59 (0.07)	0.59 (0.06)
Uncertain about things	0.37 ^c (0.07)	— ^c	— ^c
Fatigue-inertia			
Full-scale H	0.53 (0.05)	0.65 (0.05)	0.67 (0.05)
Worn out	0.59 (0.05)	0.69 (0.05)	0.71 (0.05)
Fatigued	0.57 (0.05)	0.63 (0.06)	— ^e
Exhausted	0.62 (0.04)	0.69 (0.05)	0.68 (0.06)
Sluggish	0.53 (0.06)	0.58 (0.06)	0.63 (0.06)
Weary	0.23 ^c (0.09)	— ^c	— ^c
Anger-hostility			
Full-scale H	0.47 (0.05)	0.53 (0.05)	0.612(0.06)
Angry	0.46 (0.07)	0.51 (0.07)	0.63 (0.06)
Grouchy	0.47 (0.06)	0.49 (0.07)	0.58 (0.08)
Annoyed	0.55 (0.06)	0.57 (0.06)	0.63 (0.06)
Resentful	0.41 (0.07)	— ^d	— ^d
Bitter	0.43 (0.08)	— ^d	— ^d
Ready to fight	0.47 (0.06)	0.49 (0.07)	— ^e
Rebellious	0.39 ^c (0.08)	— ^c	— ^c
Deceived	0.47 (0.07)	0.47 (0.07)	— ^e
Furious	0.51 (0.06)	0.56 (0.07)	— ^e
Bad-tempered	0.53 (0.05)	0.58 (0.05)	— ^e
Depression-dejection			
Full-scale H	0.59 (0.05)	0.63 (0.05)	0.68 (0.06)
Unhappy	0.58 (0.07)	0.61 (0.07)	0.66 (0.07)
Sad	0.65 (0.05)	0.67 (0.05)	0.72 (0.06)
Hopeless	0.62 (0.06)	0.65 (0.06)	— ^e
Discouraged	0.65 (0.06)	0.68 (0.06)	0.66 (0.06)
Lonely	0.61 (0.06)	0.65 (0.06)	0.68 (0.06)
Miserable	0.60 (0.07)	0.63 (0.07)	— ^e
Gloomy	0.63 (0.06)	0.64 (0.06)	— ^e
Desperate	0.50 (0.09)	0.51 (0.09)	— ^e
Helpless	0.52 (0.08)	0.51 (0.09)	— ^e
Worthless	0.65 (0.06)	0.67 (0.06)	— ^e
Terrified	0.32 ^c (0.07)	— ^c	— ^c

Abbreviations: MSA, Mokken scaling assessment; PKU, phenylketonuria; POMS, profile of mood states.

^aN = 110.

^bLow scalability defined as H coefficient <0.40.

^cItems removed after iteration 1 of MSA.

^dItems removed after iterations 2 and 3 of MSA.

^eItems removed through Rasch analysis.

Table 6. Initial and Final Rasch Model Parameters.

PKU-POMS Domain, Item	Category Threshold Parameters				Fit Residual	Item Fit χ^2	χ^2 , P Value
	a1	a2	a3	a4			
A. Initial Rasch Model Parameters Item Fit of 31 Items Retained After Mokken Scaling Analysis^{a,b}							
Vigor-activity							
Lively	-2.45	-0.82	0.91	2.37	1.09	0.51	.77
Active	-2.34	-0.30	0.79	2.95	0.63	1.92	.38
Energetic	-2.62	1.05	0.74	1.82	-0.54	1.84	.40
Tension-anxiety							
Panicky	-0.73	0.24	1.03	1.46	-0.03	0.73	.70
Uneasy	-1.27	-0.34	1.18	1.56	1.20	0.94	.62
Nervous	-1.99	-0.36	-0.11	1.80	-0.12	2.41	.30
Anxious	-2.51	-0.80	-0.20	1.01	-0.34	0.56	.76
Confusion-bewilderment							
Confused	-0.88	0.11	1.33	2.12	1.42	0.88	.64
Unable to concentrate	-1.85	-0.51	0.20	1.26	-0.30	1.97	.37
Forgetful ^c	-2.23	-0.61	0.73	0.33	-0.42	1.70	.43
Fatigue-inertia							
Worn out	-3.25	-0.85	-0.013	1.88	-0.28	2.38	.31
Fatigued ^c	-2.70	0.02	-0.069	2.39	0.55	0.51	.78
Exhausted	-2.08	-0.40	0.57	1.96	-0.44	2.69	.26
Sluggish	-0.89	-0.042	0.71	2.76	1.27	1.45	.48
Anger-hostility							
Angry ^c	-1.78	-0.47	1.32	0.92	0.257	1.77	.41
Grouchy ^c	-2.13	-0.56	0.83	0.06	1.3	1.24	.54
Annoyed	-3.70	-1.56	-0.23	0.70	0.095	1.13	.57
Ready to fight	0.12	0.32	0.34	1.89	-0.676	1.53	.47
Deceived ^c	0.53	0.18	0.54	1.31	-0.161	0.90	.64
Furious ^c	0.24	-0.04	1.62	-1.19	-0.628	2.34	.31
Bad-tempered	-0.85	-0.12	0.17	1.56	-1.232	4.42	.11
Depression-dejection							
Unhappy	-3.28	-0.31	0.09	1.32	0.87	0.73	.69
Sad	-2.84	-0.80	0.03	0.70	-0.37	2.55	.28
Hopeless ^c	-0.73	-0.68	2.16	1.03	-0.46	0.83	.66
Discouraged ^c	-2.88	1.24	-0.11	-0.31	-0.08	0.34	.84
Lonely	-1.80	-0.53	-0.51	0.46	0.24	0.41	.81
Miserable ^c	-0.79	-0.06	-0.53	1.29	-0.02	1.49	.48
Gloomy ^c	-1.84	-0.20	-0.32	0.98	-0.13	1.75	.42
Desperate ^c	0.04	-0.22	1.57	1.24	0.83	0.82	.66
Helpless	-0.41	0.14	0.55	1.47	0.80	0.88	.65
Worthless ^c	-0.22	-0.72	0.66	7.59	-1.33	3.80	.15
B. Final Rasch Model Parameters and Item Fit of 20 Items in Draft PKU-POMS^{b,d}							
Activity							
Lively	-2.45	-0.82	0.91	2.37	1.09	0.51	.77
Active	-2.34	-0.30	0.79	2.95	0.63	1.92	.38
Energetic	-2.62	1.05	0.74	1.82	-0.54	1.84	.40
Anxiety							
Panicky	-0.73	0.24	1.03	1.46	-0.03	0.73	.69
Uneasy	-1.27	-0.34	1.18	1.56	1.20	0.94	.62
Nervous	-1.99	-0.36	-0.11	1.80	-0.128	2.42	.30
Anxious	-2.51	-0.80	-0.17	1.01	-0.34	0.56	.76
Confusion							
Confused	-0.76	0.24	1.548774	2.19	1.36	1.17	.56
Unable to concentrate	-1.71	-0.38	0.365705	1.45	-0.41	1.98	.37
Forgetful	-2.10	-0.45	0.336567	— ^e	-0.08	2.70	.26
Tiredness							
Worn out	-3.41	-0.92	-0.14	2.00	-0.11	3.13	.21
Exhausted	-2.09	-0.47	0.42	2.08	0.32	0.89	.64
Sluggish	-1.04	-0.19	0.75	3.00	1.07	0.55	.76

(continued)

Table 6. (continued)

PKU-POMS Domain, Item	Category Threshold Parameters				Fit Residual	Item Fit χ^2	χ^2 , P Value
	a1	a2	a3	a4			
Anger							
Angry	-1.40	-0.05	2.00	3.19	0.05	1.52	.47
Grouchy	-1.85	-0.04	0.86	— ^e	1.40	0.24	.89
Annoyed	-2.9	1.25	0.12	1.68	-0.08	0.30	.86
Depression							
Unhappy	-2.53	0.14	0.86	2.76	0.74	1.10	.58
Sad	-2.15	-0.22	0.54	2.04	-0.36	2.56	.28
Discouraged	-2.05	-0.66	0.32	0.45	0.18	0.46	.79
Lonely	-1.02	-0.14	0.02	1.65	0.19	1.00	.61

Abbreviations: PKU, phenylketonuria; POMS, profile of mood states.

^aItems examined in Rasch analyses were those retained as acceptable items from Mokken scaling.

^bSeparate Rasch analyses were conducted within each of the 6 domains.

^cIndicates disordered thresholds.

^dItems examined in Rasch analyses were a reduced set from those retained from both Mokken scaling and the first Rasch analysis.

^eThe response option for “extremely” was collapsed with “quite a bit” in final model.

All 10 participants were diagnosed with PKU during the first year of life, and all were currently following a Phe-restricted diet. Of the 10 participants, 4 (40%) reported a formal diagnosis of anxiety and 2 (20%) reported a formal diagnosis of depression. The interviews demonstrated that the final 20 items in the draft PKU-POMS were relevant and well understood by patients with PKU. The instructions, recall period, response options, item content, and domains were found to be comprehensive, clear, and relevant, with no revisions needed.

Discussion

The original 65-item POMS is a widely used instrument for the assessment of mood states, and multiple revised versions of the measure have been created in order to adapt the measure to a specific patient population¹⁰⁻¹⁴ or specific culture.^{7,15-18} However, no known PRO for the assessment of mood states, or revised version of the POMS, has been created for patients with PKU. Additionally, the content validity and psychometric properties of the POMS in patients with PKU has not been investigated previously. Thus, in the current study, a 3-phase, mixed qualitative and quantitative approach was used to assess the content validity, revise the original POMS for use specifically in patients with PKU, and assess the psychometric properties of the new PKU-POMS.

In phase 1, cognitive interviews with adults with PKU were conducted on the POMS to identify and remove items that were poorly understood, thought to be unacceptable as appropriate words to describe the patient’s mood state, or considered irrelevant to patients with PKU. In total, 17 items were removed across all 6 domains from the 58 original POMS items (excluding the 7 friendliness domain items). As noted above, previous investigations conducted on multiple patient populations have reduced the original POMS by removing poorly performing items. All 17 items removed in the present study correspond

to other independent item reduction studies where the same POMS items have been removed, such as sorry for things done,^{7,12,16,32} shaky,^{7,12,14,17} unworthy,^{12-15,17,18} muddled,^{7,13,18,33} and carefree.^{7,12-14,18} These findings provide further support for the removal of the selected items in the current investigation using input from adults and adolescents (≥ 16 years) with PKU.

In phase 2, the remaining items were assessed quantitatively first using MSA and then Rasch analysis. Using MSA, an iterative process was used to remove the worst items within weak domains until all domains and all items achieved at least a moderate scalability level, and an additional 10 POMS items were removed. Following the MSA analysis, Rasch analyses were conducted on the reduced 31-item POMS, again using an iterative process within each of the 6 domains, to remove all misfitting items. Much like in phase 1, all items removed through the quantitative analysis had previously been identified as poorly performing items and removed through other revisions of the POMS, such as ready to fight,^{7,12,13,33} miserable,^{12,13,18} and helpless.^{13,14,16} The MSA technique is a useful nonparametric approach for the development of questionnaires to measure health constructs,³⁴ which relies on less stringent statistical assumptions compared to Rasch analysis.³⁵ By coupling these 2 approaches in the development of the PKU-POMS, the current study was able to first identify the most problematic items within each domain and then more specifically assess the fit of the remaining items and domains to the stricter Rasch measurement model. The fit of the draft 20-item PKU-POMS to the Rasch measurement model and scalability was found acceptable.

In phase 3, the psychometric properties of the draft 20-item PKU-POMS were examined, and cognitive interviews among adults with PKU were conducted to investigate the acceptability of the draft instrument in the target population. These analyses demonstrated that the revised domains were internally consistent and provided preliminary support for the

Table 7. PKU-POMS Convergent Validity and Responsiveness Assessed Using Pearson's Correlations Between PKU-POMS and ADHD RS-IV and Plasma Phe Levels.

PKU-POMS Domain	Convergent Validity ^a				Responsiveness ^b			
	ADHD RS-IV		Plasma Phe		ADHD RS-IV		Plasma Phe	
	N	Pearson's <i>r</i>	N	Pearson's <i>r</i>	N	Pearson's <i>r</i>	N	Pearson's <i>r</i>
Activity	114	−0.04	114	−0.06	65	−0.40 ^c	52	−0.25
Anxiety	114	0.33 ^d	114	0.09	65	0.43 ^d	52	0.34 ^c
Confusion	114	0.60 ^e	114	0.09	65	0.52 ^e	52	0.47 ^d
Tiredness	114	0.28 ^c	114	0.17	65	0.34 ^c	52	0.15
Anger	114	0.31 ^d	114	0.08	65	0.32 ^c	52	0.26
Depression	114	0.15	114	0.00	65	0.21	52	0.32 ^c

Abbreviations: ADHD RS-IV, Attention-Deficit Hyperactivity Disorder Rating Scale IV; EOS, end of study; Phe, phenylalanine; PKU, phenylketonuria; POMS, profile of mood states.

^aConvergent validity assessed using values at day 1.

^bResponsiveness assessed using change scores from day 1 to EOS.

^c*p* < .05.

^d*p* < .001.

^e*p* < .0001.

convergent validity and responsiveness of the measure. Finally, cognitive interviews on the draft PKU-POMS indicated that the measure is clear and easy to understand, with all items being relevant to patients with PKU. Additionally, patients believed that the revised domain labels accurately reflected the items contained within each domain. Thus, through this adaptation process, the PKU-POMS is shorter and less burdensome for patients to complete compared to the original 65-item POMS, and is clear, relevant, and easy for patients with PKU to understand.

Various research studies in the patients with PKU have demonstrated that even with early initiation of treatment, there is still an increased prevalence for various psychiatric, neuro-cognitive, and behavioral problems, including problems with mood,²⁻⁵ and that these psychiatric problems are associated with increases in Phe. Using the Brief Symptom Inventory (BSI), Bilder and colleagues³⁶ assessed the psychiatric symptom patterns in 64 patients with PKU, finding that 6 of the 7 subscales of the BSI were elevated in patients with PKU. Further, in a systematic review of 10 published intervention and case reports, there was a clear association between reductions in Phe with marked reductions in psychiatric symptoms in all studies.³⁷ However, no instruments of psychiatric symptoms used in any of the clinical studies identified were developed specifically for use in adults with PKU. The revised PKU-POMS developed through 3 phases of qualitative and quantitative research with adults having PKU can help more clearly elucidate the relationships between Phe and mood symptoms and serve as a valid, reliable, and responsive measure of changes in mood symptoms in clinical studies of PKU.

The strengths of this study included the use of both qualitative and quantitative methods to adapt the POMS for use in patients with PKU. However, certain limitations of the study should be noted. Specifically, the participants recruited for phase 1 included only patients taking part in the National PKU Alliance Conference or saw the recruitment flier online via a

study Web site. Thus, it is unknown if these findings generalize to a broader sample of adults with PKU. Additionally, the sample size in the PRISM-1 study at the EOS was limited (*n* = 65), and the psychometric validation analyses utilized the same data with which the quantitative data reduction analyses were conducted. Thus, a follow-up psychometric validation study on the PKU-POMS in a separate sample of patients with PKU will be used to further inform the usefulness and psychometric properties of this instrument.

Conclusion

A detailed, 3-phase item reduction process incorporating qualitative and quantitative techniques yielded the 20-item PKU-POMS instrument. This new instrument is designed for the assessment of treatment efficacy on change in mood states in clinical trials of adult patients with PKU.

Acknowledgments

The authors thank the National PKU Alliance for their support, and the participants who contributed to this research.

Declaration of Conflicting Interests

The author(s) declared potential conflict of interest with respect to the following research, authorship, and/or publication of this article: Yinpu Chen, Rishabh Jain, Laurel Konkol, Markus Merilainen, and Holly Weng are employees of BioMarin Pharmaceuticals and receive financial support in terms of salary and stock.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Elizabeth Bacci, Kathleen Wyrwich, and Katharine Gries are employees of Evidera, which received financial support from BioMarin Pharmaceuticals in connection with the study development and execution as well as the manuscript development.

References

1. Kaufman S. An evaluation of the possible neurotoxicity of metabolites of phenylalanine. *J Pediatr*. 1989;114(5):895-900.
2. Moyle JJ, Fox AM, Arthur M, Bynevelt M, Burnett JR. Meta-analysis of neuropsychological symptoms of adolescents and adults with PKU. *Neuropsychol Rev*. 2007;17(2):91-101.
3. Pietz J, Fatkenheuer B, Burgard P, Armbruster M, Esser G, Schmidt H. Psychiatric disorders in adult patients with early-treated phenylketonuria. *Pediatrics*. 1997;99(3):345-350.
4. Smith I, Knowles J. Behaviour in early treated phenylketonuria: a systematic review. *Eur J Pediatr*. 2000;159(suppl 2):S89-S93.
5. Waisbren S. Phenylketonuria. In: Goldstein S, Reynolds CR, eds. *Handbook of Neurodevelopmental and Genetic Disorders in Children*. New York, NY: Guilford Press; 1999:433-458.
6. ten Hoedt AE, de Sonnevile LM, Francois B, et al. High phenylalanine levels directly affect mood and sustained attention in adults with phenylketonuria: a randomised, double-blind, placebo-controlled, crossover trial. *J Inherit Metab Dis*. 2011;34(1):165-171.
7. Wald FD, Mellenbergh GJ. [The shortened version of the Dutch translation of the Profile of Mood States (POMS)]. *Ned Tijdschr voor Psychol*. 1990;45(2):86-90.
8. McNair DM, Lorr M, Droppleman LF. *Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Services; 1971.
9. McNair DM, Heuchert J, Shilony E. *Profile of Mood States (POMS): Technical Update*. Tonawanda, NY: Multi-Health Systems, Inc; 2003.
10. Heuchert JP, McNair DM. *POMS 2 ®: Profile of Mood States Second Edition*. Tonawanda, NY: Multi-Health Systems, Inc; 2012.
11. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A procedure for evaluating sensitivity to within-person change: can mood measures in diary studies detect change reliably? *Pers Soc Psychol Bull*. 2006;32(7):917-929.
12. Shacham S. A shortened version of the Profile of Mood States. *J Pers Assess*. 1983;47(3):305-306.
13. Sutherland HJ, Lockwood GA, Cunningham AJ. A simple, rapid method for assessing psychological distress in cancer patients. *J Psychosoc Oncol*. 1989;7(1-2):31-43.
14. Terry PC, Lane AM, Lane HJ, Keohane L. Development and validation of a mood measure for adolescents. *J Sports Sci*. 1999;17(11):861-872.
15. Fernandez EMA, Fernandez CA, Pesqueria GS. Adaptation of the profile of mood states into Spanish with a sample athletes. *Psicothema*. 2002;14(4):708-713.
16. Netz Y, Zeav A, Arnon M, Daniel S. Translating a single-word items scale with multiple subcomponents—a Hebrew translation of the Profile of Mood States. *Isr J Psychiatry Relat Sci*. 2005;42(4):263-270.
17. Perczek R, Carver CS, Price AA, Pozo-Kaderman C. Coping, mood, and aspects of personality in Spanish translation and evidence of convergence with English versions. *J Pers Assess*. 2000;74(1):63-87.
18. Shin YH. A study on verification of the profile of mood states (POMS) for Korean elders. *J Korean Acad Nurs*. 1996;26(4):743-758.
19. Albani C, Blaser G, Geyer M, et al. [The German short version of “Profile of Mood States” (POMS): psychometric evaluation in a representative sample]. *Psychother Psychosom Med Psychol*. 2005;55(7):324-330.
20. Curran SL, Andrykowski MA, Studts JL. Short Form of the Profile of Mood States (POMS-SF): Psychometric Information. *Psychol Assess*. 1995;7(1):80-83.
21. Fillion L, Gagnon P. French adaptation of the shortened version of the Profile of Mood States. *Psychol Rep*. 1999;84(1):188-190.
22. Moret L, Nguyen JM, Pillet N, Falissard B, Lombrail P, Gasquet I. Improvement of psychometric properties of a scale measuring inpatient satisfaction with care: a better response rate and a reduction of the ceiling effect. *BMC Health Serv Res*. 2007;7:197.
23. Steiger JH. Understanding the limitations of global fit assessment in structural equation modeling. *Pers Individ Differ*. 2007;42(5):893-898.
24. Yu CY. *Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes* [doctoral dissertation]. Los Angeles, CA: University of California; 2002.
25. van Shurr WH. Mokken scale analysis: between the Guttman scale and parametric item response theory. *Polit Anal*. 2003;11(2):139-143.
26. Loevinger J. The technic of homogeneous tests compared with some aspects of scale analysis and factor analysis. *Psychol Bull*. 1948;45(6):507-529.
27. Mokken RJ. *A Theory and Procedure of Scale Analysis*. The Hague/Berlin: Mouton/De Gruyter; 1971.
28. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. (Republished). Chicago, IL: The University of Chicago Press; 1980.
29. Embretson SE, Reise SP. *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
30. Linacre JM. What do infit and outfit mean-square and standardized mean? *Rasch Meas Trans*. 2002;16(2):878.
31. Bollen KA. *Structural Equations with Latent Variables*. Hoboken, NJ: John Wiley and Sons; 1989.
32. Andrade E, Arce C, Torrado J, Garrido J, De Francisco C, Arce I. Factor structure and invariance of the POMS Mood State Questionnaire in Spanish. *Span J Psychol*. 2010;13(1):444-452.
33. Grove JR, Prapavessis H. Preliminary evidence for the reliability and validity of an abbreviated Profile of Mood States. *Int J Sport Psychol*. 1992;23(2):93-109.
34. Stochl J, Jones PB, Croudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Med Res Methodol*. 2012;12:74.
35. Sijtsma K, Molenaar IW. *Introduction to Nonparametric Item Response Theory*. London, UK: Sage Publications; 2002.
36. Bilder DA, Burton BK, Coon H, et al. Psychiatric symptoms in adults with phenylketonuria. *Mol Genet Metab*. 2013;108(3):155-160.
37. Bilder DA, Noel JK, Baker ER, et al. A systematic review (SR) of the effectiveness of reducing blood phenylalanine (Phe) levels in adults with phenylketonuria (PKU) on neuropsychiatric symptoms. *J Inherit Metab Dis*. 2014;37(suppl 1):S27-S185; Abstract P-069.