Examination of the PROMIS Upper Extremity Item Bank

Man Hung, PhD[1,2,3], **Maren W. Voss, MS**[1], **Jerry Bounsanga**[1], **Anthony B. Crum, BS**[1], and **Andrew R. Tyser, MD**[1]

[1]University of Utah Department of Orthopaedics, Salt Lake City, Utah 84108

[2]University of Utah Division of Epidemiology, Salt Lake City, Utah 84108

[3]University of Utah Division of Public Health, Salt Lake City, Utah 84108

Abstract

Study Design—Clinical measurement

Purpose—The Patient-Reported Outcomes Measurement Information System (PROMIS) v1.2 Upper Extremity (UE) item bank was recently developed in response to research demonstrating UE function to be a separate domain from general physical function as measured with PROMIS instruments. In particular, the PROMIS v1.2 UE item bank assesses upper extremity physical function in patients with hand and upper extremity conditions. Though validation on a general sample was conducted prior to its release, the psychometric properties of the PROMIS v1.2 UE item bank have not been fully evaluated in specific populations. This study assesses the performance of the PROMIS v1.2 UE within the UE orthopaedic patient population.

Methods—A total of 1,197 adult patients presenting to a tertiary orthopaedic clinic specializing in hand and upper extremity conditions were administered the PROMIS v1.2 UE items and three other PROMIS scales. Descriptive statistics of patient demographics and item raw scores were performed, as well as Rasch analysis of the PROMIS v1.2 UE item bank.

Results—The PROMIS v1.2 UE item bank fits a unidimensional Rasch model. It demonstrated adequate reliabilities, item local independence and model fit. There was essentially no issue with gender differential item functioning. The test items mostly targeted low levels of function, thus revealing large ceiling effect.

Conclusions—The PROMIS v1.2 UE item bank is a useful clinical assessment tool with good psychometric properties that can be implemented in clinical settings. Consideration should be

Corresponding Author: Man Hung, PhD, Associate Professor, [1]University of Utah School of Medicine, 590 Wakara Way, Salt Lake City, UT. 84108 USA, Phone: 801-587-5372, Fax: 801-587-5411, Man.Hung@hsc.utah.edu.
Jerry Bounsanga, [1]University of Utah, 590 Wakara Way, Salt Lake City, UT 84108 USA, Jerry.Bounsanga@utah.edu
Maren W. Voss, M.S., [1]University of Utah, 590 Wakara Way, Salt Lake City, UT 84108 USA, Maren.Voss@utah.ed
Anthony B. Crum, [1]University of Utah, 590 Wakara Way, Salt Lake City, UT 84108 USA, Bryan.Crum@utah.edu
Andrew R. Tyser, MD, [1]590 Wakara Way, Salt Lake City, UT 84108 USA, Andrew.Tyser@hsc.utah.edu

given to supplemental testing for patients at higher levels of function with other legacy tests until the ceiling effects of the PROMIS v1.2 UE item bank are addressed.

Keywords

Orthopaedics; Patient-reported outcomes; Physical function; PROMIS upper extremity; Rasch

1.0 Introduction

The Patient-Reported Outcomes Measurement Information System (PROMIS), sponsored by the National Institutes of Health, has developed numerous patient reported outcome (PRO) instruments in the last decade. The PROMIS Physical Function (PF) instrument was one instrument developed as a part of this project [1], and has been increasingly utilized in clinical and research settings. The PROMIS PF instrument assesses five categories of physical functioning which include upper extremity (UE), lower extremity (LE), axial, central, and instrumental activities of daily living (IADL). Early analysis indicated that the PROMIS PF items provided a reasonable fit with unidimensional item-response theory (IRT), suggesting that the items are able to provide information about distinct and measurable domains in physical function. One weakness, however, is that the PROMIS PF was not able to discriminate well at higher levels of functioning.[2] The subjective nature of PRO instruments have raised valid concerns that they may fail to detect extremes in functioning[3], a problem that leads to floor and/or ceiling effects. The floor is when the measure is not sensitive enough to provide information about the lower levels of functioning and the ceiling is when high levels functioning are not well captured by the measure.

Despite rigorous development, validation of measures within specific populations can reveal limitations that need to be addressed. Unidimensionality is the concept that a test is measuring one thing and that the test items provide information about that single construct. The unidimensionality of the original PROMIS PF has been questioned and an additional category has since been added. Rose and colleagues found that items in the PF domain assessing upper body function did not fit well within the assumptions of the unidimensional IRT model, that they were measuring something different.[4] Hung, et al. found non-trivial variance in lower and upper extremity measures [5] indicating a necessary distinction between the upper extremity (UE) and lower extremity (LE) domains. In addition, a pronounced ceiling effect for UE PF items was noted in their large sample of patients with UE disorders (n=865). This distinction between upper and lower extremity function was also found in analysis a pediatric population. Researchers found that UE and LE (or mobility) subdomains created better item fit over a single PROMIS PF dimension for a pediatric population.[6] When these independent and different subdomains are aggregated into a single score for physical function, the responsiveness of the measure is minimized, it has decreased precision, and these limitations might impair the ability to make clinical decisions from its output.[6]

This work has led ultimately to the development of two subscales of the original PROMIS PF item bank: one for UE and one for LE (or mobility) function.[7] The PROMIS v1.2 UE item bank was developed by identifying 16 items with adequate item fit within a UE domain

and with items showing unique variance from the overall PROMIS PF domain.[7] Separating out items that distinctly measure UE function allows the scale to be unidimensional and it can then be validated within a unidimensional IRT model. That development project analyzed retrospective data from 21,773 participants who participated in the PROMIS wave 1 data collection. That analysis from a general population indicated a problematic ceiling effect and the need for further development of UE items for higher levels of function. The developers found that even those with low levels of physical function had a 50% probability of selecting the most extreme high function scores.[7] This finding of a problematic ceiling effect and a lack of sensitivity to detecting higher function may or may not be significant when the measure is tested on clinical populations which are more likely to experience impairment. Therefore, testing the measure in specific clinical populations is an important next step in validating its usefulness.

Research on the recently developed PROMIS v1.2 UE item bank is limited. Past research has investigated the usefulness of the general PROMIS PF in clinical populations when compared to commonly used legacy instruments which evaluate UE function, such as the DASH and *Quick*DASH.[8, 9] However, this research on clinical populations has not been conducted for the newly developed UE item bank. The PROMIS v1.2 UE measure of physical function has not been extensively validated or tested in clinical populations. The present purpose is to evaluate the validity in an orthopaedic patient population with UE disorders. The present study described the psychometric properties of the PROMIS v1.2 UE item bank in a UE orthopaedic population.

2.0 Methods

Data

Data were collected from 1,197 patients who were aged 18 years or older and presented to the hand and upper extremity clinic at a tertiary university orthopaedic center for upper extremity (non-shoulder) complaint. All new, return, and post-op patients were administered four PROMIS computer adaptive tests (CAT)s via tablet computers during 2014–2015 while waiting to be seen by medical staff, as part of routine clinical care. These measures included the PROMIS v1.2 UE CAT, the PROMIS v1.2 Physical Function CAT, the PROMIS v.10 Anxiety CAT and the PROMIS v1.1 Pain Interference CAT.. Anxiety and pain scales were included in data collection due to the effect of anxiety [10] and pain [11] on physical function. All of the data were securely collected and stored at the institution's enterprise data warehouse. Institutional Review Board approval was obtained for the current study.

Analysis

Descriptive statistics of patient demographics and item raw scores were performed, then followed by Rasch analysis with WinSteps 3.81 [12] software to examine the psychometric properties of the PROMIS v1.2 UE item bank. The Rasch Partial Credit Model was selected because the instrument had more than two response categories in which the locations were not common to all items.[13] We started with unidimensionality testing by conducting a principal component analysis (PCA) of the residuals. Unidimensionality is a fundamental requirement of measurement. It refers to the notion that there is only one attribute existing in

the instrument and helps to make clear interpretation of findings free of confounding attributes. Unidimensionality may be established if the PCA reveals an unexplained variance by the first contrast less than 5% [14, 15], or an eigenvalue less than 2.0.[16]

Local item independence and item fit were conducted to examine how the items were fit to the model. Items that do not fit the model add random noise to the instrument and should be isolated and dealt with appropriately. Item residual correlations of greater than 0.7 indicate item dependence and where items are redundant they could be eliminated. Infit and outfit mean square (MNSQ) values were used to measure item fit, with MNSQ values less than 2.0 considered fitting to the model.[15, 17, 18]

Also known as item bias, differential item functioning (DIF) is an indicator of whether items function the same across subgroups. DIF was conducted in this study to see if items exhibit difference in difficulty measures across gender, which was categorized as either being female or male. A DIF contrast of greater than 1.0 logits was considered as significantly different.[19]

Person and item reliability were also analyzed, which refers to the consistent reproducibility of a person's score and the ordering of item difficulty under similar conditions. Expressed as an r-value between 0 and 1, person and item reliabilities greater than or equal to 0.80 are considered good whereas values greater than 0.90 are considered excellent.[20–23] Similar to reliability, both the person separation index (PSI) and item separation index (ISI) were also conducted. An advantage of PSI and ISI is that their values do not have an upper bound such as in the case of an r-value. The ISI is an indicator of construct validity, that is, the hierarchy of the item difficulty measures. An ISI of at least 2.0 is needed to indicate sufficient sample size for confirming the item hierarchy.[14, 15] A PSI of greater than 2.0 is needed to reach a reliability of 0.80.[14, 15]

Instrument coverage or targeting was analyzed to identify which items were able to cover or adequately target a range of abilities, traits, or functions of a patient. The person-item map can reveal whether the items in an instrument are targeted to measure the patients' ranges of function or ability. Displayed vertically, person and item distribution are separated from left to right respectively on the map (Fig. 1). Higher functioning levels are at the top of the vertical graph whereas lower functioning levels are at the bottom. All items in the instrument should target or cover patients at varying levels of functioning in order to have adequate coverage. It is suggested that if the difference between the average of the item measures and the average of the person measures is greater than 1, the instrument is said to be off targeting.[24]

3.0 Results

The current study included a total of 1,197 patients visiting a tertiary orthopaedic clinic for upper extremity conditions, with demographic information summarized in Table 1.

All of the PROMIS measures were recorded in T-scores. The mean PROMIS UE score was 35.58, (SD=10.00; min=14.30; max=56.26). The PROMIS UE measure was highly

correlated with three other PROMIS measures – PROMIS Anxiety (r=−0.48; p <0.0001), PROMIS Pain Interference (r=−0.65; p <0.0001), and PROMIS PF (r=0.71; p<0.0001).

3.1 Unidimensionality

An assessment of dimensionality returned an eigenvalue of 1.96 and 4.2% of the unexplained variance attributed to the first contrast in the Rasch dimension's residuals, providing evidence of unidimensionality. The total raw variance explained by the measures was 65.9%, of which 53.7% was by the persons and by the items was 12.2%.

3.2 Local Independence

Inter-item residual correlations were computed for pairs of items from the PROMIS v1.2 UE item bank. (Table 3). These correlations ranged from −0.37 to 0.34. They are all below 0.70, implying sufficient local independence of items.

3.3 Item Fit

The outfit MNSQ for the items ranged from 0.64–1.70, showing adequate fit to the Rasch model. Similarly, the infit MNSQ also showed adequate fit, ranging from 0.61–1.61. Overall, the PROMIS v1.2 UE items fit the Rasch model well.

3.4 Differential Item Functioning

Male and female subjects showed few systematic differences in the item measures. DIF contrasts in gender ranged from −1.03 to 0.81. All items in the PROMIS v1.2 UE bank functioned similarly across gender, with the exception of PFB30 which asked "Are you able to open a new milk carton?" which had a DIF contrast of −1.03.

3.5 Reliability

The instrument also showed adequate reliabilities. Person reliability was 0.84 while item reliability was 0.82. Additionally, both PSI and ISI showed adequate separation, with respective values of 2.33 and 2.38.

3.6 Targeting

Instrument targeting is assessed at the item and person level. The item mean measure was 0.00 and the person mean measure was 1.32. To look at whether the difficulty of measures match the items with the persons, the average item measure is subtracted from the average person measure, resulting in a difference of 1.32. This value is greater than 1.00, which suggests the instrument is not well targeted to the patient population. Looking at the person-item map (Fig. 1) it is evident that more of the PROMIS v1.2 UE items mapped to lower functioning ability levels of the patients with little coverage at the upper ability range. This evidences a ceiling effect.

4.0 Discussion

The ability to accurately assess functional outcomes in the treatment of upper extremity (UE) musculoskeletal conditions is an important element to diagnosis and in monitoring

intervention effectiveness. Rational treatment decisions are empowered by tools that predictably and reliably assess patient responses to treatment, and as a result, patient-reported outcome (PRO) instruments have been increasingly recognized and utilized in both clinical and research settings.[25] It is therefore of significant importance that these PROs exhibit high quality measurement characteristics. They should fit within a conceptual model, have sufficient reliability, validity, interpretability of scores, and minimize responder burden.[26]

In this study the PROMIS v1.2 UE CAT was administered to patients experiencing a wide range of upper extremity issues including longer term disability and/or disease as well as shorter term acute conditions. The psychometric properties of its item bank were evaluated in this specific orthopaedic UE patient population to contribute a greater understanding of the measurement properties of this scale. The instrument showed high correlations with the PROMIS Anxiety CAT, Pain Interference CAT, and PF CAT, as can be expected for self-reported measures of physical function which are influenced by patient perceptions and psychological states.[27, 28]

We confirmed that the PROMIS v1.2 UE items were a good fit with the Rasch model, confirming that the items met the requirement of unidimensionality. This means that upper extremity function demonstrated to be a single domain with independent characteristics compared to overall physical function. Further, the instrument demonstrated adequate reliability, with both person and item reliabilities over 0.80. It had adequate PSI and ISI, demonstrating that it was able to distinguish between different participants, a measure of construct validity. DIF between males and females was not a problem for 15 of the 16 items, suggesting equal applicability across genders. These findings correspond with the validation studies of the PROMIS v1.2 UE which demonstrated strong psychometric properties.[7]

In every aspect of psychometric examination, the PROMIS v1.2 UE instrument performed well, with one exception. The instrument was more targeted to lower levels of function, with a limited number of questions targeted toward those with higher function. This finding corresponds to the analysis from a general population indicating ceiling effect in the UE item bank, however this finding is of greater concern in this analysis.[7] While perhaps this limited ability to discriminate high-end functioning of the upper extremity might not be overly concerning in a general population of healthy adults, our findings that this effect held true even for individuals with upper body disability or presumed impairment is particularly worrisome. The inability of the PROMIS v1.2 UE item bank to measure higher functioning in an impaired population may limit its usefulness as a measurement tool. Adding additional questions to the item banks—specifically questions that target higher functioning ability levels—would likely improve the utility of the PROMIS UE CAT in hand and upper extremity orthopaedics.

This problem with floor and/or ceiling effects for physical function measurement is not unique to the PROMIS v1.2 UE item bank and demonstrates a general limitation in measurement of UE function. Other commonly used measures have evidenced significant concerns with targeting toward low levels of function. The Disability of the Arm, Shoulder, and Hand (DASH) had a ceiling effect in orthopaedic populations [8] and the DASH and Short

Musculoskeletal Functional Assessment (SMFA) evidenced ceiling effects in a mature population (aged 60 or older).[29] While each measurement tool has its limitations, the PROMIS v1.2 UE has compared well with legacy instruments already in standard practice.[30, 31] The current study builds on this prior research in providing a psychometric analysis of the PROMIS v1.2 UE item bank in an UE orthopaedic patient population with suggestions for improving its targeting to higher levels of function.

Further studies are needed to evaluate the performance of the PROMIS v1.2 UE item bank in comparison to other legacy instruments that assess upper body function. The advantages of the PROMIS UE CAT in reducing testing time and burden can make it a highly efficient and useful tool in clinical practice. In addition, the electronic administration of the instrument creates an opportunity to link with medical records and improve data fidelity. While electronic administration requires an investment in equipment and training, there are advantages to integrating technology into clinical practice in the long run.

The study has limitations. This analysis of the PROMIS v1.2 UE item bank was a cross-sectional examination of patient responses and did not incorporate repeated measures or assess the ability of the instrument to detect changes in function. Longitudinal analysis would be useful, particularly in an UE orthopaedic patient population undergoing treatments that can produce dramatic improvement. Detecting change in function is important in the clinical application of PRO measures and deserves further investigation.

5.0 Conclusion

The PROMIS v1.2 UE item bank exhibits good psychometric properties, with the exception that it is not well targeted to the entire range of function in our patient population. Most of the items are targeted to the lower functioning patients, which is an indication of ceiling effect. Additional items are need to cover the upper range of function.

Acknowledgments

References

1. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med. Care. 2007; 45:S3.

2. Hays RD, Liu H, Spritzer K, Cella D. Item response theory analyses of physical functioning items in the medical outcomes study. Med. Care. 2007; 45:S32–S38. [PubMed: 17443117]

3. Fries JF, Krishnan E. What constitutes progress in assessing patient outcomes? J. Clin. Epidemiol. 2009; 62:779–780. [PubMed: 19282147]

4. Rose M, Bjorner JB, Becker J, Fries J, Ware J. 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; 61:17–33. [PubMed: 18083459]

5. Hung M, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. J. Orthop. Res. 2011; 29:947–953. [PubMed: 21437962]

6. DeWitt EM, Stucky BD, Thissen D, et al. Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: built using item response theory. J. Clin. Epidemiol. 2011; 64:794–804. [PubMed: 21292444]

7. Hays RD, Spritzer KL, Amtmann D, et al. Upper-extremity and mobility subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. Arch. Phys. Med. Rehabil. 2013; 94:2291–2296. [PubMed: 23751290]

8. Tyser AR, Beckmann J, Franklin JD, et al. Evaluation of the PROMIS Physical Function Computer Adaptive Test in the Upper Extremity. The Journal of hand surgery. 2014; 39:2047–2051. e2044. [PubMed: 25135249]

9. Overbeek CL, Nota SP, Jayakumar P, Hageman MG, Ring D. The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. Clinical Orthopaedics and Related Research®. 2015; 473:311–317. [PubMed: 25099262]

10. Scopaz KA, Piva SR, Wisniewski S, Fitzgerald GK. Relationships of Fear, Anxiety, and Depression With Physical Function in Patients With Knee Osteoarthritis. Arch. Phys. Med. Rehabil. 2009; 90:1866–1873. [PubMed: 19887210]

11. Liechtenstein MJ, Dhanda R, Cornell JE, Escalante A, Hazuda HP. Disaggregating pain and its effect on physical functional limitations. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 1998; 53:M361–M371.

12. Linacre, JM. Winsteps Rasch measurement computer program. Beaverton, Oregon: 2016.

13. Rasch G. Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. 1960

14. Hung M, Hon SD, Cheng C, et al. Psychometric Evaluation of the Lower Extremity Computerized Adaptive Test, the Modified Harris Hip Score, and the Hip Outcome Score. Orthopaedic Journal of Sports Medicine. 2014; 2 2325967114562191.

15. Wright BD, Masters GN. Rating Scale Analysis. Rasch Measurement: ERIC. 1982

16. Gothwal VK, Bagga DK. Vision and Quality of Life Index: validation of the Indian version using Rasch analysis. Investigative ophthalmology & visual science. 2013; 54:4871–4881. [PubMed: 23800765]

17. Bond, TG., Fox, CM. Applying the Rasch model: Fundamental measurement in the human sciences. Psychology Press; 2013.

18. Hung M, Carter M, Hayden C, et al. Psychometric assessment of the patient activation measure short form (PAM-13) in rural settings. Qual. Life Res. 2013; 22:521–529. [PubMed: 22466721]

19. Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. Medical care. 2004; 42:I–37.

20. Davidshofer, K., Murphy, C. Psychological testing: principles and applications. Upper Saddle River, NJ: Pearson/Prentice Hall; 2005.

21. Hung M, Baumhauer JF, Brodsky JW, et al. Psychometric Comparison of the PROMIS Physical Function CAT With the FAAM and FFI for Measuring Patient-Reported Outcomes. Foot Ankle Int. 2014; 35:592–599. [PubMed: 24677217]

22. Hung M, Baumhauer JF, Latt LD, et al. Validation of PROMIS® Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. Clinical Orthopaedics and Related Research®. 2013; 471:3466–3474. [PubMed: 23749433]

23. Hung M, Hon SD, Franklin JD, et al. Psychometric properties of the PROMIS physical function item bank in patients with spinal disorders. Spine (Phila Pa 1976). 2014; 39:158–163. [PubMed: 24173018]

24. Khadka J, Pesudovs K, McAlinden C, Vogel M, Kernt M, Hirneiss C. Reengineering the glaucoma quality of life-15 questionnaire with rasch analysis. Invest Ophthalmol Vis Sci. 2011; 52:6971–6977. [PubMed: 21810983]

25. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J. Eval. Clin. Pract. 2006; 12:559–568. [PubMed: 16987118]

26. Deutsch L, Smith L, Gage B, Kelleher C, Garfinkel D. Patient-reported outcomes in performance measurement: commissioned paper on PRO-based performance measures for healthcare accountable entities. Washington, DC: National Quality Forum. 2012

27. Stegenga BT, Nazareth I, Torres-Gonzalez F, et al. Depression, anxiety and physical function: exploring the strength of causality. J. Epidemiol. Community Health. 2012; 66:e25. [PubMed: 21693471]

28. Brown DJ, McMillan DC, Milroy R. The correlation between fatigue, physical function, the systemic inflammatory response, and psychological distress in patients with advanced lung cancer. Cancer. 2005; 103:377–382. [PubMed: 15558809]

29. Morgan JH, Kallen MA, Okike K, Lee OC, Vrahas MS. PROMIS Physical Function Computer Adaptive Test Compared With Other Upper Extremity Outcome Measures in the Evaluation of Proximal Humerus Fractures in Patients Older Than 60 Years. J. Orthop. Trauma. 2015; 29:257–263. [PubMed: 26001348]

30. Waljee JF, Carlozzi N, Franzblau LE, Zhong L, Chung KC. Applying the Patient-Reported Outcomes Measurement Information System to Assess Upper Extremity Function among Children with Congenital Hand Differences. Plast. Reconstr. Surg. 2015; 136:200e–207e.

31. Döring A-C, Nota SP, Hageman MG, Ring DC. Measurement of upper extremity disability using the patient-reported outcomes measurement information system. The Journal of hand surgery. 2014; 39:1160–1165. [PubMed: 24799143]

HIGHLIGHTS

This study attempts at exploring the PROMIS v1.2 UE item bank in the orthopaedic population.

- This is the first study evaluating the psychometric properties of PROMIS v1.2 UE item bank in patients with hand concerns.

- The PROMIS v1.2 UE item bank exhibits good psychometric properties.

- More items need to be developed to target higher functioning orthopaedic patients.

```
        MEASURE    PERSON        ITEM

        5 .###########  +
                        |
                        |
                        |
                .#####  T|
                .     |
        4       .     +
                      |
                .######  |
                .     |
                .     |
        3       .###  +
                .##   |
                .##  S|
                .###  |
                .##   |
                .###  |
        2       .##   +
                ##    |
                .####  |
                .###   |T
                .#####  |   PFA28
                .#####  |   PFA18
        1         #   M+
        .########  |   PFA48
                .#  |S PFA29R1
                ######  |   PFB21     PFB30
                .#  |
                ######  |
        0       .#  +M PFA20
        .######  |   PFA16R1   PFA38      PFB22
                #  |
        .######  |   PFA17
                .  |S PFA35     PFA54      PFB33
                ### S|   PFB36
        -1      .  +   PFA44
                .###  |
                .   |
                .###  |T
                      |
                .###  |
        -2      .   +
                .   |
                .#  |
                .  T|
                .   |
                .   |
        -3          +
                .   |
                .   |
                      |
                .   |
                      |
        -4      .#  +

        EACH "#" IS 9 PEOPLE: EACH "." IS 1 TO 8 PEOPLE.
```
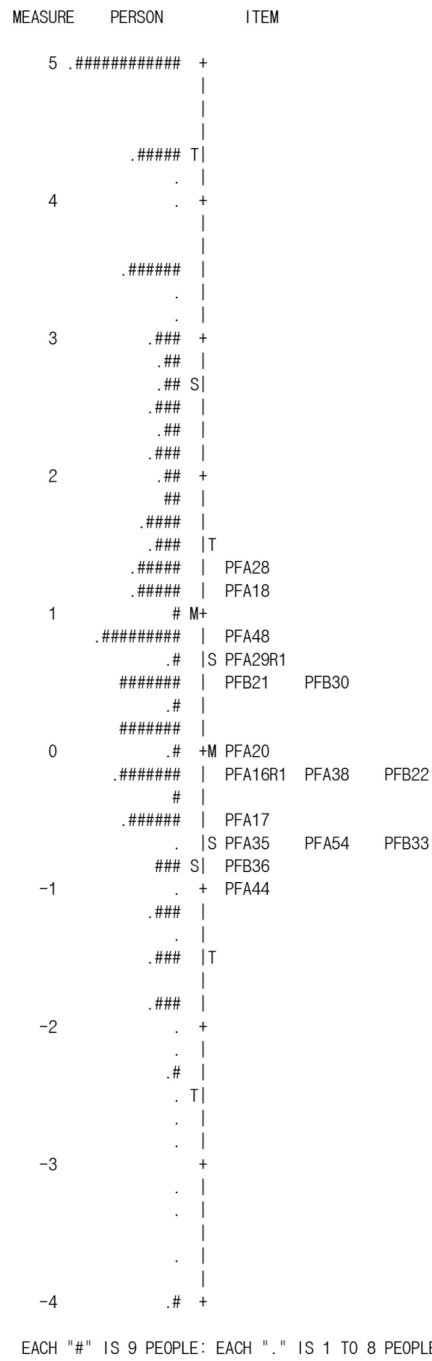
Fig 1.
Person-item map

Table 1

Patient Characteristics (N = 1,197).

Variable	Mean (SD)	n (%)	Range
Age (years)	46.14 (16.53)		18 – 93
Sex			
Male		564 (47.1)	
Female		633 (52.9)	
Employment Status			
Not Employed		205 (17.1)	
Full-Time		596 (49.8)	
Disabled		65 (5.4)	
Part-Time		71 (5.9)	
Retired		150 (12.5)	
Self-employed		56 (4.7)	
Student		45 (3.8)	
Unknown		9 (0.8)	
Tobacco User			
Never		725 (60.6)	
Yes		212 (17.7)	
Quit		195 (16.3)	
Passive		8 (0.7)	
Missing		2 (0.2)	
Not Asked		55 (4.6)	

Table 2

Dimensionality analysis of the PROMIS v1.2 Upper Extremity item bank.

	Eigen	%
Total raw variance explained (total)	49.92	100.0%
By measures	30.92	65.9%
By persons	25.19	53.7%
By items	5.73	12.2%
Raw unexplained variance (total)	16.00	34.1%
1st contrast	1.96	4.2%
2nd contrast	1.69	3.6%
3rd contrast	1.32	2.8%
4th contrast	1.19	2.5%

Table 3

Inter-item residual correlations.

Items	PFA16r1	PFA17	PFA18	PFA20	PFA28	PFA29r1	PFA35	PFA38	PFA44	PFA48	PFA54	PFB21	PFB22	PFB30	PFB33	PFB36
PFA16r1	1.0															
PFA17	−.0901	1.0														
PFA18	−.1345	−.1624	1.0													
PFA20	.0237	.0093	−.0267	1.0												
PFA28	−.0833	−.1779	−.1336	−.1000	1.0											
PFA29r1	−.3738	.0507	−.1861	−.0790	−.1393	1.0										
PFA35	.0569	−.0569	.0015	.0864	−.0149	−.0789	1.0									
PFA38	.0137	−.1267	−.0816	−.0514	−.0814	−.3101	.0580	1.0								
PFA44	.0646	−.1249	−.0441	−.0540	−.0472	−.3488	.0779	.1499	1.0							
PFA48	.0063	−.0541	−.0393	.1402	−.0905	−.0188	.1600	−.0217	.0043	1.0						
PFA54	.0063	−.0472	.0392	.2919	−.0681	−.0425	.0766	−.0553	−.0348	.3435	1.0					
PFB21	.0697	−.1130	−.1471	.0448	−.1192	−.1163	.0147	−.0447	−.0075	.0768	.1539	1.0				
PFB22	−.2511	−.1318	−.0980	.0059	−.0549	−.1098	−.1109	−.1791	−.2369	−.0454	−.0220	−.0418	1.0			
PFB30	.0010	−.0369	−.0793	−.0516	−.0228	−.0425	.0473	−.0151	.0063	.1446	−.1361	−.0632	−.0182	1.0		
PFB33	−.1093	−.0228	−.0409	−.1139	.0695	.0387	−.1584	−.1135	−.1018	−.1230	.0015	−.0094	.0349	−.0537	1.0	
PFB36	.0303	−.0547	−.1094	−.0317	−.0988	−.1113	.0631	.1175	.1320	−.0488	−.0773	−.0282	−.0913	−.0503	−.0874	1.0