# Oral Presentations

**(Listed Alphabetically by Presenting Author)**

| Presenting Author | Abstract Number & Title | Session | Room | Page # |
|---|---|---|---|---|
| Heidi Anttila | (6) Integrating PROMIS Short Forms to a Patient-Driven Mobile ICanFunction | Concurrent Session 4: Data Collection Methods in Clinical Practice | Iceland | 3 |
| Benjamin Arnold | (19) Translation of the Concept of Fatigue in PROMIS Measures | Concurrent Session 1: Translation and Cross-Cultural Validation | Norway | 8 |
| Susan Bartlett | (73) PROMIS-29 Health Profile and Related | Concurrent Session 3: Applications in Musculoskeletal Conditions | Iceland | 30 |
| Susan Bartlett | (71) Differences in PROMIS Responses in People with RA Recruited Online vs an Academic Medical Center | Concurrent Session 2: Interpreting PROMIS Scores | Iceland | 29 |
| Clifton Bingham | (75) Treatments Should Target Physical Function, Mood, and Fatigue to Enhance Participation in People with Rheumatoid Arthritis | Concurrent Session 4: Data Collection Methods in Clinical Practice | Iceland | 31 |
| Ryan Calfee | (66) Floor Effect of Depression on Patients with Orthopedic Conditions | Concurrent Session 2: Interpreting PROMIS Scores | Iceland | 26 |
| Ryan Calfee | (68) Patient-Reported Musculoskeletal Function: Comparing PROMIS Upper-Extremity to PROMIS General Physical Function Scores | Concurrent Session 4: New Developments of the Physical Function Item Bank | Finland | 27 |
| Aaron M. Chamberlain | (44) Correlation of PROMIS Upper Extremity, Pain Interference and Depression scores with SST and ASES scores | Concurrent Session 2: Applications Across Clinical Populations | Finland | 19 |
| John Chaplin | (25) Progress in the Cultural Adaptation of Four PROMIS® Pediatric Item Banks for Use in Sweden. | Concurrent Session 1: Translation and Cross-Cultural Validation | Norway | 11 |
| Patricia Connell | (22) Functional Impact of Congenital Hand Differences: Results from the Congenital Upper Limb Differences (CoULD) Registry | Concurrent Session 2: Applications Across Clinical Populations | Finland | 9 |
| Agnes Dardas | (67) Variability of PROMIS Domains Across Orthopedic Conditions | Concurrent Session 1: Item Bank Development and Validation | Iceland | 27 |
| Gerard Flens | (2) Longitudinal Measurement Invariance of the Dutch Flemish PROMIS Item Banks for Anxiety and Depression | Concurrent Session 2: Advanced Validation Methods | Norway | 1 |
| Christopher B. Forrest | (24) Development and Psychometric Evaluation of the Pediatric Oral Health Item Banks | Concurrent Session 1: Item Bank Development and Validation | Iceland | 10 |
| Rob J. Fredericksen, | (40) Development and Content Validation of a PROMIS Measure for Medication Adherence | Concurrent Session 1: Item Bank Development and Validation | Iceland | 17 |
| Carlos García Forero | (47) Multidimensional Higher-Order Factor Structure and Diagnostic Accuracy of PROMIS Mental Health Domain | Concurrent Session 2: Advanced Validation Methods | Norway | 19 |

| Presenting Author | Abstract Number & Title | Session | Room | Page # |
|---|---|---|---|---|
| Heather L. Gelhorn | (55) PROMIS Physical Function: Application to a Complex Rare Disease Population - Tenosynovial Giant Cell Tumor (TGCT) | Concurrent Session 3: Extending the Reach and Relevance of Existing PROMIS Item Banks | Finland | 22 |
| Charlotte Ibsen | (11) Development of a Patient-Reported Outcome Instrument for Patients with Lumbar Radicular Pain - STUDY PROTOCOL | Concurrent Session 3: Extending the Reach and Relevance of Existing PROMIS Item Banks | Finland | 5 |
| Alexander Joeris | (10) AO Patient Outcomes Center – A Novel Software to Implement PROMIS in Orthopedic Surgeons' Clinical Routine | Concurrent Session 4: Data Collection Methods in Clinical Practice | Iceland | 5 |
| Aaron Kaat | (69) An Upper Extremity Physical Functioning Item Bank | Concurrent Session 4: New Developments of the Physical Function Item Bank | Finland | 28 |
| Michael Kallen | (49) Determining When Multidimensionality Impacts Total Scale and Subscale Scores: Application of Confirmatory Bifactor Modeling to Osteoarthritis of the Knee (OAK) Patient Responses to the KOOS | Concurrent Session 2: Advanced Validation Methods | Norway | 20 |
| Shanthini Kasturi | (62) Validity and Reliability of PROMIS in Systemic Lupus Erythematosus (SLE) | Concurrent Session 3: Applications in Musculoskeletal Conditions | Iceland | 24 |
| Fionna Klasen | (14) German Translation of Six Item Banks from the Patient-Reported Outcomes Measurement Information System (PROMIS) | Concurrent Session 1: Translation and Cross-Cultural Validation | Norway | 6 |
| Kaisa Kokko | (79) Descriptive, Psychometric, and Feasibility Summaries of PROMIS Physical Function Instruments for the TOIMIA Database | Concurrent Session 4: New Developments of the Physical Function Item Bank | Finland | 32 |
| Gregor Liegl | (34) Reducing Ceiling Effects in PROMIS Physical Function Measures by Using an Extended Response Scale | Concurrent Session 4: New Developments of the Physical Function Item Bank | Finland | 14 |
| Christopher M. McAndrew | (57) Insurance Status and Patient Outcomes in the Orthopaedic Outpatient Setting | Concurrent Session 1: PROMIS on the Front Line: Assessing Quality of Care | Finland | 23 |
| Christopher M. McAndrew | (58) PROMIS Physical Function, Pain Interference, and Depression: Assessing Outcomes in an Orthopaedic Trauma Setting | Concurrent Session 3: Applications in Musculoskeletal Conditions | Iceland | 23 |
| David Mitten | (37) Impact of PROMIS on Patient Satisfaction | Concurrent Session 1: PROMIS on the Front Line: Assessing Quality of Care | Finland | 15 |
| Ashley Neese | (82) The Clinical Spectrum of PROMIS Physical Function Scores Over Time in Patients with Operative Lumbar Pathology | Concurrent Session 2: Interpreting PROMIS Scores | Iceland | 34 |
| M. Owen Papuga | (38) The Use of PROMIS CAT Post-Surgical Scores in Assessing “Value” in ACL Reconstruction Surgery. | Concurrent Session 1: PROMIS on the Front Line: Assessing Quality of Care | Finland | 16 |

| Presenting Author | Abstract Number & Title | Session | Room | Page # |
|---|---|---|---|---|
| Christina W. Schnohr | (9) Danish Translation of the PROMIS Item Bank for Physical Functioning – Results from Cognitive Testing | Concurrent Session 1: Translation and Cross-Cultural Validation | Norway | 4 |
| Wouter Schuller | (16) Responsiveness of the DF-PROMIS- PB and the DF-PROMIS- PI Item Banks in Patients Presenting in Musculoskeletal Practice | Concurrent Session 3: Applications in Musculoskeletal Conditions | Iceland | 7 |
| Eisuke Segawa | (70) Comparing IRT Pattern Scoring and IRT Summed Score Conversion Tables for PROMIS Short Forms | Concurrent Session 2: Interpreting PROMIS Scores | Iceland | 28 |
| Ewout B. Smit | (5) Development of a Dutch-Flemish PROMIS Physical Functioning Short Form for Geriatric Rehabilitation | Concurrent Session 3: Extending the Reach and Relevance of Existing PROMIS Item Banks | Finland | 2 |
| Amanda Spraggs-Hughes | (59) Using PROsetta Stone to Translate PROMIS Depression Scores for Meaningful Use in Orthopaedic Trauma | Concurrent Session 1: PROMIS on the Front Line: Assessing Quality of Care | Finland | 23 |
| Kai Sun | (50) Chinese-American Rheumatology Patients Using Traditional Chinese Medicine Have Worse PROMIS Health Status Scores | Concurrent Session 2: Applications Across Clinical Populations | Finland | 21 |
| Gemma Vilagut | (48) Construct Validity and Differential Item Functioning of Spanish PROMIS Depression, Anxiety and Anger | Concurrent Session 2: Advanced Validation Methods | Norway | 20 |
| Lari Wenzel | (72) Validation of PROMIS Emotional Distress Short Forms in a Cervical Cancer Survivor Population | Concurrent Session 1: Item Bank Development and Validation | Iceland | 29 |
| Lena Wettergren | (15) The Ability To Capture Sexual Dysfunction In Young Cancer Patients Using Items of the SexFS | Concurrent Session 2: Applications Across Clinical Populations | Finland | 6 |
| Susan Yount | (36) Creating Condition-Specific PROMIS Assessments: An Overview of the Gap Analysis Method | Concurrent Session 3: Extending the Reach and Relevance of Existing PROMIS Item Banks | Finland | 15 |
| Randi Zegman | (31) Using a Multi-Modal, Longitudinal Collection Strategy to Collect PROMIS Scales Across a Healthcare Delivery System | Concurrent Session 4: Data Collection Methods in Clinical Practice | Iceland | 13 |

# Poster Session

**(Listed Alphabetically by Presenting Author)**

## 1. Reliability of PROMIS Physical Function CAT, Pain Interference CAT, and Depression CAT Scores in Patients with Hip Intra-Articular Pathology Before and After Arthroscopy

Allison W. McIntyre, MPH, CCRC[1]; Jeff Houck, PT, PhD[2]; Brian Giordano, MD[1]; Kaili Pecorella, MS[1]

**Presenting Author: Allison W. McIntyre, MPH;** allisonw_mcintyre@urmc.rochester.edu

**Institutional Affiliations:** [1]University of Rochester Department of Orthopaedics, Rochester, NY [2]George Fox University, Newberg, OR

**Objective:** Analyze PROMIS scores collected for clinical use and research from the same patients, on the same day.
**Methods:** 1062 PROMIS scores (Physical Function (PF), Pain Interference (PI), Depression (DP)) were collected from 65 patients. Scores were obtained clinically and for research at pre-op (n=62), first follow up (n=55), 5-7 weeks (n=44), and 3-4 months (n=16) post arthroscopy. Intra-class correlation coefficient (ICC) and standard error of the measurement (SEM) were calculated for each time point. Good reliability = ICC of 0.8-0.89, excellent reliability = ICC of >0.9. Minimal detectable change (MDC) estimated as 1.96 times the SEM. Average change from first follow up to 3-4 months was calculated to determine if changes were larger than the MDC.
**Results:** Mean age 36; 68% female. Clinical scores missing from 10% of sample. Good to excellent reliability for all but one time point on PF instrument (ICC 0.78 to 0.95) and for all on PI instrument (ICC 0.84 to 0.92). DP instrument reliability was excellent (ICC 0.92 to 0.97). The MDC values ranged from ±3.8 to ±5.7 for PF, from ±4.3 to ±6.6 for PI, and from ±3.7 to ±5.2 for DP. The average change from first follow up to 3-4 months (n=15) on each scale {PF (19.3 ±13.1), PI (-12.1±9.8), and DP (-4.8.1±9.5)} were larger than the MDC for PF and PI scales.
**Conclusions:** The PROMIS PF, PI, and DP scales demonstrated good to excellent reliability. The MDC ranged from 3.7 to 6.6 across scales and time points. Smaller MDC are desirable to pick up change in patient status, however, the interval from first follow up to 3-4 months was sufficient to exceed a MDC except with DP scale. More MDC studies are needed to determine if for shorter intervals or smaller changes in PROMIS PF, PI, and DP are clinically useful.

## 2. Longitudinal Measurement Invariance of the Dutch Flemish PROMIS Item Banks for Anxiety and Depression

Gerard Flens, MSc[1], Niels Smits, PhD[2], Caroline B. Terwee, PhD[3], Philip Spinhoven[4], and Edwin de Beurs, PhD[1,4]

**Presenter**: **Gerard Flens, MSc**; gerard.flens@sbggz.nl

**Institutional Affiliations:**[1] Foundation for Benchmarking Mental Health Care, Bilthoven, The Netherlands;[2] Research Institute of Child Development and Education, University of Amsterdam;[3] Department of Epidemiology and Biostatistics, The EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands;[4] Department of Psychiatry, Institute of Psychology, Leiden University

**Objective**: In Dutch mental health institutions, symptoms are routinely assessed with self-report scales. The item scores obtained from the patients result in latent trait estimates that are used to assess individual changes in symptoms over time. In order to properly measure change, a scale must measure the same construct(s) at each time point, that is, must show longitudinal measurement invariance.
**Methods**: We have developed Dutch-Flemish versions of the patient-reported outcomes measurement information system (PROMIS) adult V1.0 item banks for anxiety and for depression. Our goal with these item banks is to use them as input for Computer Adaptive Tests (CAT) in routine assessment during treatment. To determine if repetitive usage of the banks is legitimate, we investigated measurement invariance over time. Each item bank was administered at two time points to N = 500 respondents who started outpatient treatment for an anxiety- or depression disorder. The first administration was at the start of treatment, the second after a minimal interval of 3 months. Dimensionality was studied using confirmatory factor analysis (cfa) and exploratory factor analysis (efa). Measurement invariance was investigated using a series of confirmatory measurement models with an increasing number of model restrictions.
**Results**: we present preliminary results for a subset of cases (approximately N = 100) as data is still being collected.
**Conclusions:** We discuss the findings so far and their implications when using the depression and the anxiety item bank for the assessment of change in symptoms.

## 3. Feasibility and Reliability of the Dutch-Flemish PROMIS Computer Adaptive Tests for Depression and Anxiety

Gerard Flens, MSc[1], Leo. D. Roorda[2], Niels Smits, PhD[3], Caroline B. Terwee, PhD, Philip Spinhoven,PhD[5], and Edwin de Beurs, PhD[1,5]

**Presenter**: **Gerard Flens, MSc**; gerard.flens@sbggz.nl

**Institutional Affiliations**: [1]Foundation for Benchmarking Mental Health Care, Bilthoven, The Netherlands; [2]Amsterdam Rehabilitation Research Center, Reade, Amsterdam, The Netherlands, [3] Research Institute of Child Development and Education, University of Amsterdam, [4] Department of Epidemiology and Biostatistics, the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands, [5]Department of Psychiatry, Institute of Psychology, Leiden University

**Objective**: In the Netherlands, several Dutch-Flemish PROMIS item banks have been evaluated on their psychometric properties, with more to come. At this moment the Dutch-Flemish PROMIS item banks Depression, Anxiety and Pain Behavior are ready for Computerized Adaptive Testing (CAT) and implementation

in daily clinical practice. In this study, we investigate the feasibility and reliability of the depression and anxiety CATs in clinical subjects.
**Methods**: A Dutch-Flemish Assessment Center is being founded to administer CATs to clinical subjects. In this Assessment Center, the item banks, the item parameters and the CAT software (i.e. item selection procedure, scoring procedure, starting and stopping rules) have been implemented. The Center provides CATs through a dedicated website or connects to PROM providers who provide assessments service to Mental Health Institutes. In this pilot study, we connected the Dutch-Flemish Assessment Center to one of the Dutch PROM providers. Moreover, Dutch-Flemish PROMIS Anxiety and Depression CATs were administered in a clinical sample from a Dutch Mental Health Institute. The CATs were administered with a legacy instrument (Symptom Questionnaire [SQ48]) that contains an anxiety and a depression subscale.
**Results**: We will present the Dutch platform that is used to collect CAT data. The three levels of the platform are the Dutch-Flemish Assessment Center, the PROM provider and the Mental Health Institute. Furthermore, we will present the CAT data that are collected so far with this platform.
**Conclusions:** We will discuss the information flow between the various levels of the platform (which information is stored at which level) and explain how we dealt with privacy issues. In addition, we discuss the number of selected items of the CAT's, the total administration time, and the reliability of the scores. These statistics are compared with the outcomes of the SQ48 to assess how promising the PROMIS item banks for Depression and Anxiety are in daily clinical practice.

### 4. Content Validity of the PROMIS Sleep Disturbance and Sleep Related Impairment Item Banks in Adolescents

Jojanneke A.M.C. van Kooten MSc[1], Caroline B. Terwee PhD[2], Gertjan J.L. Kaspers MD PhD[1], Raphaële R.L. van Litsenburg MD, PhD[1]

**Presenting author: Caroline Terwee, PhD**; cb.terwee@vumc.nl

**Institutional affiliations:** [1]Department of Pediatric Oncology-Hematology, VU University Medical Center Amsterdam, The Netherlands, [2]Department of Epidemiology and Biostatistics, VU University Amsterdam, the Netherlands

**Objective:** Sleep problems are common in adolescents and can have a negative impact on daily functioning and quality of life; therefore recognition of sleep problems is important. The PROMIS (Patient-Reported Outcomes Information System) sleep disturbance (SD) and sleep related impairment (SRI) items banks are internationally used, well-validated instruments developed for and tested in adults. To evaluate aspects of sleep specific to adolescents the Sleep in Adolescents Scale (SAS) was designed to complement the PROMIS-SD and PROMIS-SRI for use in adolescents. This study evaluates the content validity of the self- and proxy versions of the PROMIS-SD, PROMIS-SRI and SAS for use in adolescents.
**Methods:** Experts (n=6), adolescents (n=24, 12-18 years) and their parents (n=7) commented on the relevance, comprehensiveness and comprehensibility of the questionnaires.
**Results:** Experts considered all items relevant and only few items were found irrelevant by adolescents and parents. Missing subjects related mainly to nighttime wakening and sleep cognitions. The majority of items were comprehensible. The ability of parents to report on their adolescent's sleep was limited.
**Conclusions:** The PROMIS-SD and PROMIS-SRI have adequate content validity in adolescents. The additional SAS was adjusted according to the feedback to ensure content validity and now evaluates the effect of social media, screen time, school and social activities on five domains (bedtime, sleep quality, daytime fatigue/sleepiness, nighttime awakenings and sleep cognitions). The validity of the proxy scales is limited considering the difficulties reported by the parents. Further psychometric evaluation of these scales in adolescents is ongoing.

### 5. Development of a Dutch-Flemish PROMIS Physical Functioning Short Form for Geriatric Rehabilitation

Ewout B. Smit, MD[1]; Hylco Bouwstra, MD, PhD[1] Caroline B. Terwee, PhD[2].

**Presenter**: **Ewout B. Smit, MD**; eb.smit@vumc.nl

**Institutional Affiliations**: [1]VU University Medical Center, Department of General Practice & Elderly Care Medicine, Amsterdam, the Netherlands
[2]VU University Medical Center, Department of Epidemiology and Biostatistics, Amsterdam, the Netherlands

**Objective:** Development of a Dutch-Flemish PROMIS® Physical Functioning (PF) Short Form that is valid and acceptable for patients in a geriatric inpatient rehabilitation setting.
**Design:** This study consists of two phases. First, a developmental phase, where the short form will be constructed by means of an expert and patient focus group meeting. In the second phase, the draft short form will be tested in 8 to 10 specialized geriatric rehabilitation wards.
*Developmental phase*
To develop a short form with good content validity, experts and patients will be included. First, a focus group meeting will be organized with 6 to 8 experienced professionals working in the field of geriatric intramural rehabilitation. The purpose of this meeting is to make a preselection of relevant items from the 121 items of the PROMIS® data bank PF. In addition, experts will be

asked to add additional items that they consider important for measuring physical functioning in geriatric rehabilitation patients. Second, a focus group meeting will be organized with a diverse sample of 6-8 geriatric rehabilitation patients. They will first be asked to mention relevant limitations in activities. Subsequently, the identified activities by patients will be compared to the selected items by experts and discrepancies will be discussed with the patients. The patients' opinions will be given priority in the selection of items for the short form.
*Test phase*
In the second phase, the draft short form will be tested in another sample group of geriatric rehabilitation patients before finalization. The following psychometric properties of the draft short form will be evaluated: Internal consistency, test-retest reliability, structural validity, responsiveness and interpretability.
**Results:** Results will be presented at the PHO conference.

### 6. Integrating PROMIS Short Forms to a Patient-driven Mobile ICanFunction Assessment tool (mICF)

Heidi Anttila, PhD, PT[1], Kaisa Kokko MSc, PT[1], Sinikka Hiekkala PhD, PT[2], Jaana Paltamaa PhD, PT[3], Petteri Weckström B.Eng. [3]

**Presenting Author: Heidi Anttila, PhD;** heidi.anttila@thl.fi

**Institutional Affiliations**: [1]National Institute for Health and Welfare (THL); Welfare department; Ageing, Disability and Functioning Unit, Helsinki, Finland, [2]The Finnish Association of People with Physical Disabilities (FPD), Helsinki, Finland, [3]Jyväskylä University for Applied Sciences, Jyväskylä, Finland

**Objective:** A feasibility study of a Patient-driven Mobile ICanFunction Assessment tool (mICF) aims to design and develop a first prototype (POC, Proof of Concept) for person-centered assessment of functioning based on ICF (International Classification of Functioning, Disability and Health). PROMIS instruments can be used to operationalize many ICF concepts. The long term aim is to translate several PROMIS item banks into Finnish and to integrate PROMIS Computerized Adaptive Testing (CAT) to the mICF.
**Methods:** Starting March 2016, the items of the Adult Physical Function item bank and the Pediatric Mobility item bank are being translated into Finnish using FACIT translation methodology. After translation, relevant items of these two PROMIS instrument are going to be implemented as a part of the mICF assessment tool. The user experience and acceptance of this technological approach will be tested with children and adults with severe disabilities (short grown people) and their service providers in September-October 2016.
**Results:** Expected benefits of the mICF solution for service user are easier and more reliable self-assessment as well as clearer communication with health care professionals. For professionals mICF can provide a quick summary of the service user's life situation. The solution utilizes national interoperability standards for information systems. The content is based on international ICF codes and functioning information model, which allow structured information collection of both descriptive text and instruments.
**Conclusion:** Reliable health-related information and electronic services, such as mICF will play an important role in reaching the current needs for shared decision-making and goal-setting in healthcare. PROMIS measurements integrated to the mICF solution will enable standardized, precise and reliable self-reported information of functioning.

### 7. Validation of Eight Dutch-Flemish PROMIS Pediatric item banks in Children with Juvenile Idiopathic Arthritis

Hedy A. van Oers, MSc[1], Anouck Splinter, MSc[1], Marion A. van Rossum, MD PhD[2], J. Merlijn van den Berg, MD PhD[3], Dieneke Schonenberg, MD PhD[3], Koert M. Dolman, MD PhD, Martha A. Grootenhuis, PhD[1], Caroline B. Terwee, PhD[4] Lotte Haverman, PhD[1]

**Presenting Author: Lotte Haverman, PhD**; l.haverman@amc.nl

**Institutional Affiliations:** [1]Psychosocial Department, Emma Children's Hospital AMC, Amsterdam, the Netherlands, [2] Amsterdam Rheumatology and immunology Center, Reade, Department of Pediatric Rheumatology, Amsterdam, The Netherlands, [3]Department of Pediatric Hematology Immunology & Infectious Disease, Emma Children's Hospital AMC, Amsterdam, the Netherlands, [4]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, the Netherlands

**Objective**: We aim to validate eight Dutch-Flemish pediatric PROMIS item banks (Anger, Anxiety, Depressive Symptoms, Fatigue, Physical Function - Mobility, Physical Function - Upper Extremity, Pain Interference, and Peer Relationships) in Dutch children with Juvenile Idiopathic Arthritis (JIA).
**Methods**: Children aged 8-18 years with JIA undergoing treatment in three hospitals in Amsterdam: Academic Medical Centre, Reade or Sint Lucas Andreas Hospital were asked by an invitational letter to complete questionnaires online: Pediatric Quality of Life Inventory 4.0. (PedsQL4.0.), Childhood Health Assessment Questionnaire (CHAQ), and eight pediatric PROMIS item banks. Ten days after completing the questionnaires (T1), children were asked to complete the pediatric PROMIS item banks again (T2). Pediatric rheumatologists were asked to provide information on disease activity and number of infected joints. Scores on the pediatric PROMIS item banks will be compared with the PedsQL4.0 and CHAQ sores. Unidimensionality will be assessed by Confirmatory Factor Analysis. Construct validity will be assessed by correlating the PROMIS item banks to the PedsQL4.0, CHAQ and

measures of disease activity. Differential Item Functioning for age and gender will be examined. In addition, test-retest reliability and measurement error will be assessed. Children with JIA undergoing treatment in the University Medical Centre in Leiden (LUMC) will also be asked to participate, to obtain the objective of 200 participants (T1).
**Results**: Results will be presented at the conference. So far, 108 children completed the T1 questionnaires (response rate 39.0%) and 84 children completed T2.
**Conclusions**: If the results are positive, the next step in implementing PROMIS into clinical pediatric practice will be the establishment of Dutch norms in the general population.

### 8. Translation of Selected Items of the PROMIS SexFS v2.0 into Swedish

Johanna Strandquist, PhD, MD[1], Lars E Eriksson, PhD[2, 3] Jeanette Winterling, PhD[1], Maria Wiklander, PhD[1], Kathryn E Flynn, PhD[4], Kevin Weinfurt, PhD[5],Lena Wettergren, PhD[1]

**Presenting author: Lena Wettergren, PhD;**
lena.wettergren@ki.se

**Institutional Affiliations**: [1]Department of Neurobiology, Care Sciences and Society,Karolinska Institutet, Huddinge, Sweden; [2]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden; [3]School of Health Sciences, City University London, Northampton Square, London, United Kingdom; [4]Department of Medicine, Center for Patient Care and Outcomes Research, Medical College of Wisconsin, Milwaukee, WI, USA; [5]Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA

**Objective:** To describe the translation of selected items of the PROMIS Sexual Function and Satisfaction Measures (SexFS) version 2.0 into the Swedish language, and to describe the linguistic validation of the items.
**Methods:** Using forward and backward techniques, we translated 54 items from 13 SexFS domains: Screener, Satisfaction with Sex Life, Interest in Sexual Activity, Orgasm Ability, Orgasm Pleasure, Vaginal Discomfort, Vulvar Discomfort (Labial/Clitoral), Anal Discomfort, Vaginal Lubrication, Erectile Function, Sexual Activities, and Bother. The Swedish back translations and the original English version were compared to assure there were no differences in the meaning of the items. The provisional Swedish version (48 items; anal discomfort excluded) was tested in cognitive interviews with 22 adolescents and adults representing a diversity of gender, age (median 29; range 16-62 years), sexual orientation, ethnicity and education level. The sample included participants with (n=8) and without (n=14) a cancer experience. Interviewees completed the SexFS items and then indicated whether any of the questions were difficult to understand or answer, confusing, upsetting/offensive, and whether he/she would have asked the question in a different way. The translation process was conducted in collaboration with the PROMIS Translation Director and according to the manual provided by FACITtrans and PROMIS.
**Results:** The cognitive interviews revealed that the translated items overall were reasonably well understood. The interviewees proposed minor changes regarding language and order of items, and wanted a clearer highlighting of important words in some items. Furthermore, interviewees wanted to add the response option of "Never do this" for some items assessing sexual activities. Item modifications were made that did not change the meaning of the original items.
**Conclusions:** A final and linguistically validated version of a Swedish PROMIS SexFS v2.0 (selected items), approved by the PROMIS Translation Director, is now available for use in research and clinical practice.

### 9. Danish Translation of the PROMIS Item Bank for Physical Functioning – Results From Cognitive Testing

Christina W. Schnohr, MSc, PhD[1], Henning Langberg, PhD[1], Charlotte L. Rasmussen, MSc[1], Jakob B. Bjorner, MD, PhD[1,2]

**Presenter** – **Christina W. Schnohr, MSc, PhD**;
cwsc@sund.ku.dk

**Institutional Affiliations:** [1]Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark, [2]Optum Patient Insight, Lincoln, RI, USA.

**Objective**: The Danish translation of the PROMIS item bank for physical functioning was performed using standardized methodology. Here, we present results from cognitive testing of the Danish translations.
**Methods:** The cognitive interviews included with 30 adults recruited from rehabilitation centers in the municipality of Copenhagen. To avoid excessive response burden and but ensuring multiple evaluations of each item, five groups of five respondents each evaluated 20% of the items. A final evaluate of revised items were performed by five additional respondents.
**Results**: Generally, the Danish versions of the items were easily understood and easy to administer among the respondents. A few items needed wording changes to fit the Danish context. Although items were developed in an American context, differences in cultural understanding were rare and lead to wording changes only for one item (personal care). However, cognitive interviews identified four types of challenges in answering the questions:
Mismatch between respondents' functional level and item difficulty. Items that were much too hard or much too easy could in some instances cause annoyance or problems in understanding.
Uncertainty regarding use of aids. Elderly respondents and cancer rehabilitation patients were sometimes uncertain whether aids (such as the use of a cane) should be considered when answering the items.

Activities not usually performed by the respondent. Some respondents had difficulty assessing their performance on activities that they did not usually do.
Lack of clarity regarding the context of an item. E.g. on the item: “Are you able to carry a laundry basket up a flight of stairs?” some respondents asked: is the basket empty or full?
**Conclusions**: The standard translation methodology was successful in eliminating problems in translation, but also pointed out the importance of performing cognitive interviews to increase validity of the questionnaire.

### 10. AO Patient Outcomes Center – A Novel Software to Implement PROMIS in Orthopedic Surgeons' Clinical Routine

Alexander Joeris, MD, MSc [1], Mark Vrahas, MD[2], Nan Rothrock[3]

**Presenting Author: Alexander Joeris, MD, MSc;** alexander.joeris@aofoundation.org

**Institutional Affiliations:** [1]AO Clinical, Duebendorf, Switzerland: [2]Massachusetts General Hospital, Boston, USA: [3]Feinberg School of Medicine, No Northwestern University, Chicago, USA

**Objective:** A data collection and reporting software (AO Patient Outcomes Center [AOPOC]) was developed to increase user-friendliness of administering patient-reported outcome measures (PROMs) and therefore provide orthopedic trauma surgeons a PROM tool for use in routine clinical care. All patients used PROMIS Computer Adaptive Tests (CAT). The results of alpha and beta testing of AOPOC are presented.
**Methods:** Alpha phase testing at three hospitals evaluated time to complete each instrument, CAT score distribution, a usability survey completed by patients, and a log of user problems. Additionally, semi-structured interviews were conducted with clinical users to identify usability issues. Beta phase testing at 17 US orthopedic surgery clinics evaluated system performance with an increased volume of users and usability.
**Results:** In alpha testing, 935 patients completed an assessment. The vast majority of patients had no difficulties completing the CATs. 91% expressed willingness to complete similar assessments at future visits. Eleven interviewed clinicians rated user-friendliness of AOPOC as high. Clinic workflow disruption varied by clinic. Challenges included institutional rules concerning patient data in external software, problems with specific data collection devices, and an unclear process for registering patients. Mean completion time for Mobility, Upper Extremity, and Pain Interference CATs was 4.8 minutes (SD=5.2). The Upper Extremity CAT showed 17% of the sample receiving the highest score possible (T-score=56), with longer assessment times and less score precision. In beta testing with 6300 patients, load testing identified one bottleneck contributing to reduced speed. No performance issues were reported. Semi-structured interviews with 5 clinicians reassured user-friendliness of AOPOC. The biggest challenges were in administrative approval for use of AOPOC and difficulty interpreting PROMIS scores.
**Conclusions:** AOPOC is a useful software capturing PROMs in routine orthopedic care. Upper Extremity CAT displayed ceiling effects. PROM collection in clinical practice must address institutional restrictions in data storage and use.

### 11. Development of a Patient-Reported Outcome Instrument for Patients with Lumbar Radicular Pain - STUDY PROTOCOL

Charlotte Ibsen, MSc[1], Berit Schiøttz-Christensen, Professor[3], Claus Vinther Nielsen, Professor[2], Mogens Hørder, Professor[4], Thomas Maribo, Associate Professor[2]

**Presenting Author: Charlotte Ibsen**, **MSc**; chaibs@rm.dk

Institutional Affiliations: [1]DEFACTUM, Central Denmark Region, Denmark,[2]Department of Public Health, Section of Clinical Social Medicine and Rehabilitation, Aarhus University,[3]Spine Centre of Southern Denmark, Hospital Lillebaelt, Middelfart,Denmark,[4]Institute of Public Health, Unit for Patient Perspective, University of Southern Denmark

**Objective:** To develop and pilot test a patient-reported outcome (PRO) instrument based on the International Classification of Functioning, Disability and Health (ICF) to assess functioning in patients with lumbar radicular pain.
**Methods:** The PRO will be developed with direct input from patients with lumbar radicular pain (n=10) and clinicians with extensive experience assessing and treating the population (n=10). The development of the PRO will be based on ICF Comprehensive Core Set for Low Back Pain and Rehabilitation Set. The development process will be divided into three steps. First, patients and clinicians select meaningful categories from the two ICF sets. A three round web based consensus process will be used, and specified criteria for consensus will determine which ICF categories should be included in the PRO. Secondly, items from the PROMIS® Physical Functioning item bank will be linked to ICF, and items corresponding to the included categories will be identified. Thirdly, patients and clinicians will select suitable PROMIS® items to be included in the PRO by means of another web based consensus process. The PRO will then be pilot tested (Alfa test) among the participating patients and clinicians. Adjustments will be done due to results from the Alfa test, and a Beta test among patients in the target population will be performed (n=100). In both Alfa and Beta test the patients have to respond additional questions to test face validity, construct- and content validity. The development process starts in august 2016, with completion December 2017.
**Results**: The development process and preliminary results will be presented at the conference.
**Conclusions:** The study develops a PRO, based on ICF and PROMIS® Physical Functioning item bank, to be used in clinical practice in order to systemize and qualify the

description of functioning among patients with lumbar radicular pain.

### 14. German Translation of Six Item Banks from the Patient-Reported Outcomes Measurement Information System (PROMIS)

Fionna Klasen, PhD[1], Michaela Dabs[1], Veronika Ottová-Jordan[1], Laura Schlepper[1], & Ulrike Ravens-Sieberer[1]

**Presenting Author: Fionna Klasen, PhD;**
f.klasen@uke.de

**Institutional Affiliations:**
[1]University Medical Center Hamburg-Eppendorf, Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Research Section Child Public Health

**Objective:** The German pediatric PROMIS project is part of the "Patient Reported Outcome Measurement Information System" (PROMIS®) initiative. The current project aims to translate and linguistically validate six pediatric item banks for self- and proxy-reports from American English into universal German.
**Methods** Six pediatric item banks (anxiety, anger, depressive symptoms, fatigue, pain, and peer relationships) were translated by using standardized methodology approved by the PROMIS Statistical Center. The translation process included the following steps (i) preparation of materials, (ii) two forward translations for German, Austrian, and Swiss, respectively, (iii) reconciliation, (iv) back translation of independent translator, (v) back translation review, (vi) and harmonization. Subsequently, the items were tested in cognitive interviews with 58 children and adolescents from Germany (16), Austria (22), and Switzerland (20) for the self-report as well as with 42 parents and other care-givers (Germany (12), Austria (17), and Switzerland (13)) for the proxy-report.
**Results** Translators rated the translation difficulty of most items (94.9%) as easy or feasible. Overall, subsequent pretesting showed that items of the universal German version were understood as they were intended, as only 14 out of 82 items of the self-report and 15 out of 82 items of the proxy-report required rewriting. We will report in detail on translation and pre-testing processes.
**Conclusions:** The methodology used and experience gained in this study can be used as an example for researchers in other countries interested in translating PROMIS scales. The German item banks will shortly be ready for use. Implications for future research and prevention programs will be discussed.

### 15. The Ability to Capture Sexual Dysfunction in Young Cancer Patients Using Items of the SexFS

Lena Wettergren, PhD[1], Johanna Strandquist[1], Claudia Lampic[1], Lars E Eriksson[2, 3]

**Presenting Author: Lena Wettergren, PhD;**
lena.wettergren@ki.se

**Institutional Affiliations**: [1]Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Huddinge, Sweden, [2]Department of Learning, Informatics, Management and Ethics, Medical Management Center (LIME, MMC), Karolinska Institutet, Stockholm, Sweden;; [3]School of Health Sciences, City University London, Northampton Square, London, United Kingdom.

**Objective:** We aimed to investigate the ability of the PROMIS Sexual Function and Satisfaction measures (SexFS v2.0) to capture sexual dysfunction among young adults with cancer.
**Methods:** Participants having sexual problems who had been diagnosed with cancer during the last 5 years were recruited through advertisement in newspapers/journals, bulletin boards, patient organizations and social media. Those signing up (n=13) participated in an exploratory trial to test the feasibility of a web-based intervention to alleviate sexual problems in young adults with cancer. The intervention, carried out over 8 weeks, had educational and behavior change content including multimedia (pictures, video vignettes and audios), interactive online activities (e.g. self-monitoring) and partial feedback support (discussion forum, tailored feedback from experts).
Sexual function was measured using 48 items covering 12 domains of the PROMIS SexFS, assessing sexual function the last 30 days in those being sexually active. The SexFS was answered before and after the intervention. Additionally, selected participants were interviewed about their sexual problems. The sexual problems reported in the SexFS were compared with what was recorded in the interviews.
**Results:** Despite that all of the participants had voluntary signed up to the program because they experienced sexual problems this was not as prominent in the results PROMIS SexFS.
**Conclusions:** It appears that PROMIS SexFS might have limitations in its ability to capture sexual dysfunction among patients with cancer, as patients may avoid sexual activities due to severe sexual problems e.g. pain, bleeding, lack of desire or erectile dysfunction.

## 16. Responsiveness of the DF-PROMIS-PB and the DF-PROMIS-PI Item Banks in Patients Presenting in Musculoskeletal Practice

Wouter Schuller, MD [1,2], Caroline B. Terwee, PhD [1], Francisca Galindo-Garre, PhD [1], Niels Smits, PhD [1], Daphne Rohrich, MSc [1], Raymond W.J.G. Ostelo, PhD [1,3], Henrica C.W. de Vet, PhD [1]

**Presenting Author: Wouter Schuller, MD;**
w.schuller@vumc.nl

**Institutional affiliations:** [1]VU University Medical Center, Department of Epidemiology & Biostatistics and the EMGO+ Institute for Health and Care Research, van der Boechorststraat 7, 1081 BT, Amsterdam, The Netherlands, [2]Spineclinic, Mahoniehout 10-12, 1507 ED, Zaandam, The Netherlands, [3]Department of Health Science of the faculty of earth and life sciences and the EMGO+ Institute for Health and Care Research, VU University, Amsterdam, The Netherlands

**Objective:** The aim of our study was to examine the responsiveness of the Dutch-Flemish PROMIS Pain Behaviour item bank (DF-PROMIS-PB) and the Dutch-Flemish Pain Interference item bank (DF-PROMIS-PI) in patients presenting in musculoskeletal practice.
**Methods:** Patients were recruited to fill in a web based survey at baseline and after a follow-up period of three months. All patients completed the full DF-PROMIS-PB and DF-PROMIS-PI item banks, supplemented at follow-up by 7-point Global Perceived Effect (GPE) rating scales. Tailored to the main complaint subgroups of patients completed legacy instruments measuring functional limitations: the Roland Disability Questionnaire, the Neck Disability Index, the Lower Extremity Function Scale, the Disabilities of Arm Shoulder or Hand and the Headache Impact Test-6. Areas under the ROC Curve (AUC), Effect Sizes (ES) and Standardized Response Means (SRM) of all instruments were compared, as well as correlations between the PROMIS T-score changes, the change scores of the legacy instruments and the GPE. Hypothesis were specified.
**Results:** 923 Patients completed both baseline and follow-up measurements. Correlation with the GPE was 0.46 for the DF-PROMIS-PI item bank and 0.28 for the DF-PROMIS-PB item bank. Correlations between the DF-PROMIS-PI item bank and the legacy instruments ranged from 0.34 (NDI) to 0.69 (DASH). Correlations between the DF-PROMIS-PB item bank and the legacy instruments ranged from 0.38 (RDQ) to 0.53 (DASH). Responsiveness parameters for the DF-PROMIS-PB item bank were lower than those of the legacy instruments (ES 0.36, SRM 0.45, AUC 0.62). Responsiveness parameters for the DF-PROMIS-PI item bank were comparable to those of the legacy instruments (ES 0.49, SRM 0.57, AUC 0.74).
**Conclusions:** For the DF-PROMIS-PI item bank 75% of our hypotheses were met, demonstrating good responsiveness, comparable to the legacy instruments. For the DF-PROMIS-BP almost none of our hypotheses were met, showing limited responsiveness in a population of patients with musculoskeletal pain.

## 17. Sexual Interest, Satisfaction, and Associated Factors in People with Chronic Pain

Dagmar Amtmann[1], PhD; Fraser Bocell[1] PhD, Kendra S. Liljenquist[1], PhD; Rana Salem[1], MA, Karon Cook[2], PhD, Kevin Weinfurt[3], PhD

**Presenting Author: Dagmar Amtmann, PhD**;
dagmara@uw.edu

**Institutional Affiliations:** [1]University of Washington, [2] Northwestern University**,** [3] Duke University

**Objective:** To examine the level of sexual interest and satisfaction with sex life in people with chronic pain (CP), and their associations with physical (i.e., pain, fatigue) and psychological (e.g., depression, anxiety) functioning.
**Methods:** Self-reported survey was used to collect demographic, clinical and health information in a cross-sectional study with individuals with CP. Measures included PROMIS interest in sexual activity (interest), satisfaction with sex life (satisfaction), interference of pain on sexual satisfaction (interference), pain during sex, pain intensity, physical function, fatigue, depression, and anxiety. For PROMIS sexual function domains, a t-score of 50 represents the average for sexually active U.S. adults. Linear regression models were used to examine the relationship between specific patient characteristics and the outcomes of sexual interest and satisfaction with sex.
**Results:** Participants (N=230) reported on average less interest (M=42.2, SD=9.7) and less satisfaction (M=46.7, SD=6.6) than the norm (both $p<0.0001$). Of those who engaged in sexual activity in the past 30 days (n=83, 36%), 43% reported moderate to severe pain (≥5 on a 0-10 scale) during sexual activity and 18% reported that pain interfered with their satisfaction quite a bit to very much. Male sex and younger age were associated with more sexual interest.
**Conclusions:** Fewer than 40% of people with CP engaged in sexual activity in the past month. People with CP in this study reported lower average sexual interest and satisfaction than the average U.S. adult. About two fifths reported pain during sexual activity, which led to significant interference with sexual satisfaction for some. Pain didn't seem to be related to whether or not people had had sex, but when they did have sex, a significant minority said that pain interfered with their satisfaction. Chronic pain appears to negatively affect sexual function. More research is needed to better understand sexual function in people with CP.

**18. Sensitivity to Change of the PROMIS Pain Interference Scale in People with Low Back Pain**

Fraser D. Bocell, PhD[1], Dagmar Amtmann, PhD[1]

**Presenting Author: Fraser D Bocell**, **PhD**;
fbocell@uw.edu

**Institutional Affiliation:** [1]University of Washington

**Objective:** The goal of the study was to provide evidence for the sensitivity to change of the PROMIS pain interference (PROMIS-PI) scores in people with low back pain (LBP) receiving an epidural steroid injection (ESI).
**Methods:** The data were collected in a prospective, longitudinal study. Adults with LBP that lasted at least six weeks and were scheduled for an ESI completed a web-based survey at baseline (BL) prior to injection, and one month (T2) and three months (T3) post-injection. Participants also completed a global rating of change scale (GRC) on a 5 point scale (from "much worse" to "much better") at T2 and T3. Change in PROMIS-PI scores from BL to T2 and BL to T3 were calculated. Descriptive statistics, including the Standardized Response Mean (SRM), were calculated for all levels of GRC. GRC categories were then combined into "worse", "about the same" and "better", to limit the number of tests. One-way ANOVAs were conducted to evaluate differences in PROMIS-PI by GRC levels, followed by pairwise contrasts. Differences in PROMIS-PI scores at T1 and T2 between different GRC levels were also tested.
**Results:** The means and SRM followed the expected patterns, with positive or near zero mean differences for participants reporting worse pain, and negative for those reporting better pain. The ANOVAs were statistically significant ($p < .001$, $\omega^2 = 0.13$ to .25). Pairwise contrasts indicated the PROMIS-PI differentiated between patients reporting worse and better pain, and between no change and better pain. All significant differences were further supported by large effect sizes (Cohen's d = 0.57 to 1.16).
**Conclusions:** Results provide strong evidence for the responsiveness of the PROMIS-PI to change in pain interference. The PROMIS-PI instruments can be used to measure change in people with LBP across time.

**19. Translation of the Concept of Fatigue in PROMIS Measures**

Benjamin Arnold, MA[1], Emily Parks-Vernizzi[1], Helena Correia, Lic[2]

**Presenting Author: Benjamin Arnold**, **MA**;
barnold@facit.org

**Institutional Affiliations:** [1]FACITtrans, Elmhurst, IL, USA**,** [2]Northwestern University Feinberg School of Medicine

**Objective:** Translation and cross-cultural adaptation in the context of item banking requires special preparation. Ideally, all items in each bank would be translated simultaneously by the same translation team. This is not always possible due to financial and time constraints, as well as the inclusion of previously translated legacy items in the banks. The objective of this study is to report on challenges and solutions when translating multiple fatigue-related concepts within the same measure, using the translation of the PROMIS Fatigue Short Form 8a into 10 languages as an example.
**Methods:** The PROMIS Fatigue item bank has 95 items, with several short forms derived from those items. The SF8a was translated into Chinese (Simplified & Traditional), Korean, Malay, Polish, Romanian, Russian, Serbian and Ukrainian, according to the FACIT Translation Methodology, an iterative process of forward, back-translation, expert review, harmonization across languages and cognitive interviewing.
**Results:** The PROMIS Fatigue bank contains ten terms to describe degrees of fatigue: tired, fatigue, sluggish, run-down, physically drained, exhausted, bushed, totally drained, wiped out and extreme exhaustion. All were considered when translating the three terms appearing in the SF8a. It was found that many languages do not have as many terms to describe this fatigue spectrum. In addition, it was difficult to match the intended degree of severity between English and other languages, particularly in cases when items had already been translated in separate translation efforts and could not be changed.
**Conclusions:** When translating fatigue-related items from items banks, it is essential that the content of the entire bank and any existing translations be taken into consideration from the beginning of the process. In extreme cases in which conceptual equivalence cannot be achieved for all terms used in the source, it may be necessary to omit items from the target language version of the bank.

**20. A strategy for the Integration of Item-Banking into National Medical Quality Registries.**

John Chaplin, PhD[1], Karin Lycke[2], Lillemor Egertz[3], Eva Helmersson[4], Leszek Stawiarz[2]; Jan Hillert[2]

**Presenting Author: John Eric Chaplin, PhD**;
john.chaplin@gu.se

**Institutional Affiliations:** [1]Inst. Clinical Sciences, Sahlgrenska Academy at University Gothenburg; [2]Karolinska Institute, Stockholm; [3]Stockholm County

Council, Stockholm; [4]Swedish MS Registry, Stockholm, Sweden.

**Objective:** To develop a strategy for the translation and evaluation of a subset of Neuro-QoL item banks for use in the Swedish Neuro-registries/Multiple Sclerosis registry (NEUROreg/MSreg). The SMSreg, which was officially launched in 2001, has been web-based since 2004. It currently includes data on 15,600 of Sweden's estimated 19,500 prevalent patients with MS. Recent development includes a drive for patient-reported outcomes (PRO). Capturing PRO data via the Neuro-QoL item banking is being explored. Neuro-QoL provides a clinically relevant and psychometrically robust tool for assessment of quality of life (QoL) in neurological disorders.
**Method:** A multidisciplinary team including patient representation was established. A subset of Neuro-QoL measures were identified representing the needs for data collection in MS and best suited for a successfully proof of concept. The subset included two full item banks (cognition and fatigue) and four short-form item banks (upper and lower extremity function, positive affect, satisfaction, and ability to participate in social roles and activities). A protocol for the translation and cultural adaptation of the item banks was agreed, which followed the standardised approach outlined by the Neuro-QoL assessment centre. This included two forward translations, one back translation, three reviewers, cognitive debriefing and a finalisation review.
**Results:** There was a close agreement with the original items, however some subtle differences in meaning between the English and Swedish translations, relating to clarity of the original constructs, mean that before Neuro-QoL can be considered valid for use in Sweden, through cognitive debriefing and review steps should be taken.
**Conclusions:** The methodology used and experience gained in this study will be used as an example for collaboration with other projects looking at the integration of PROMIS Item banking into the Swedish national quality registry work. NeuroQoL will contribute to the integration of the patient's perspective.

### 21. The Cultural Adaptation of the PROMIS®-25 Pediatric Profile v1.1 to a Swedish Population: Progress and Outlook

John Eric Chaplin, PhD[1], Karin Enskär[2], Charles Taft[3], Aina Danielsson[1,4]

**Presenting Author: John Eric Chaplin, PhD**; john.chaplin@gu.se

**Institutional Affiliations:** [1]Institute of Clinical Sciences, Sahlgrenska Academy at the University of Gothenburg; [2] CHILD research group, School of Health and Welfare, Jönköping University; [3] Institute of Health and Care Sciences, Sahlgrenska Academy at the University of Gothenburg; [4] Department of Orthopaedics, Sahlgrenska University Hospital, Sweden

**Objective:** To report on the Swedish translation of the PROMIS-25 Pediatric Profile, discuss issues arising from linguistic validation and provide an outlook for future research in the Swedish Pediatric Orthopaedic Quality registry which collects information regarding the development, treatment and outcome of the most common conditions in children's orthopedics. New methods to record patient-reported treatment outcomes (PROs) are needed. The PROMIS-25 is a collection of short forms containing "high information" items which are suited to inclusion in registry work.
**Methods:** Collaboration was established with the team conducting the translation of pediatric item banks in Sweden and the items not included in that project were translated into Swedish. The FACIT translation methodology was followed with both projects working in tandem. Translations of all items in the PROMIS-25 were cognitively debriefed with 20 native-speaking children (aged 8-16 years; 11 boys) to assess relevance, comprehensibility, and appropriateness. Qualitative analysis of the comments was used to culturally adapt the translation.
**Results:** Translated items were well understood by participants; however, differences were found regarding some terms that were directly translated versus culturally adapted. For example, the question "It was hard for me to walk one block when I had pain" generated several alternative Swedish equivalents for the term 'block'. Age differences in understanding were also noted. For example, the question "Other kids wanted to be my friend" was difficult for younger children to answer due to its abstract nature. Translations were revised as needed, and re-tested to ensure cultural appropriateness, conceptual equivalence and harmonization with the English version.
**Conclusions:** Swedish pediatric items are conceptually equivalent to the English source version and can be used in registry research and clinical research and practice. The PROMIS-25 Pediatric Profile is now being used in a validation study in Sweden. Future development of the instrument will include Swedish calibration.

### 22. Functional Impact of Congenital Hand Differences: Results from the Congenital Upper Limb Differences (CoULD) Registry

Donald S. Bae MD[1], Patricia L. Connell MPH[1], Maria F. Canizares MD[1], Patricia E. Miller MS[1], Summer Roberts MA[2], Andrea Bauer MD[1], Carley Vuillermin MBBS FRACS[1], Lindley B. Wall MD[2], Peter M. Waters MD[1], Charles A. Goldfarb MD[2].

**Presenting Author: Patricia Connell, MPH;**
patricia.connell2@childrens.harvard.edu

Institutional Affiliations: [1]Boston Children's Hospital, Boston, MA. [2]Shriners Hospital for Children, St. Louis, MO.

**Objective:** To characterize the functional, emotional and social impact of congenital upper limb differences on affected children and families using validated functional outcome instruments.

**Methods:** From June 2014 to April 2016, children with congenital upper limb differences from two pediatric hospitals have been enrolled in the prospective Congenital Upper Limb Differences (CoULD) registry. Pertinent clinical data as well as functional outcomes using the Pediatric Outcomes Data Collection Instrument (PODCI) and Patient Reported Outcomes Measurement Information System (PROMIS) were collected. PROMIS modules included upper extremity function (UE), pain, anxiety, depression, and peer relationships.

**Results:** Overall, 213 PODCI and 76 PROMIS questionnaires were completed by parents of children <11 years; 82 PODCI and 60 PROMIS were completed by adolescents ≥11 years. Median patient age was 7.8 years (range 2 to 17.1 years). Median PODCI scores were >90 for all domains except pain and upper extremity function. Median PROMIS UE was 49 (IQR 40 to 57), but pain, anxiety, depression, and peer relationships scores were 41 (IQR 34 to 51), 42 (IQR 34 to 51), 35 (IQR 35 to 51) and 57 (IQR 47 to 64), respectively. Patients with entire limb involvement had higher PROMIS pain scores (46 vs. 34, p=0.01), lower PODCI UE (78 vs. 88, p=0.003) and global functioning (90 vs. 93, p= 0.007) than those with differences limited only to the hand. Compared with those with bilateral involvement, patients with unilateral differences reported higher scores for PODCI sports (96 vs. 89 p=0.006) and global functioning (94 vs. 89, p=0.003), lower pain scores (100 vs. 100 p=0.001), and better PROMIS UE (41 vs. 35, p=0.03). Additional orthopaedic conditions and medical comorbidities negatively influenced all PODCI scores and PROMIS pain and UE scores (p<0.05).

**Conclusions:** Children with congenital hand differences report decreased upper limb function but better scores in psychosocial domains.

### 23. Improvement in PROMIS Physical Function After Treatment for Copperhead Snake Envenomation: Findings from a Randomized Clinical Trial Comparing CroFab® (Crotalidae Polyvalent Immune Fab [ovine]) and Placebo

Rajiv Mallick, PhD[1], Maria Gasior, MD[1], Eric J Lavonas, MD[2].

**Presenting Author: Rajiv Mallick, PhD**;
rajiv.mallick@btgplc.com

**Institutional Affiliations:** [1]BTG International Inc., West Conshohocken, PA**,** [2]Denver Health and Hospital Authority, Denver, CO

**Objective:** Out of 9,000 patients treated in US emergency departments each year for crotaline snakebite, approximately half are due to bites by copperhead snakes (*Agkistrodon contortrix*). Although not fatal, most patients develop pain and swelling and experience impairment on activities of daily living. As part of a randomized, double-blind, study comparing CroFab(R) (Crotalidae Polyvalent Immune Fab [ovine]) versus placebo with rescue treatment for copperhead snake envenomation, we assessed return to US norms on patient-reported physical function (PF) based on the Patient-Reported Outcomes Measurement System (PROMIS) Physical Function 10a (PF 10a) short form.
**Methods:** Eligible patients with copperhead envenomation were randomized (in a ratio of 2:1) to either treatment with CroFab (R) or with placebo. In the first two years of the trial, PF was assessed in terms of the SF-36 PF domain; subsequent to a study amendment, it was measured using the PROMIS PF 10a short form at days 7 (d-7), 10, 14, 17, 21), 24, and 28 after envenomation. The SF-36 PF scores were converted to PROMIS PF T-scores (normalized to US norms mean 50; standard deviation (SD 10) based on a recently published cross-walk algorithm. Treatment group differences in PROMIS T-scores over time were compared using repeated measures, stratified, ANOVA linear mixed-effects-based least square (LS) means. Both modified intent to treat (m-ITT) population (all patients who received study treatment) and per protocol (PP) population (patients who completed the primary d-14 assessment on the primary endpoint, the Patient Specific Functional Scale (PSFS), with no major protocol deviations) analyses were conducted.
**Results:** In 74 m-ITT patients (Crofab® 45; placebo 29), LS mean PROMIS T-scores increased from mean 45.1 (SE 1.9) and 42.8 (2.0) in the Crofab and placebo groups (p=0.241) respectively, at d-7 to 52.4 (1.9) and 48.7 (2.0) (p=0.066) at d-14. Findings were similar in the PP population (p=0.062 across treatment groups at d-14). PROMIS PF T-scores further increased to 56.2 (1.9) and 54.4 (2.0) at d-21 and to 59.2 (2.0) and 58.2 (2.0) at d-28, in the m-ITT Crofab® and placebo groups respectively.
**Conclusions:** Copperhead snake-envenomated patients were at least 0.5 SD below PROMIS PF US norms 7 days after envenomation but recovered to US PF norms 14 days after envenomation in the Crofab® group, and 21 days after envenomation in the placebo group.

### 24. Development and Psychometric Evaluation of the Pediatric Oral Health Item Banks

Christopher B. Forrest, MD, PhD, [1,2] Brandon D. Becker, MPH, [1] Rachel E. Teneralli, MS, [1] Jean Hee Moon, PhD, [1] and Katherine B. Bevans, PhD[12]

**Presenting Author: Christopher B. Forrest, MD, PhD;** forrestc@email.chop.edu

**Institutional Affiliations:** [1]Children's Hospital of Philadelphia; [2]University of Pennsylvania School of Medicine

**Objective:** To describe the development and psychometric evaluation of pediatric patient reported outcome (PRO) measures of oral pain/inflammation and wellbeing (satisfaction with appearance and socio-emotional experiences related to oral health).
**Methods**: We used the PROMIS methodology to generate initial oral pain/inflammation and wellbeing item pools based on systematic literature reviews, semi-structured interviews (n = 6 dentists, 24 parents, 27 youth), cognitive interviews (n = 17 parents, 21 youth), and reading level analysis. Item pools, socio-demographic and health status questionnaires, and existing oral health PRO measures were administered to 2,265 children (8-17 years) and 2,863 parents (of children 5-17 years). Participants were recruited from a probability-based internet research panel. Psychometric analyses were conducted using data weighted according to US population benchmarks for child age, gender, race, ethnicity, geographic region, and household income.
**Results:** Pediatric dentists, parents, and children identified oral pain/inflammation and wellbeing as important and meaningful outcomes that are best assessed by patient and/or proxy report. Initial item pools included 62 pain/inflammation and 42 wellbeing items. Cognitive interviews results in the removal of 9 pain/inflammation items. Final calibrated items banks met item response theory assumptions of unidimensionality, item local independence, and monotonicity. They discriminated among children with a wide range of oral health experiences and without differential item functioning. Final child- report banks include 45 pain/inflammation items and 26 wellbeing items. Final parent- report banks include 45 pain/inflammation items and 24 wellbeing items. Items were selected for inclusion in 8- and 4-item short forms to ensure adequate precision across the full range of the latent traits and to maximize clinical utility. Construct validity was demonstrated through expected group differences and associations with existing measures of oral health.
**Conclusions:** Child- and parent-report oral pain/inflammation and wellbeing item banks and accompanying short

## 25. Progress in the Cultural Adaptation of Four PROMIS® Pediatric Item Banks for Use in Sweden.

Martin Stackelberg, PhD[1], Karin Blomberg, RN, PhD[2] Mats Eriksson, RN, PhD[1], John Eric Chaplin, PhD[2]

**Presenting Author: John Eric Chaplin, PhD;** John.chaplin@gu.se

**Institutional Affiliations:** [1]Institute for Health Sciences, Örebro University Sweden [2]Institute of Clinical Sciences, Sahlgrenska Academy at the University of Gothenburg, Sweden

**Objective:** To report on the Swedish translation of four pediatric item banks - Pediatric Anxiety, Peer Relationships, Fatigue and Depressive Symptoms - and the corresponding parent items, discuss issues arising from linguistic validation and provide an outlook for future research.
**Methods:** The FACIT translation methodology was followed with rigorous forward and backward translation. Collaboration was established with the team conducting the translation of PROMIS-25 Pediatric Profile. Translations of all items were cognitively debriefed to assess relevance, comprehensibility, and appropriateness. Qualitative analysis of the comments was used to culturally adapt the translation.
**Results:** Translated items were well understood by participants; however, differences were found regarding some item where Swedish didn't have words to distinguish between two different English terms. For example, the questions "I felt alone" and "I felt lonely" became difficult as both "alone" and "lonely" translates into the same Swedish word "ensam". It was thereby hard to find a fitting synonym for "ensam" to grasp the differences in alone/lonely. Other differences were also found. For example, the question "I was too tired to go out with my family" has no good equivalents as there is cultural difference in how children/teenagers are expected to spend time with their families. Translations were revised as needed, and re-tested to ensure cultural appropriateness, conceptual equivalence and harmonization with the English version.
**Conclusions**: The four Swedish pediatric item banks with corresponding parent items are conceptually equivalent to the English source version and can be used in research and clinical practice. Future research using these item banks is now being planned as a validation study in Sweden.

## 26. Improving and Expediting Access to PROMIS CAT Data in Orthopaedic Practice

Jason Guattery, MS, Ryan P. Calfee, MD MSc

**Presenting Author: Jason Guattery, MS;** guatteryj@wudosis.wustl.edu

**Institutional Affiliations:** Department of Orthopaedic Surgery, Washington University in St. Louis School of Medicine

**Objective:** To increase the meaningful use of PROMIS scores by improving access to score data alongside a standard set of queried fields from a native EMR. This access allows providers to look at their patient data for both QI and review preparatory to research applications in an expedited and vastly more efficient manner than traditional methods.

**Methods:** A multi-disciplinary workgroup including members from the department of Orthopaedic Surgery, Reporting department, IT department discussed possible options for data access and retrieval and the most cost-efficient manner of access is to generate a custom report within existing business reporting software that providers can log in to access as they need It for QI purposes and review preparatory to research. This standard report is comprised of 13 fields as well as PROMIS CAT scores that members of the department of Orthopaedic Surgery deemed to have near universal applicability for QI and research across different divisions. The data can be exported in numerous common formats to accommodate a number of different data analysis programs.
**Results:** A custom report was generated within Tableau reporting software that included PROMIS CAT scores as well as 13 different data fields from the EMR. Access permissions were developed within Tableau reporting software for providers to see this report on their own patient population. As mentioned above, this data is solely for QI purposes and review preparatory to research. It is important to note that use beyond this scope requires IRB approval and gathering through traditional channels.
**Conclusions:** Streamlining and expediting access to PROMIS CAT scores can be done for providers utilizing existing reporting technology. While there are limitations on the use of this data, the ability to speed up both QI research and review preparatory to research represent a major improvement in the efficiency of data access for the department.

### 28. Understanding the Contribution of PROMIS to PCORI-Funded Patient-Centered Comparative Clinical Effectiveness Research

Mary Kay Margolis, MPH, MHA[1]; Lori Frank, PhD[1]; Jason Gerson, PhD[1]; Rachel Witsaman, MPH, PMP[1]; Ethan Basch, MD, MSc[2]; David R. Flum, MD, MPH[3]; Cynthia J. Girman, DrPH[4]

**Presenting Author: Lori Frank, PhD**; lfrank@pcori.org

**Institutional Affiliations**: [1]The Patient-Centered Outcomes Research Institute (PCORI), [2] University of North Carolina, [3]University of Washington Department of Surgery, [2]CERobs Consulting, LLC

**Objective:** PCORI funds patient-centered comparative clinical effectiveness research (CER) to help patients make informed health and healthcare decisions. The goals of patient-centered CER and the measurement goals of the Patient Reported Outcomes Measurement Information System (PROMIS) - to provide valid and responsive measures of health and well-being – overlap. Given the specific PCORI funding for PROMIS-focused studies, a review of the use of PROMIS measures within the PCORI portfolio was completed to understand the contribution of PROMIS in patient-centered CER.
**Methods:** As part of a comprehensive coding taxonomy developed for PCORI's funded projects, use of PROMIS measures was flagged.
Research plans for these projects were reviewed to describe outcome domains assessed.
**Results:** In the 359 projects awarded from December 2012 through February 2016, 56 (16%) used at least one 1 PROMIS measure. These projects were funded through each of the five national priority areas: Accelerating PCOR and Methodological Research (n=11); Addressing Disparities (n=6); Assessment of Prevention, Diagnosis, and Treatment Options (n=24); Communication and Dissemination Research (n=4); and Improving Healthcare Systems (n=11). Eight of the 11 Methodological Research projects were funded through the specific PROMIS targeted funding. Of the 45 CER-specific projects, 16 (35%) used PROMIS Global Health (10 in combination with at least one other PROMIS measure). Each of the seven PROMIS domains were measured in 10 or more projects, with measurement of Social Roles in 10 projects, Sleep Disturbance in 11 projects, Physical Function in 15 projects, Pain Interference and Fatigue in 16 projects each, Depression measurement in 17 projects, and Anxiety measurement in 19 projects.
**Conclusions:** Use of PROMIS within PCORI-funded studies extends beyond projects funded through the PROMIS-targeted funding. The use of the seven PROMIS domains all contribute to the link between PROMIS measurement and patient and clinician decision-making in this set of patient-centered CER projects.

### 29. Implementation of PROMIS in Sweden

Evalill Nilsson[1] & John Eric Chaplin, PhD[2]

**Presenting Author: Evalill Nilsson, MD, PhD**; evalill.nilsson@ki.se

**Institutional affiliations:** [1]Quality Register Center QRC Stockholm, Stockholm, Sweden & Karolinska Institutet, Stockholm, Sweden**,** [2]Inst. Clinical Sciences, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

**Objective:** The objective is the smooth integration of PROMIS item banking into the Swedish National Quality Registries (NQRs).
**Background:** Sweden, as a pioneer of NQRs in the 1970s, now has over 100 registries, each focusing on specific diagnoses and interventions. NQRs have enabled continuous quality improvement at the clinical and the political/administrative levels. NQRs are now in the process of incorporating patient-reported outcome data. A thorough investigation conducted in 2015 concluded that Sweden should implement PROMIS at a national level in order to take advantage of future computer-assisted technology options, the possibilities of large-scale data analysis across conditions and healthcare systems.
**Methods:** A project team, financed by the NQR organisation, has been established to undertake the following actions during 2016 as the first phase of integration:

1) survey the existing PROMIS instruments and item-banking projects in Sweden.
2) pilot a PROMIS item bank, made available to the NQRs through the new national e-health platform.
3) establish a collaborative national integration network, and Scandinavian and international partnership.
4) integrate the patients' perspective through co-design from the start of the process.
5) develop a long-term strategy plan for item banking in Sweden, and initiate the next phase of the implementation process.

**Discussion:** The implementation of a sustainable nationwide change in the way PROM data is collected is a major challenge. Such a strategy will have to address the resistance to replacing established PRO instruments, the current limited awareness of item banking in Sweden, the training requirements for staff and patients, and the integration of prior PRO data. It is considered vital for smooth integration to involve patient groups and future users as co-designers and advisors from the start of the project. Co-design will ensure that item banking will be rapidly integrated into the Swedish NQRs.

### 31. Using a Multi-Modal, Longitudinal Collection Strategy to Collect PROMIS Scales Across a Healthcare Delivery System

Randi Zegman, BA[1]Josie Elias, MBA, MPH[2], Peter Meyers,BS[3], Ronnen Cretu,BA[4], Neil Wagle, MD, MBA[5]

**Presenting Author: Randi Zegman, BA;**
rzegman@partners.org

Institutional Affiliation: Partners Healthcare, Boston, MA

**Objective**: The PROMs (Patient Reported Outcome Measures) program at Partners Healthcare is spearheading a patient care initiative to collect longitudinal PRO data system-wide. Since the inception of the program in 2012, more than 130,000 questionnaires have been collected across Partners institutions in over 20 service lines. The PROMIS-10 global health scale, along with additional PROMIS scales, have been incorporated into many questionnaires collected to date. The Partners PROMs team regularly advocates for the inclusion of PROMIS measures in new questionnaires. This presentation describes the PROMs team's strategy to encourage widespread adoption of PROMIS scales through a multi-modal digital platform.

**Methods**: Questionnaires are collected electronically in-clinic via tablet devices (Apple iPads). The PROMIS-10 scale appears in every questionnaire in three month intervals, regardless of condition, to provide a baseline assessment of global health. Additional PROMIS scales are used in different specialties. In the initial phase of the PROMs program, questionnaires were collected through a vendor's (Tonic) iPad application. Patients received follow-up emails periodically and were able to complete the questionnaire from their home computers. The questionnaire platform for the PROMs program has now transitioned to Epic Systems. Patients receive appointment-based questionnaires on iPads via Epic's Welcome kiosk application. Patients also have the opportunity to complete questionnaires in their patient portal (MyChart) before an appointment. Real-time questionnaire responses, including the PROMIS-10 and other PROMIS scales, are incorporated directly into the patient's chart for review by a clinician during the encounter. Patient responses and scores are displayed in a graphical format to show the patient's progress over time.

**Results:** The expansion of the PROMs program (from 3 pilot clinics to over 60 clinics in three years) has led to significant growth in PROMIS collection rates. For example, more than 85,000 PROMIS-10 scales were completed in the Tonic system. In the last ten months, more than 20,000 PROMIS-10 scales have been completed in Epic. The multi-modal, longitudinal platform enables collection at a steadily increasing rate of over 1,000 questionnaires per week. Adding the PROMIS scales to questionnaires across service lines (e.g. cardiology, orthopedics, and urology) ensures standardization throughout the network. The combination of in-clinic and home collections via electronic tools captures a broader patient base than in-clinic collections alone.

**Conclusion:** The Partners PROMs program facilitates the widespread collection of PROMIS scales by enabling in-clinic and home collection. The large amount of data will allow the analytical team to make comparisons across conditions in the Partners network. Once Epic makes CAT functionality available, there will be further adoption of PROMIS instruments throughout Partners. PROMs plans to expand to all clinics in the Partners system within the next several years.

### 32. Dutch norm scores for three Dutch-Flemish PROMIS Item Banks

Martine H.P. Crins, MSc[1], Leo D. Roorda, PhD, MD, PT[1], Thomas Klausch, PhD[2], Niels Smits, Phd[3], Henrica C.W. de Vet, Prof[2], Rene Westhovens, Prof, MD[4,5], Maarten Boers, MD[2,6], Joost Dekker, Prof[7], Caroline B. Terwee, PhD[2].

**Presenting Author: Martine H.P. Crins, MSc;**
M.crins@reade.nl

**Institutional Affiliations:** [1] Amsterdam Rehabilitation Research Center | Reade, Amsterdam, The Netherlands, [2] Department of Epidemiology and Biostatistics, The EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands, [3] Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands,[4] Department of Development and Regeneration, Skeletal Biology and Engineering Research Center, KU Leuven, Louvain, Belgium,[5] Rheumatology, University Hospitals, KU Leuven,

Louvain, Belgium, [6] Amsterdam Rheumatology and immunology center, VU University Medical Center, Amsterdam, The Netherlands,[7] Department of Rehabilitation Medicine and Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands.

**Objective:** To facilitate interpretation and implementation of PROMIS in the Netherlands, we aimed to obtain Dutch norm scores for the Dutch-Flemish PROMIS Physical Function (DF-PROMIS-PF), Pain Behavior (DF-PROMIS-PB) and Pain Interference (DF-PROMIS-PI) Item Banks.
**Methods:** About 2300 persons of the general Dutch population will complete a web-based survey, of which 1300 persons will complete the full DF-PROMIS-PF (121 items, 5-point Likert scale) and the other 1000 persons will complete the full DF-PROMIS-PB (39 items, 6-point Likert scale) and the full DF-PROMIS-PI (40 items, 5-point Likert scale). The sample will be stratified for gender, age, education, and ethnicity according to the distribution of the 2015 census of the general Dutch population. A one-factor confirmatory factor analysis will be performed to assess unidimensionality per item bank. A graded item response model (GRM) will be fitted per item bank to evaluate the item characteristics of the item banks and to facilitate future development of computer adaptive tests (CATs). Ordinal regression models will be used to evaluate Differential Item Functioning (DIF) for language (Dutch vs. English) as a measure of cross-cultural validity. Furthermore, the impact of using Dutch versus English (US) item parameters on T-score estimates will be evaluated.
Based on the results it will be decided if Dutch-specific item parameters are needed or if US item parameters can be used in Dutch populations. Based on that decision Dutch norm scores will be computed.
**Results and conclusion:** The data-collection is in progress until June 2016. The results will be presented at the conference.

### 33. Calibration and Validation of the Dutch-Flemish PROMIS Physical Functioning Item Bank in Patients Treated with Physiotherapy

Martine H.P. Crins, MSc[1], Philip van der Wees, PhD, PT[2], Thomas Klausch, PhD[3], Simone van Dulmen, PhD PT[2], Leo D. Roorda, PhD, MD, PT[1], Caroline B. Terwee, PhD[3]

**Presenting Author: Martine H.P. Crins, MSc;**
M.crins@reade.nl

**Institutional Affiliations:** [1] Amsterdam Rehabilitation Research Center | Reade, Amsterdam, The Netherlands, [2] IQ healthcare, UMC St. Radboud, Nijmegen, The Netherlands, [3] Department of Epidemiology and Biostatistics, The EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands

**Objective:** To assess the psychometric properties of the Dutch-Flemish PROMIS Physical Function item bank (DF-PROMIS-PF) in Dutch patients treated with physiotherapy.
**Methods:** 753 Dutch patients who were treated with physiotherapy (in the year prior to the survey) completed a web-based survey, including the full DF-PROMIS-PF (121 items, 5-point Likert scale). A one-factor confirmatory factor analysis was applied to assess unidimensionality. A graded item response model (GRM) was fitted to evaluate the item characteristics of the item bank and to facilitate future development of a computer adaptive test (CAT). Relationships with legacy instruments were calculated to evaluate construct validity. Ordinal regression models were used to evaluate Differential Item Functioning (DIF) for language (Dutch vs. English) as a measure of cross-cultural validity. Furthermore, DIF for body region (affection in the neck vs low back vs knees etc.) was evaluated. The study results were compared to the results of previous studies of the DF-PROMIS-PF in patients with chronic pain and rheumatoid arthritis. In addition, Dutch norm scores are being collected for the DF-PROMIS-PF.
**Results:** The interim analyses support unidimensionality of the DF-PROMIS-PF (CFI=0.974 and TLI=0.973). The data of the item bank fit the GRM, and showed good coverage across the physical function continuum. Analyses of construct validity, cross-cultural validity and DIF for body region are in progress and will be presented at the conference.
**Conclusions:** The interim results indicate that the items of the DF-PROMIS-PF fit a GRM and indicate that the DF-PROMIS-PF can be used to develop a CAT for measuring physical function in Dutch patients treated with physiotherapy.

### 34. Reducing Ceiling Effects in PROMIS® Physical Function Measures by Using an Extended Response Scale

Gregor Liegl, Mag[1], Felix Fischer, PhD[1], Barbara Gandek, PhD[2 3], Jacob Bjorner, MD, PhD[4 5], John Ware, PhD [2 3], Matthias Rose, MD, PhD [1 2], James F. Fries, MD[6], Sandra Nolte, PhD [1 7]

**Presenting Author: Gregor Liegl, Mag**;
gregor.liegl@charite.de

**Institutional Affiliations:** [1]Department of Psychosomatic Medicine, Center for Internal Medicine and Dermatology, Charité – Universitätsmedizin Berlin, Berlin, Germany [2]Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA [3] John Ware Research Group, Worcester, MA, USA [4]Optum

Patient Insights, Lincoln, RI, USA [5]Department of Public Health, University of Copenhagen, Copenhagen, Denmark [6]Department of Immunology and Rheumatology, Stanford University School of Medicine, Palo Alto, CA, USA [7]Population Health Strategic Research Centre, School of Health and Social Development, Deakin University, Burwood, Australia

**Objective:** To investigate whether an extended response scale increases the range of measurement in static PROMIS® Physical Function (PF) measures.
**Methods:** In the course of PROMIS wave 1 data collection, a 5-item PF short form was presented in three different item formats. Two of these formats are used in the PROMIS PF item bank, both utilizing a response scale with five response categories in which the highest PF category indicates performing an activity "without any difficulty" or degree of limitation expressed as "not at all". In this study, a third item format was presented using six response categories in which the highest levels of function in performing an activity are "easy" and "very easy". We compared format-specific item information curves between the three versions of the 5-item short form, using a Graded Response Model which included the different versions of the 5-item short form and all remaining items of the PROMIS PF item bank. We additionally simulated response patterns for five groups with different PF levels ("very low" to "very high"; n=10,000/group) to calculate relative validity (RV) coefficients for each format-specific short form indicating the relative power to distinguish between known groups compared to the full item bank (RV=1.0).
**Results:** More than n=8,500 participants responded to a subset of the five items presented in different item formats. Compared to the two standard item formats (difficulty/limitation), the extended six-category format increased the range of measurement by more than 0.5 standard deviations on the latent PF continuum. It also showed higher power to distinguish between groups with "high" and "very high" PF (RV=0.92) than the difficulty/limitation formats (RV=0.79 and RV=0.78, respectively).
**Conclusions:** Using an item format with an extended response scale is a promising and efficient way to increase measurement range in patient-reported PF measures with the potential to reduce ceiling effects.

### 36. Creating Condition-Specific PROMIS® Assessments: An Overview of the Gap Analysis Method

Karen Kaiser, PhD[1], Karen Schifferdecker, PhD MPH[2], Kathleen Carluzzo, MS[2], Amy Eisenstein, PhD[1], George Greene, PhD[1], Susan Yount, PhD[1], Dave Cella, PhD[1], Elliott Fisher, MD MPH[2]

**Presenting Author: Susan Yount, PhD**;
s-yount@northwestern.edu

**Institutional Affiliations:** [1]Northwestern University Feinberg School of Medicine; [2]Geisel School of Medicine at Dartmouth

**Objective:** The Patient Reported Outcomes Measurement Information System® (PROMIS®) provides generic measures that are applicable across populations and chronic conditions. Creating a methodology for integrating condition-specific measures with generic measures such as PROMIS will provide both general health and tailored disease-specific assessments. As part of a mixed-methods study, we developed a "gap analysis" method to aid in creating condition-specific assessments. Here we describe the gap analysis process used for osteoarthritis of the knee (OA-K).
**Methods:** The gap analysis utilized data from 8 focus groups (FG) with OA-K patients and consisted of: 1) coding FG transcripts using a thematic analysis approach; 2) grouping FG findings within PROMIS® domains and identifying existing PROMIS items to represent those findings; 3) reducing the item pool based upon subjective and psychometric item quality; 4) writing new items based on FG findings not aligned with existing PROMIS content (i.e., "gap areas"); and 5) grouping new items within existing PROMIS domains or creating new domains.
**Results:** An initial pool of 200 PROMIS items from 9 domains was created to represent the coded data. After review, 52 were retained. Thirty-seven potential gap areas were identified based upon codes not aligned with PROMIS content. Areas deemed not relevant or not amenable to patient report were removed; 24 items were written to cover the remaining gaps. Most new items fell within 6 existing PROMIS domains. Three new domains (life satisfaction, symptoms, independence) were identified.
**Conclusions:** The gap analysis method requires a team with expertise in measurement and qualitative methods. A more deductive approach to coding that utilizes the PROMIS framework could shorten this process. Despite its shortcomings, the gap analysis method pinpoints existing PROMIS content and aids in augmenting PROMIS with new, condition-specific content.

### 37. Impact of PROMIS on Patient Satisfaction

Marsha Jensen, BSN, RN, CPHQ1, David Mitten MD[1], Judith Baumhauer MD MPH[1]

**Presenting Author: David Mitten, MD**;
Dave_Mitten@urmc.rochester.edu

**Institutional Affiliations: University** of Rochester Medical Center, Rochester NY

**Objective:** To examine the impact of real-time implementation of the PROMIS on patient satisfaction as measured by CAHPS scores.
**Methods:** Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys were obtained from patients treated in a busy Orthopaedic practice 6 months before and after implementation of PROMIS (CAT Physical Function,

Pain Interference, and Depression). Orthopaedic provider ID numbers were obtained and these providers were divided into two groups: Those who viewed the PROMIS scores more than 150 times as documented in Erecord (high PROMIS utilizers) and those who viewed PROMIS less than 150 times (low PROMIS utilizers) . Chi-square analyses were used to compare the 9 CAHPS questions pre- and post- PROMIS implementation between the two groups. A second analysis examined the individual provider CAHPS scores pre and post PROMIS. A p-value of <0.05 denotes significance.
**Results:** 172 Pre-PROMIS, 276 Post PROMIS CAHPS surveys were available for review for the high PROMIS utilizers, and 1185 Pre-PROMIS, 2010 Post PROMIS surveys were available for the low PROMIS utilizers. The between group finding was the high PROMIS utilizers scored significantly higher than the low utilizers on "Recommend this provider office" ($p<0.03$) in the Post PROMIS period. The individual high PROMIS utilizers scored significantly higher post PROMIS implementation on "Show respect for what you say" ($p<0.04$) than in the Pre-PROMIS period.
**Conclusion:** Viewing PROMIS scores in the clinic positively impacts patient satisfaction scores.

### 38. The Use of PROMIS CAT Post-Surgical Scores in Assessing "Value" in ACL Reconstruction Surgery.

Owen Papuga PhD, Michael Maloney MD, Ilya Voloshin MD, Gregg Nicandri MD, John Goldblatt MD, Robert Bronstein MD, Lucien Rouse MD

**Presenting Author: M. Owen Papuga, PhD;**
**Owen**_Papuga@urmc.rochester.edu

**Institutional Affiliations**: University of Rochester, Department of Orthopaedics

**Objective:** There is a lack of standardized outcome measures or methods to study the effects of treatment and to effectively monitor and document patient outcomes. We have used Patient Reported Outcome Measurement Information System (PROMIS) computed adaptive testing questionnaires collected during routine orthopaedic office visits to describe the pre-surgical and postsurgical physical function (PF) and pain interference (PI) to compare outcomes after ACL reconstruction surgery performed by 6 separate providers using various surgical techniques.
**Methods:** A retrospective analysis of 403 patients undergoing ACL reconstructive was conducted to compare the baseline PROMIS PF and PI scores to those scores after surgery. Longitudinal comparisons of PROMIS scores were then made to identify any significant differences between providers. PROMIS scores were assessed using a general linear Analysis of Variance (ANOVA).
**Results:** Average PROMIS PF and PI scores were similar for all providers. Figures 1and 2 below show the average outcomes for all providers along with 95% confidence intervals. The ANOVA general linear (PROVIDER x TIME) model found no statically significant differences between the providers with at least 10 patients for each chosen time point (p=.073). However, Tukey Post hoc analysis identified significant differences between individual providers PROMIS PF and PI scores at the 6 month follow up. No significant differences in patient baseline scores were found between providers (p=.498)
**Conclusion:** If fully utilized, these outcome measures can help inform physicians in individual surgical cases, as well as be used to identify surgical techniques with improved outcomes. However, if there is are no differences in outcome, techniques that require less time, have reduced costs, or are less complex could be considered to have a higher value. There is the potential for cost savings and improved patient safety.

### 39. A Pilot Study of PROMIS and Patient Reported Outcomes in Thalassemia

John Chapin, MD[1],Dorothy Kleinert, NP[2], Lisa A. Mandl, MD MPH [3], Sujit Sheth, MD[2], Patricia-Jane Giardina, MD[2]

**Presenting Author: John Chapin, MD;**
jchapin29@gmail.com

**Institutional Affiliations:** [1]Weill Cornell Medicine-New York Presbyterian Hospital Department of Medicine, Division of Hematology-Oncology, New York, NY, [2]Weill Cornell Medicine -New York Presbyterian Hospital Department of Pediatrics, Division of Hematology-Oncology, New York, NY, [3]Hospital for Special Surgery, Division of Rheumatology, New York, NY

**Objective:** Adult patients with transfusion-dependent thalassemia (TDT) can expect to live into their 5th decade as a result of safer transfusion practices and improved management of iron overload. However, TDT patients develop multiple co-morbidities as they age. Patient reported outcomes (PROs) are important in understanding disease impact, and are key measures of efficacy for new therapeutics. The NIH's Patient Reported Outcome Measurement Information System (PROMIS) utilizes item response theory, a psychometric measurement method that is administered electronically via Computer Adaptive Testing (CAT). This allows greater measurement precision, while at the same time decreasing responder burden. We hypothesize that PROMIS is feasible and acceptable to adult patients with TDT.
**Methods:** Patients will be enrolled from the New York Comprehensive Thalassemia Center. A lead-in phase of written qualitative responses to disease-specific questions will be conducted. Subsequently, PROs will be administered at scheduled transfusions. Domains will be compared to similar domains from disease-specific and validated instruments (TranQoL, SF-36). **Results:** Primary outcomes will be the ability to complete PROMIS, and the comparison with domains of other instruments. Qualitative responses to questions will be analyzed by Word Cloud and analysis for elicited themes. Secondary outcomes will be

the impact of co-morbidities on scores and qualitative patient experiences. **Conclusion:** We endeavor to demonstrate that PROMIS can be expanded to a network of Thalassemia Centers in North America to prospectively measure the impact of new therapeutics on quality of life.

### 40. Development and Content Validation of a PROMIS Measure for Medication Adherence

Fredericksen, RJ., PhD MPH,[1] Brown, S., MSW MPH,[1] Edwards, TE, PhD, [2] Yang F., PhD, [3] Alperowitz-Bichell, K., MD, [4] Godfrey, M., MPH, [5] Wang, A.[4] Church, A. [1] Fitzsimmons, E. [1] Gutierrez, C. [6] Paez, E. [7] Dant, L., MPH, [6] Loo, S., MPH, [6] Walcott, M., PhD, [8] Mugavero, M., MD, [8] Mayer, K., MD, [6]Mathews, C., MD, [7] Patrick, DL., PhD, [2] Crane, PK., MD, [1] Crane, HM., MD[1]

**Presenting Author: Rob J. Fredericksen, PhD, MPH**; rfrederi@uw.edu

**Institutional Affiliations:** [1]University of Washington, [2]University of Washington/Seattle Quality of Life Group, [3]Georgia Regents University, [4]Chase Brexton Health Services, [5] Beaufort Jasper Hamilton Community Health Services, [6] Fenway Community Health, [7] University of California-San Diego, [8]University of Alabama-Birmingham

**Objective**: Develop a clinically relevant measure of medication adherence behaviors/barriers for clinical care using mixed qualitative/quantitative methods. The qualitative methods are described here.
**Methods:** We identified validated measures of adherence behaviors and barriers and placed them into an item pool and winnowed pool content using PROMIS Qualitative Item Review (QIR) criteria. Item pool content areas informed development of an interview guide for exploring how primary care patients conceptualize adherence behaviors and barriers. We conducted 1:1 interviews with 218 patients prescribed ≥1 of the following medication classes: antiretrovirals, antihypertensives, cardiovascular medications, diabetes medications, anticonvulsants, antidepressants, mood stabilizers, and antipsychotics. We interviewed English and Spanish speakers from 6 geographically diverse U.S. community health and HIV clinics. We coded interviews into 2 behavior and 8 barrier content areas and matched interview excerpts to item pool content. We evaluated unmatched excerpts for potential new item development. We then conducted two rounds of cognitive interviews with patients (n=37) to assess comprehension and interpretation of developed items.
**Results:** Participants were 36% female, 38% African American, 22% Latino, and 14% spoke Spanish as their primary language. We developed a measure that included 2 behavior areas: changes to regimens or pills, and adherence-promoting behaviors. The 8 barrier areas were health-related barriers, activity-related barriers, attitudes/beliefs, access issues, regimen attributes, disease/medication stigma, social factors, and simple forgetting.
**Conclusion**: Through QIR and content validation of new and legacy adherence behavior and barrier items, we developed a clinically relevant, multi-dimensional measure of medication adherence behaviors and barriers for use with patient populations living with chronic disease for which close medication adherence is necessary in order to avert illness or death.

### 41. Impact of Depression on Physical Function and Pain following Foot and Ankle Fracture

Judith Baumhauer MD MPH[2], Bryant Ho, MD[2], A. Samuel Flemister, MD[2], John Ketz, MD[2], Benedict DiGiovanni, MD[2],Irvin Oh, MD[2], Jeff Houck, PT, PhD[1]

**Presenting Author: Judith Baumhauer, MD, MPH;** Judy_Baumhauer@urmc.rochester.edu

**Institutional Affiliations:** [1]George Fox University, [2]University of Rochester

**Objective:** Describe the number of patients affected by depression (mild to severe) during the first 6 months after a foot and ankle fracture. Determine if experiencing mild or greater depression on initial presentation after an ankle fracture is associated with 6 month pain, physical function, or depression PRO scores.
**Methods:** PROMIS physical function (PF), pain interference (PI), and depression (DEP) were administered to all orthopedic foot and ankle patients (From April 2015 to May 2016) at the University of Rochester Medical Center. Complete data sets, including a time point <30 days after and >180 days after initial treatment were used for analyses. ICD-9 code diagnoses included (n=44): bimalleolar (n=4), Trimalleolar (n=7), Lateral malleolar (n=9), Astragus (n=1), Ankle (n=6), Calcaneal (n=8), Cuboid (n=2), metatarsal (n=5) and foot bones (n=4). PROMIS Depression scores were classified: 55-64=mild, 65-74=moderate, >74=severe. A two way ANOVA was used to compare the 6-9 month PROMIS PF and PI between fracture patients who experienced mild or greater depression to those who did not experience depression.
**Results:** A total of 21/44 (47.7%) of patients experienced mild or higher depression [7/44 (15.9%) were moderate depression and 1/44(2.3%) were classified as severe depression]. A total of 7/44 (15.9 %) patients continued to have mild or greater depression at 6 months. Average length of depression lasted 135 ± 94.8 days. Only 2/44 (4.5%) patients began to experience depression after initial evaluation during the course of treatment. However, 4/44 (9.1%) patients experienced improvement/decline in depression across the 6 months of care. At 6 months there was no main effect (p=0.30) or interaction effect (p=0.31)

indicating differences in PF or PI for patients that experienced depression.
**Conclusions:** Mild clinical depression is common after a foot or ankle fracture. However, few patients experience moderate or severe depression. While depression rarely started after the initial evaluation and resolved for most patients by 6 months, the length of depression can last 1-4 months. Depression does not affect 6-9 month PF or PI outcomes.

### 42. Comparison of Patient Reported Outcomes 6-13 Months after Ankle Fracture and Hip Fracture

Jeff Houck, PT, PhD[1], Bryant Ho, MD[2], Kyle Judd[2], Samuel Flemister, MD[2], John Ketz, MD[2], Benedict DiGiovanni, MD[2],Kyle Judd, MD[2], Irvin Oh, MD[2], Judy Baumhauer, MD MPH [2]

**Presenting Author: Jeff Houck, PT, PhD;**
jhouck@georgefox.edu

**Institutional Affiliations:** [1]George Fox University, [2]University of Rochester

**Objective:** To determine differences in recovery using PROMIS physical function (PF), pain interference (PI), and depression (DEP) between two lower extremity fractures (ankle and hip).
**Methods:** PROMIS PF, PI, and DEP were administered to all orthopedic patients (From April 2015) at a University Hospital. Complete data that included a time point 6 to 13 months after their post-surgery visit. Ankle fracture records included diagnoses (ICD-9 codes): bimalleloar, triamlleolar, ankle, and lateral malleoli fracture (n= 32). For hip fracture diagnoses (ICD-9 code): neck, midcervical, intertrochanteric, and intracapsular (n=24). Because initial post-surgery PROMIS scores were similar, a two way ANOVA was used to compare the 6-13 month PROMIS measure between fracture type (ankle versus hip) controlling for *age and gender*. A logistic regression analysis was also used to predict achieving a minimal clinically important difference (MCID) using baseline variables.
**Results:** PROMIS PF was significantly lower at follow up for patients with hip fracture ($p<0.01$), however, PI ($p=0.47$) and DEP ($p=0.08$), were comparable. The lower PF after hip fracture compared to ankle fracture was marked (Difference=7.8 95%CI (2.3–13.3)). Significantly greater ankle fracture patients (69.7%) achieved an MCID as compared to hip fracture patients (37.5 %) ($p=0.047$). However, neither fracture type achieved a mean PF score of 50 by 6-13 months (Hip=34.1±9.2 versus Ankle=42.6±8.0). Logistic regression revealed that baseline PF ($p<0.01$) and age ($p<0.01$) were independent predictors of an MCID or greater change in PROMIS PF (overall accuracy 75%).
**Conclusions:** This data demonstrates the ability of PROMIS PF to discriminate differences in outcome between two common orthopedic trauma diagnoses. The strong differences in outcome could be estimated from baseline PF data and age. The markedly low scores at 6-13 month follow up suggest further research and ongoing clinical strategies to address physical function are warranted for both fracture types.

### 43. Comparison of PROMIS Physical Function, Pain Interference, and Depression at Initial Clinical Assessment of Foot and Ankle Soft Tissue Injuries

Jeff Houck, PT, PhD[1], Bryant Ho, MD[2], Samuel Flemister, MD[2], John Ketz, MD[2], Benedict DiGiovanni, MD[2], Irvin Oh, MD[2], Judy Baumhauer, MD MPH [2]

**Presenting Author: Jeff Houck, PT, PhD;**
jhouck@georgefox.edu

**Institutional Affiliations:** [1]George Fix University, [2]University of Rochester

**Objective:** Describe and compare initial evaluation PROMIS physical function (PF), pain interference (PI), and depression (DEP) of patients with orthopedic foot and ankle soft tissue injuries.
**Methods:** PROMIS PF, PI, and DEP were administered to all orthopedic foot and ankle patients (From April 2015) at a University Hospital. Complete data collected at the initial treatment for specific ICD-9 codes were included (n=131). A two way ANOVA was used to compare differences across health domains (PF, PI, and DEP) and soft tissue diagnoses (Achilles (n=35), Foot Tenosynovitis (n=24), Posterior Tibialis (n=19), Peroneal (n=10), Ankle Sprain (n=43)).
**Results:** A significant interaction effect indicated patient responses depended on both health domain and diagnosis ($p<0.001$). Pairwise comparisons indicated that PF for patients with ankle sprain (36.2 ± 10.8) was significantly lower than some tendon injuries (Achilles 46.2 ± 12.6 ($p<0.01$), Foot Tenosynovitis 43.3 ± 10.3 ($p<0.01$)) and similar to others (Posterior Tibialis 41.6 ± 5.8 ($p=0.06$)). Peroneal (40.3 ± 5.0 ($p =0.267$)). PI was significantly higher after Ankle Sprain (63.1 ± 1.2) as compared to Achilles (55.9 ± 1.3($p<.01$)) and Foot Tenosynovitis (58.2 ± 1.6)($p=0.01$), however was similar to PT (61.6 ± 1.7($p=0.43$) and Peroneal (61.0 ± 2.4($p=0.48$). Pairwise comparisons for DEP indicated significantly higher scores for ankle sprain (50.4 ± 11.3) compared to Foot Tenosynovitis (44.6 ± 8.8) ($p = 0.02$), however, all diagnoses were less than 50.
**Conclusions:** Knowing expected PROMIS PI, PF and DEP scores for specific foot and ankle problems are useful in clinical decision making. These data indicate that ankle sprains, an acute injury, result in similar severity as chronic tendinopathies such as posterior tibial and peroneal tendon problems. None of the foot and ankle injuries were associated with depression (DEP>55). The severity of PT and Peroneal tendinopathy motivates continued efforts to address PF for these problems.

**44. Correlation of PROMIS Upper Extremity, Pain Interference and Depression scores with SST and ASES scores**

Aaron M. Chamberlain MD, MSc[1]; Ryan P. Calfee MD, MSc[1]; Jay D. Keener MD[1]; Ken Yamaguchi MD, MBA[1]; Devon C. Nixon MD[1]

**Presenting Author: Aaron M. Chamberlain, MD;** chamberlaina@wustl.edu

**Institutional Affiliations:** [1]Department of Orthopedic Surgery, Washington University in St. Louis, St. Louis, MO

**Objective**: The American Shoulder and Elbow Surgeons (ASES) score and the Simple Shoulder Test (SST) are traditional and widely accepted patient-reported outcome (PRO) measures for assessing shoulder outcomes. PROMIS surveys have been created to quantify patient outcomes data in multiple domains including the Upper Extremity (UE CAT), Pain Interference (PI CAT) and Depression and Physical Function (PF). No study has examined the correlation of PROMIS UE, PI and Depression CAT scores with ASES score and SST. Also, no prior study has compared the UE CAT and PF CAT in patients with Shoulder complaints.
**Methods**: We retrospectively reviewed all patient visits from August 2015 – March 2016 who presented to a single Shoulder clinic. Patients completed ASES and SST paper-based questionnaires and the UE, PI, Depression and PF PROMIS surveys. Pearson correlation coefficients were used to describe relationships for all instruments. Regression analyses were used to determine if significant variables from correlations and group analyses predicted ASES, SST, and PROMIS UE CAT scores. Mean UE CAT and PF CAT scores were compared.
**Results**: We identified 119 patients with complete ASES, SST, and PROMIS scores. PROMIS UE, PI and Depression scores correlated with the ASES and SST scores ($p < 0.001$). Step-wise regressions indicated that PROMIS PI scores alone explained statistically significant portions of the variance for each respective PRO. Mean UE CAT scores were 31.6 (SD 9.11) and mean PF CAT scores were 40.5 (SD 8.91).
**Conclusions**: Our results describe the strong correlative relationship of PROMIS UE, PI and depression CAT with ASES and SST scores. PROMIS PI results were the strongest predictors of disability for all three PRO instruments. UE CAT scores describe greater disability compared to PF CAT scores in the same patient population. Our findings support the validity of PROMIS UE, PI and depression scores in describing shoulder outcomes.

**47. Multidimensional Higher-Order Factor Structure and Diagnostic Accuracy of PROMIS Mental Health Domain**

Carlos G. Forero, PhD[1], Gemma Vilagut, MSc[2], Mónica Astals, PhD[3] Adelina Abellanas[3], Cristóbal Díez-Aja, MD[3], José-Ignacio Castro- Rodríguez, MD MPh[3], José Manuel López-Santín MD[3], Miguel Garriz, PhD[3], Carmen Sanchez-Gil, MD[3], Itxaso Alayo, MSc[1], Jordi Alonso, MD, PhD[2]

**Presenting Author: Carlos García Forero, PhD;** cgarcia@imim.es

**Institutional Affiliations:** [1]CIBER Epidemiología y Salud Pública (CIBERESP), Spain, [2]Health Services Research Unit, Hospital del Mar Medical Research Institut (IMIM) Barcelona, Spain, [3]Institute of Neuropsychiatry and Addictions (INAD), Barcelona, Spain

**Objective:** To compare different multidimensional models explaining the factor structure of PROMIS Mental Health item Banks
**Methods:** Cross-sectional study with Internet panel data, comparable to Spanish adult general population (n=1,807). We applied multidimensional Graded Response IRT Models to PROMIS Depression, Anxiety and Anger item Banks. We modelled multidimensional structures on item responses: 3 independent traits; second-order trait governing Depression, Anxiety and Anger dimensions, and a bifactor structure with general and specific factors. Fit was assessed based on limited information estimation (criteria RMSEA<0.05, CFI, TLI>0.95). We used Area Under the Curve (AUC) in Receiver Operating Curve analysis to assess diagnostic accuracy of multidimensional factor scores in detecting major depression (MDE) and generalized anxiety disorder (GAD) according to DSM criteria. Population disorder variability explained by the models was assessed (R^2) using MDE and GAD diagnostics as dependent variables in structural models.
**Results:** As expected by the high intercorrelations between dimensions (r>0.75), the independent traits model showed poor fit (RMSEA=0.15, CFI=0.30, TLI=0.29). The Bifactor model best fitted the data (RMSEA=0.026, CFI=TLI=0.98), followed by second-order factor model (RMSEA=0.032. CFI=TLI=0.97). All models showed high and almost identical diagnostic accuracy for MDE (AUC=0.94 to 0.95), and slightly lower for GAD (AUC=0.90). Explained disorder variance was similar across models, significantly higher for MDE (R^2=0.74) than for GAD (R^2=0.52). The second-order and the general factors were the only significant predictors of pathology in their respective models.
**Conclusions:** Dimensionality of the PROMIS Mental Health Item Banks is best explained by a Bifactor model with general distress factor capturing the commonalities between items and specific Depression, Anxiety and Anger dimensions. While unidimensional score estimates suffice for accurate disorder detection, the second-order and bifactor structures use a single piece of information with

similar performance. Multidimensional structures can be efficient and useful for guiding CAT applications aimed for diagnosis.

**48. Construct Validity and Differential Item Functioning of Spanish PROMIS Depression, Anxiety and Anger**

Gemma Vilagut, MSc[1], Carlos G. Forero, PhD[2], Adelina Abellanas[3], Mónica Astals, PhD[3], José Ignacio Castro-Rodríguez, MD MPH[3], Cristóbal Díez-Aja, MD[3], Miguel Garriz, PhD[3], José Manuel López-Santín MD[3], Carmen Sanchez-Gil, MD[3], Karina Mayoral[1], Jordi Alonso, MD, PhD[1]

**Presenting Author: Gemma Vilagut, MSc;** gvilagut@imim.es

**Institutional Affiliations**: [1]Health Services Research Unit, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), Carrer del Doctor Aiguader, 88, Edifici PRBB, 08003, Barcelona, Spain, [2]CIBER Epidemiología y Salud Pública (CIBERESP), Spain, [3]Institute of Neuropsychiatry and Addictions (INAD), Parc de Salut Mar, Barcelona, Spain

**Objective:** Evaluating construct validity and differential item functioning of PROMIS depression, anxiety and anger item banks in a Spanish general population sample.
**Methods:** Cross-sectional study using Internet panel data with distributions by age, sex and region comparable to Spanish adult general population (n=1,807). Calibration and validation of the item banks followed PROMIS standards, using item response (IRT) and classical test theory approaches. IRT assumptions were evaluated, using a Graded Response Model for calibration. Construct validity was evaluated with multi-trait multi-method matrix (MTMM) with legacy and with disability measures. Differential item functioning (DIF) by age, sex and education was assessed with ordinal logistic regression (McFadden's pseudo-R2 change>0.02). Cross-cultural DIF was assessed for language (English vs. Spanish) with the original English PROMIS calibration data.
**Results:** All item Banks fulfilled IRT assumptions. Unidimensionality models showed adequate fit in all item banks, (CFI>0.95 and RMSEA<0.08). MTMM correlations showed a method effect in PROMIS item Banks (correlations >0.7). PROMIS depression correlated higher with its legacy measures than with measures from other domains; anxiety and anger item banks showed some unexpected high MTMM correlations, suggesting overlap in contents of these constructs across instruments. No DIF was observed regarding age, sex and education in any of the item banks.
Regarding language, 1 depression, 2 anxiety and 4 anger items showed signs of DIF (pseudo-$R^2$ change>0.02). However, their impact on the overall scores, as compared to the purified score, was small (absolute value differences <0.06).
**Conclusions:** Results supports good metric properties of the Spanish PROMIS item banks evaluated, especially for depression. Precaution is advised about items showing potential bias, especially if they are selected in computer adaptive test assessments. Further studies addressing potential construct validity issues affecting PROMIS anxiety and anger are recommended.

**49. Determining When Multidimensionality Impacts Total Scale and Subscale Scores: Application of Confirmatory Bifactor Modeling to Osteoarthritis of the Knee (OAK) Patient Responses to the KOOS**

Michael A. Kallen, PhD, MPH [1]; David Cella, PhD [1]; Susan Yount, PhD [1]; Karen E. Schifferdecker, PhD [2]; Kathleen L. Carluzzo, MPH [2]

**Presenting Author: Michael A. Kallen, PhD, MPH;** michael.kallen@northwestern.edu

**Institutional Affiliations:** [1]Northwestern University; [2]Dartmouth University

**Objective:** Item responses to measures may exhibit multidimensionality, yet multidimensionality itself does not support subscale creation nor recommend against the use of total scores. We employed a two-phase approach to confirmatory bifactor modeling to investigate multidimensionality and determine its impact on reliable variance associated with KOOS total and subscale scores.
**Methods:** The KOOS is a 42-item, 5-subscale knee disability measure; subscale scores, not total scores, are typically reported. 1) We compared "gold standard" slope estimates from a multidimensional (six factor) bifactor analysis to those from a unidimensional CFA to explore anticipated unidimensional model slope estimate distortion, implying potential total score estimation distortion. 2) We employed lambda estimates from our multidimensional bifactor analysis (general factor and each of the five specific factors) to calculate omega, omega-H, and omega-S indices and determine attributable reliable variance.
**Results:** N=600 OAK patients completed the KOOS (60% female, 29% 60+ years old, 49% had TKA). 1) Thirteen unidimensional model slope estimates exceeded the 95% CIs of gold standard estimates, including all five Sports items and all four QOL items. 2) *Total score support*: General factor omega-H was .95, compared to its omega of .99, indicating total score reliability was little inflated by multidimensionality; 95% of reliable variance was attributable to the general factor. Specific factor omega-Hs ranged from .00 to .03, supporting the general factor as predominant influence on total score variance. *Lack of subscale score support*: Specific factor omega-Ss ranged from .03 to .37, associated omegas ranged from .90 to .98, and specific factor ECVs ranged from .04 to .39. Thus, specific factors had little reliable variance unique from the general factor.

**Conclusions:** KOOS total score reporting, traditionally not recommended, is preferred to KOOS subscale reporting. Confirmatory bifactor modeling offers a reliable variance-based approach to practical decision-making on total and subscale score reporting and interpretation.

### 50. Chinese-American Rheumatology Patients Using Traditional Chinese Medicine Have Worse PROMIS® Health Status Scores

Kai Sun, MD[1 2], Lily Lee, MS[1], Henghe Tian, MD[3], Jennifer C. Leng, MD, MPH[4], Lisa A. Mandl, MD, MPH[1 2]

**Presenting Author: Kai Sun, MD;** sunk@hss.edu

Institutional Affiliations: [1]Hospital for Special Surgery, [2]Weill Cornell Medicine, [3]New York University Medical Center, [4]Memorial Sloan Kettering Cancer Center

**Objective**: To assess whether self-reported health status is different among users and non-users of traditional Chinese medicine (TCM) among Chinese-American patients with rheumatic diseases.
**Methods**: Subjects were recruited from a rheumatology clinic that serves a predominantly Chinese-American immigrant population. Patients were included if they spoke English or Mandarin Chinese and were being treated for a systemic rheumatic disease or osteoarthritis of the hands, knees, or hips. PROMIS short forms available in English and Simplified Chinese were administered by a bilingual researcher. Domains included sleep disturbance, applied cognition, anxiety, depression, pain interference, physical function, fatigue, and ability to participate in social roles and activities (SRAs). Patients also reported details of TCM use.
**Results**: 52 subjects enrolled, mean age 58 (range 22-97), 63% female, and 79% Medicaid. 79% spoke no English, and 50% reported using any TCM modality. The most common diagnoses were rheumatoid arthritis (35%), lupus (29%), spondyloarthropathies (17%), and other (19%). TCM users had worse anxiety (mean T-score 53 vs. 45, p=0.004), depression (mean T-score 53 vs. 44, p=0.004), pain (mean T-score 59 vs. 54, p=0.04), fatigue (mean T-score 56 vs. 50, p=0.04), function (mean T-score 42 vs. 47, p=0.048), and ability to participate in SRAs (mean T-score 54 vs. 61, p=0.007). Although approved PROMIS Chinese translations were used, 5 (10%) found questions on SRAs confusing and not applicable to their lives.
**Conclusions**: TCM users had worse scores in multiple important health domains compared with non-users. This cross-sectional association should be explored longitudinally. TCM use could be a marker of inadequately treated disease, or may signal non-adherence to prescribed western therapies. Despite using approved Chinese translations, some content was irrelevant to this group. More work is needed to ensure cultural validity of PROMIS questionnaires in the growing Chinese-American immigrant population.

### 53. Using PROMIS to Find Frail Patients Undergoing Total Joint Replacement

Abigail Schmucker, BA[1], Nathaniel Hupert, MD, MPH[2], Mayu Sasaki, MPH[1], Charlotte Ching, BA[1], Steven K. Magid, MD[1 2], Lisa A. Mandl, MD, MPH[1 2]

**Presenting Author: Lisa A. Mandl, MD, MPH**; mandll@hss.edu

**Institutional Affiliations:** [1]Hospital for Special Surgery, New York, NY, [2] Weill Cornell Medicine, Cornell University, New York, NY

**Objective:** Arthroplasty is usually successful, however, 15-25% of patients, especially knee cases, have poor outcomes despite excellent technical success. Pre-operative frailty may be associated with worse outcomes, but frailty assessment is labor- intensive. We investigated whether PROMIS29 domains correlate with grip strength, a strong correlate of frailty, as well as other standard PROs.
**Methods:** Adults > 65yo scheduled for elective knee (TKR) or hip (THR) replacement were recruited. Subjects completed the disease specific KOOS/HOOS as well as the global Short Form-12 (SF-12 v1), CES-D 10, and PROMIS29 v2; surveys were compared using Spearman correlations. Grip strength in the dominant hand was grouped in quartiles by age and gender.
**Results:** 109 community-dwelling patients (median 71yo, 94.5% white, 68.81% female, 56.9% TKR, 43.1% THR). Grip strength was significantly but weakly correlated with PROMIS Fatigue among combined TKR and THR cases and TKR alone, (r=0.20 and 0.26, P-values 0.046 and 0.048), but not with any other domains. PROMIS Physical Function and Social domains correlated with KOOS/HOOS QOL (TKR r=0.5 and 0.66; THR, 0.73 and 0.65; P-values <0.0001); PROMIS Physical Function and Pain Interference correlated with SF-12 PCS (TKR r=0.73 and 0.75; THR 0.75 and 0.62; P- values <0.0001); and PROMIS depression correlated with CES-D score (TKR r=0.57 and THR 0.59; P-values <0.001). In TKR patients, PROMIS anxiety and depression correlated with SF-12 MCS (r=0.63 and 0.53; P- values <0.0001); THR patients showed weaker correlations (r=0.45 and 0.44; p-values <0.01).
**Conclusion:** PROMIS 29 domains show excellent correlation with both disease specific and global instruments in this patient population. PROMIS29 domains appear to be poorly correlated with grip strength. Since this is a key component of the frailty phenotype, further studies should be done to investigate whether other PROMIS short forms might be more useful to identify frail elders.

## 54. The Association of PROMIS® CATS with Clinical Characteristics in Systemic Lupus Erythematosus (SLE) Outpatients

Shanthini Kasturi, MD[1], Jayme C. Burket, PhD[1], Jessica Berman, MD[1], Kyriakos A. Kirou, MD[1], Alana B. Levine, MD[1], Lisa R. Sammaritano, MD[1], and Lisa A. Mandl, MD, MPH[1]

**Presenting Author: Shanthini Kasturi, MD**; kasturis@hss.edu

**Institutional Affiliation:** [1]Hospital for Special Surgery, Weill Cornell Medicine, New York, NY

**Objective:** To evaluate the association of PROMIS CATs with clinical characteristics in SLE outpatients.
**Methods:** Adult SLE patients were recruited from a SLE Center of Excellence. Subjects completed 14 PROMIS CATs in domains of relevance to SLE patients. Subjects self-reported diagnoses of anxiety, depression, and fibromyalgia. SLE disease flare status was measured with the SELENA-SLEDAI. Differences between groups were compared using Mann Whitney U and Kruskal Wallis tests as appropriate.
**Results:** 204 outpatients with SLE completed PROMIS CATs. Subjects with self-reported anxiety (28.4%), depression (27.5%), and fibromyalgia (14.2%) scored worse than those without across all PROMIS domains (p-values all <0.04). Compared to other SLE patients, those with anxiety or depression were approximately one standard deviation worse in the anger, anxiety, and depression CATs, while those with fibromyalgia were one standard deviation worse in the pain interference and ability to participate in social roles CATs (p-value <0.001). Subjects with current SLE flare (19.6%) had worse scores across PROMIS domains than non-flaring subjects (p-values < 0.05) except in fatigue, sleep disturbance, and sleep-related impairment CATs. There was no relationship between flare severity and PROMIS scores. Subjects with arthritis flares (8.8%) scored worse than those with nephritis flares (3.9%) or no flare (80.4%) in CATs relating to mobility, pain, social roles, and cognitive abilities (p-value <0.05). Subjects with nephritis flares scored worse on the anger CAT than all other participants (p-value = 0.01).
**Conclusions:** PROMIS CATs reflect the worse health status of SLE patients with co-morbid anxiety, depression, and fibromyalgia and flaring lupus. Notably, PROMIS scores did not directly vary with severity of flare, and domains of fatigue and sleep were similarly impaired in flaring and non-flaring patients. These findings, as well as the high anger scores in patients with nephritis flares, may reflect unmet needs in this population and require further investigation.

## 55. PROMIS Physical Function: Application to a Complex Rare Disease Population - Tenosynovial Giant Cell Tumor (TGCT)

Heather L. Gelhorn, PhD[1], Xin Ye, PhD, MS[2], Sandra Tong, MD[3], John H. Healey, MD[4], Susan V. Bukata, MD[5], Richard D. Lackman, MD[6], Lindsey Murray, MPH[1], Rebecca M. Speck, PhD[1], Grant Maclaine, BMedSc, MBBS, MEc, MA[7], William R. Lenderking, PhD[1], Henry H. Hsu, MD[3], Paul S. Lin, MD, MBA[3], William D. Tap, MD[4]

**Presenting Author: Heather L. Gelhorn, PhD;** heather.gelhorn@evidera.com

**Institutional Affiliations:**[1]Evidera, Bethesda, MD; [2]Daiichi Sankyo Pharma Development, Edison, NJ; [3]Plexxikon Inc., Berkeley, CA; [4]Memorial Sloan Kettering Cancer Center, New York, NY; [5]Ronald Reagan UCLA Medical Center, Los Angeles, CA; [6]Cooper University Health Care, Orthopaedic Oncology Center, Camden, NJ; [7]Daiichi Sankyo Development Ltd, Buckinghamshire, United Kingdom

**Objective:** Tenosynovial giant cell tumor (TGCT), a rare, locally aggressive neoplasm of the synovium of joints and tendon sheaths, is associated with joint destruction, pain and swelling. Impacts on physical function (PF) vary depending on tumor extent and the location of the affected joint. No validated method exists to measure PF among patients with TGCT. This study was aimed at developing measures of upper and lower extremity PF that could be analyzed on a single latent scale using items selected from the PROMIS Physical Function (PROMIS-PF) item bank.
**Methods:** Patients were recruited for qualitative research interviews to identify predominant symptoms and impacts. Patients completed a checklist containing all 121 PROMIS-PF items, indicating relevant and TGCT-related impacts. The publicly available PROMIS-PF item response theory (IRT) parameters were used to select relevant items representing the range of the latent PF trait in this population.
**Results:** Participants (n=22) were 73% female, mean age 42.5 years. TGCTs were located in the knee (n=15), hip (n=3), ankle (n=2), elbow (n=1), and forearm (n=1). Fifty-four PROMIS-PF items were identified as relevant by ≥20% of the participants. PF concepts discussed by participants during the qualitative interviews were also used to select relevant items. Items with maximal slopes (range: 2.66-4.399) and appropriately targeted thresholds (range: -3.14–0.31) were selected. Items (n=15) were used to create physical function subscales specific to upper (11 items) or lower (13 items) extremity tumors, 9 items overlapped both subscales.
**Conclusions:** We describe a novel method of combining qualitative research and IRT-based item information to select a relevant and content valid subset of items to assess heterogeneous impacts on PF in a rare disease population. The TGCT-specific upper and lower extremity subscales developed using the PROMIS-PF item banks are relevant and valid to assess PF among patients with TGCT.

## 57. Insurance Status and Patient Outcomes in the Orthopaedic Outpatient Setting

Amanda Spraggs-Hughes, MA[1], William M. Ricci, MD[1], Michael J. Gardner, MD[2], Christopher M. McAndrew, MD, MSc[1]

**Presenting Author: Christopher McAndrew, MD;** mcandrewc@wudosis.wustl.edu

**Institutional Affiliations**: [1]Washington University School of Medicine, St. Louis, MO, [2]Stanford University, Stanford, CA

**Objective:** Recent literature has focused on the difference in outcomes of US patients with state funded insurance versus those with private insurance. The primary aim of this study was to examine patient insurance status and PROMIS Physical Function (PF), Pain Interference (PI), and Depression (DE) scores collected in the orthopaedic outpatient setting.
**Methods:** In 2015, the Washington University department of orthopaedic surgery implemented a new check-in process for patients that included the collection of the PROMIS Physical Function, Pain Interference, and Depression CATs via tablet. During an 11-week pilot phase at a single outpatient clinic location, data was collected from 6204 patient visits. The study team received IRB approval to examine patient reported PROMIS PF, PI, and DE scores along with basic demographic data, including insurance type. Descriptive statistics were performed, and a one-way ANOVA with Games-Howell post-hoc analysis was completed.
**Results**: Analyses determined that there were multiple statistically significant differences in PROMIS PF, PI, and DE domains between the insurance types. Of note, state funded Medicaid patients had significantly higher PI ($\bar{x}$=66.1) and DE ($\bar{x}$=53.7) scores as well as significantly lower PF ($\bar{x}$=36.7). Table 1 demonstrates the significant differences between Medicaid and other insurances with the Games-Howell post-hoc comparisons.
**Conclusions:** In a preliminary analysis of patient visits to a single outpatient orthopaedic clinic, state funded Medicaid patients reported significantly higher pain interference and depression and had lower physical function than the visits with other insurance types. This study was limited by unequal group sizes, and the research team plans to continue the investigation to include the scores of patients being seen at a clinic location dedicated to the care of state funded and uninsured patients.

## 58. PROMIS Physical Function, Pain Interference, and Depression: Assessing Outcomes in an Orthopaedic Trauma Setting

Amanda Spraggs-Hughes, MA[1], William M. Ricci, MD[1], Michael J. Gardner, MD[2], Christopher M. McAndrew, MD, MSc[1]

**Presenting Author: Christopher M. McAndrew, MD;** mcandrewc@wudosis.wustl.edu

**Institutional Affiliations:** [1]Washington University School of Medicine, St. Louis, MO**,** [2]Stanford University, Stanford, CA

**Objective:** The link between physical function, pain, and depression after traumatic orthopaedic injury has been well documented. The primary aim of this study was to compare PROMIS domains of Physical Function, Pain Interference, and Depression of patients in the orthopaedic trauma outpatient setting to the general population. The secondary aim was to assess relationships between each of the domains.
**Methods:** Beginning in June of 2015, all insured patients presenting to the orthopedic attending faculty outpatient clinics associated with a single Level-1 trauma center were asked to complete PROMIS CATs for Physical Function, Pain Interference, and Depression. After an 11-week pilot period, IRB approval was obtained to retrospectively collect the PROMIS CAT scores and basic demographic data. Descriptive statistics were analyzed and Pearson correlation coefficients were utilized to assess relationships between the individual PROMIS domains and patient demographics.
**Results:** The mean scores of 810 patient visits for each PROMIS domain in this cohort are reported in Table 1. In this sample, Physical Function ($\bar{x}$=34.9) was 1.5 SD below the general population, while Pain Interference ($\bar{x}$=60.4) was one SD above. Depression ($\bar{x}$=48.5) across the sample was consistent with the normalized population score of 50. Significant negative correlations were found between Pain Interference and Physical Function ($r$=-.445), Depression and Physical Function ($r$=-.266), and age and Physical Function ($r$=-.265), and a positive correlation was found between Pain Interference and Depression ($r$=.359) (Table 2).
**Conclusions:** In this diverse cohort of insured orthopaedic trauma patients the mean scores in Physical Function were lower than population norms while Pain Interference scores were higher. The mean Depression score was just below that of the general population, which is inconsistent with previous publications. The findings are generally consistent with previous reports, with an unexpected difference regarding Depression.

## 59. Using PROsetta Stone to Translate PROMIS Depression Scores for Meaningful Use in Orthopaedic Trauma

Amanda Spraggs-Hughes, MA[1], William M. Ricci, MD[1], Michael J. Gardner, MD[2], Christopher M. McAndrew, MD, MSc[1]

**Presenting Author: Amanda Spraggs-Hughes, MA**; amandas@wustl.edu

**Institutional Affiliations:** [1]Washington University School of Medicine, St. Louis, MO**,** [2]Stanford University, Stanford, CA

**Objective:** The primary aim of this study was to translate PROMIS Depression scores collected in an orthopaedic trauma outpatient clinic to Patient Health Questionnaire-9 (PHQ-9) depression severity levels utilizing the PROsetta Stone.
**Methods:** In 2015, the orthopaedic surgery department of a Level 1 trauma center implemented a new check-in process for patients that included the collection of the PROMIS CAT Depression assessments. The PROsetta Stone cross-walk was utilized to stratify the PROMIS scores by PHQ-9 depression severity levels. Descriptive statistics were used to analyze and report on the PROMIS Depression scores and translated PHQ-9 severity levels for the patients seen for injuries by members of the Orthopaedic Trauma Service (n=810 patient visits).
**Results:** Figure 1 demonstrates the breakdown of translated PROMIS Depression to PHQ-9 severity levels for this sample. The PHQ-9 proposed treatment actions recommend initiating depression treatment for PROMIS Depression scores of ≥59.9 including immediate initiation of pharmacotherapy intervention for PROMIS ≥65.8. Based on these guidelines, 13.7% of patient visits should have resulted in recommendation for treatment of depression and 4.2% meet criteria for immediate initiation of pharmacotherapy.
**Conclusions:** Translation of PROMIS Depression scores into the depression severity levels of the PHQ-9 allowed determination of patients in need of further evaluation and treatment for depressive symptoms. While the overall mean PROMIS score of this cohort was near the population mean, data from nearly 14% of our patient visits met PHQ-9 criteria for initiation of depression treatment or beginning pharmacotherapy. By utilizing the severity stratification of the commonly used PHQ-9, PROMIS Depression scores can be interpreted for appropriate and timely treatment of depression in orthopedic trauma patients. Future directions could include the creation of Collaborative Care Models to aid in the treatment of depression in orthopaedic trauma patients.

### 61. Sleep Disturbance Among Patients with Implanted Cardiac Device, The Substudy of PRO-CARDIA-POL Trial – Protocol Design.

Andrzej Cacko MD PhD, Emanuel Tataj MScEng., Agnieszka Żukowska M.S., Joanna Michalik MD, Wojciech Glinkowski MD PhD

**Presenting Author: Andrzej Cacko MD PhD**; andrzej.cacko@gmail.com

**Institutional Affiliation:** Department of Medical Informatics and Telemedicine, Medical University of Warsaw, Warsaw, Poland

**Objective:** The primary causes of implantation of the cardiac device (pacemaker or implantable cardiac defibrillator) are arrhythmia or significant heart failure. The risks of depression and anxiety among patients with implanted cardiac device are very high. Sleep disturbance may affect the quality of life and influence the depression and anxiety. The real burden of sleep disturbance among patients with implanted cardiac device is not well recognized. The aim of the study is to assess the burden of sleep disturbances and its effect on quality of life. The adherence to an electronic method of data collection among Polish patients' with the implantable cardiac device is the anticipated secondary outcome.
**Methods:** FACITtrans standardized methodology was used to translate Sleep Disturbance and Sleep-Related Impairment Item Banks. Items were introduced in the group of patients with implanted CARDIAc device in POLand (PRO-CARDIA-POL) trial. Patients' recruitment is based on the Academic Outpatient Cardiology Center in Warsaw (Poland). The presented study was designed to analyze the presence of sleep disturbances and its influence on patients' quality of life. All patients included in the study fill out the Sleep Disturbance and Sleep-Related Impairment questionnaires during outpatient follow-up visit electronically using tablets. Patients are obliged to fill the questionnaires again at home, either traditional (paper-and-pencil) or electronic (website). However, patients are asked about the reason for their choices.
**Results:** After this preliminary study, the final study group will consist of approximately five hundred patients that will be recruited from June to September 2016. The study will assess the presence of sleep disturbance and its influence on patients' quality of life. The significant difference in questionnaires scores responded at the Clinic or home is anticipated. Individual scores will be compared for every patient. We expect that attitudes towards electronic surveys will represent the overall preparedness of cardiac patients for online testing in Poland.
**Conclusions:** The study presents the introductory phase of PROMIS instruments translated into Polish applied to the group of patients to with implanted cardiac device. The study is expected to show several health status issues and the adherence to the electronic questionnaires in the PRO-CARDIA-POL trial.

### 62. Validity and Reliability of PROMIS® in Systemic Lupus Erythematosus (SLE)

Shanthini Kasturi, MD[1], Jayme C. Burket, PhD[1], Jessica Berman[1], MD, Kyriakos A. Kirou, MD[1], Alana B. Levine, MD[1], Lisa R. Sammaritano, MD[1], and Lisa A. Mandl, MD, MPH[1]

**Presenting Author: Shanthini Kasturi, MD;** kasturis@hss.edu

**Institutional Affiliation:** [1]Hospital for Special Surgery, Weill Cornell Medicine, New York, NY

**Objective:** a) To assess the validity of PROMIS computerized adaptive tests (CATs) in SLE outpatients by correlating PROMIS CATs with legacy patient reported outcome (PRO) measures, SLE disease activity, and organ damage; b) to assess the retest reliability of PROMIS CATs in SLE outpatients.
**Methods:** Adult SLE patients were recruited from a SLE Center of Excellence. Subjects completed the SF-36, LupusQoL-US, and selected PROMIS CATs at baseline and again within one week. SLE disease activity and damage were measured with the physician-derived SELENA-SLEDAI and SLICC-ACR damage index. PROMIS domains were compared with a) similar domains in the SF-36 and Lupus QoL-US; b) SLE disease activity; and c) organ damage using Spearman's correlation coefficients (*r*). Retest reliability was evaluated among subjects reporting stable SLE activity at two assessments one week apart using intraclass correlation coefficients (ICC).
**Results:** 204 patients completed PROMIS CATs and legacy instruments, with 162 (79%) completing a retest within one week. PROMIS CATs showed favorable performance characteristics compared to legacy instruments, demonstrating normal distributions and minimal floor and ceiling effects. PROMIS CATs correlated moderately to strongly with similar domains in the SF-36 and Lupus QoL ($r$ =0.49 to 0.83, $p < 0.0001$). However, correlations between PROMIS CATs and SELENA-SLEDAI and SLICC-ACR damage index were generally weak and statistically insignificant. PROMIS CATs retest ICCs were excellent ranging from 0.72 to 0.88.
**Conclusions:** To our knowledge, these data are the first to show that PROMIS CATs are valid and reliable for many SLE relevant domains. Importantly, PROMIS scores did not correlate well with physician-derived measures, underscoring the principle that PROs measure unique constructs. Valid PRO measures are particularly important in the management of SLE, a chronic disease with myriad clinical manifestations. Further studies are needed to evaluate the role of PROMIS in optimizing longitudinal disease management in SLE.

### 63. The Effectiveness of Physical Therapy in Low Back Pain Evaluated Using PROMIS Short Forms

A. Żukowska[1,2,3,4], Wojciech Glinkowski, MD, PhD[1,2,3,4,5]

**Presenting Author: Wojciech Glinkowski, MD, PhD;** w.glinkowski@gmail.com

**Institutional Affiliations**: [1]Department of Medical Informatics and Telemedicine, Medical University of Warsaw, Warsaw, Poland, [2] Department of Orthopaedics and Traumatology of Locomotor System, Baby Jesus Clinical Hospital, Warsaw, Poland, [3]Center of Excellence "TeleOrto", Warsaw, Poland4, [4]Gabinet Lekarski, Warsaw, Poland, [5]Polish Telemedicine Society, Warsaw, Poland

**Objective:** Evaluation of the effectiveness of rehabilitation of low back pain utilizing PROMIS Physical Function SF v1.2 10a and PROMIS Pain Interference SF v1.1 6b.
**Methods:** Study group consisted of 29 patients (17 women, 12 men). All subjects have been diagnosed with lumbar spondyloarthritis and suffered from low back pain (average pain score VAS=6). The average age of the group was 62.5 years. They participated in 3-weeks physiotherapy program that consisted of electrotherapy, magnetic therapy, cryotherapy, manual therapy and individually tailored physical exercises. Each subject before and after physiotherapy program filled out pencil-and-paper versions of PROMIS Physical Function SF v1.2 and PROMIS Pain Interference SF v1.1.
**Results:**
The back pain had the significant impact on the daily lives of subjects before the physical therapy program. Patients noted limited physical functions prior the therapy. Daily activities of patients were challenging because of the pain and mobility limitations.
After physical therapy program, it has been observed that the physical functions of patients have improved. Pain no longer interfered with everyday activities, and it was easier to complete daily tasks.
**Conclusions:** PROMIS Short forms can be successfully used in the evaluation of the musculoskeletal physical therapy efficiency. The results are an individual assessment of health from the perspective of the patient and can be used to assess the effectiveness of the applied treatment protocol.

### 64. The Impact Assessment of Hip Osteoarthritis on Daily Activities

Glinkowski W.[1,2,3,4], Żukowska A.[2,3,4], Dymitrowicz M. [2]

**Presenting Author: Wojciech Glinkowski, MD, PhD;** w.glinkowski@gmail.com

**Institutional Affiliations:** [1]Department of Medical Informatics and Telemedicine, Medical University of Warsaw, Warsaw, Poland, [2]Department of Orthopaedics and Traumatology of Locomotor System, Baby Jesus Clinical Hospital, Warsaw, Poland, [3]Center of Excellence "TeleOrto", Warsaw, Poland, [4]Polish Telemedicine Society, Warsaw, Poland

**Objective:** Assessment of the impact of hip osteoarthritis on the daily activities utilizing PROMIS Physical Function SF v1.2 10a and PROMIS Pain Interference SF v1.1 6b.
**Methods:** Study group consisted of 60 people (40 women, 20 men). All subjects have been diagnosed with hip osteoarthritis and suffered from pain (average VAS=5). The average age of the group was 65.5 years (women average age 64.8 years, men average age 66.7 years). All subjects were qualified for hip replacement surgery.
The control group consisted of 40 people (26 women, 16 men). All subjects were healthy and showed no signs of hip osteoarthritis (average VAS=0). The average age of the group was 33.4 years (women average age 34.8 years, men average age 30.8 years).

Each subject filled out pencil-and-paper versions of PROMIS Physical Function SF10a v1.2 and PROMIS Pain Interference SF6b v1.1. All filled out PROMIS Short Forms were scored utilizing Assessment Centre (https://www.assessmentcenter.net/).
**Results:** Study group showed worse than average population T-scores (50). PROMIS Physical Function SF average T-score was 34.9, and PROMIS Pain Interference average T-score was 65.1.
Control group T-scores - PROMIS Physical Function SF average T-score was 56 and PROMIS Pain Interference average T-score was 41.
**Conclusions:** Study group showed worse T-scores than average population and control group. Hip osteoarthritis has a big impact on daily activities due to pain and limited physical function. PROMIS Short forms can be successfully used in the assessment of the impact of hip osteoarthritis on daily activities. The results are an individual assessment of health from the perspective of the patient and can be used to assess the disease burden.

### 65. Content Debriefing Selected PROMIS Short Forms in Patients with Rheumatoid Arthritis

Clifton O. Bingham III[1], Ana Kristina Gutierrez[1], Alessandra Butanis[1], Vivian Bykerk[1], Jeffrey Curtis, MD3, Susan J. Bartlett, PhD[4]

**Presenting Author: Clifton Bingham**, **MD**;
cbingha2@jhmi.edu

[1]Johns Hopkins University, [2]Hospital for Special Surgery, [3]University of Alabama at Birmingham, [4]McGill University

**Objective**: Rheumatoid arthritis is a chronic, painful and debilitating disease with high levels of symptoms and impacts. We previously evaluated performance of selected PROMIS item banks on key symptoms and impacts in people with RA. Here, we report results of cognitive debriefing of selected PROMIS short forms (SFs) in people with RA.
**Methods**: Participants with RA were recruited from 3 US academic rheumatology centers. Phone interviews were used to cognitively debrief participants using a "talk through" format of the following SFs: Physical Function (PF 20a), Pain Interference (PI; 8a), Fatigue (FAT; 7a, 8a), Participation in Social Activities and Roles (PSRA 8a). Content relevance, anchors, response options, and perceived relevance of the items/responses were probed. Conversations were recorded and transcribed.
**Results**: Thirty-two participants were mostly female (72%) with a mean (SD) age of 54 (13) years and well-established disease (13 [10] yrs.). Participants were racially (66% white, 19% black, 13% native American, 9% Asian, 13% multi-race) and geographically diverse (28% South, 41% East, 34% mid-Atlantic; 28% rural). Participants reported higher PI (57.9 [10.6]) and FAT-7a (53.2 [9.9]), FAT-8a (57.9 [10.3]) and worse PF (42.2[11.0]) and PSRA (46.7[10.9]) than the general US population. Almost all rated items as very/somewhat relevant (PF 87-100%; PI 87-100%; PSRA 84-100%) and question content/response options as very/somewhat easily understood (PF 91%, PI 90%, PSRA 91%). Ratings were similar for fatigue except one item ("How often were you too tired to take a bath or shower") where 25% rated it "not at all relevant." When selecting responses for pain and fatigue, some focused only on intensity while others considered impacts on function/activities.
**Conclusions:** People with RA rated PROMIS Physical Function, Pain Interference, Fatigue, and Participation in Social Roles and Activities SF items as relevant to their experience, supporting their use in research and clinical management.

### 66. Floor Effect of Depression on Patients with Orthopedic Conditions

[1]AZ Dardas, BA [1]J. Guattery, MS, [1]M. Kelly, MD, [1]A. Chamberlain, MD, MSc, [1]RP Calfee, MD MSc

**Presenting Author: Ryan Calfee, MD, MSc;**
calfeer@wudosis.wustl.edu

**Institutional Affiliation:** [1]Department of Orthopedic Surgery, Washington University School of Medicine

**Objective:** PROMIS CAT scores have been normalized according to a standard population distribution. We sought to determine whether patients seeking orthopedic care present with a normal distribution of depression, physical function, and pain interference scores.
**Methods:** This cross-sectional evaluation analyzed 14,977 consecutive initial outpatient clinic visits of adult patients, ages 18 or older, presenting to a tertiary orthopedic practice in the United States from 10/2015-4/2016. All patients completed electronic PROMIS Depression, Physical Function, and Pain Interference CATs during registration as routine clinical care. Univariate descriptive analyses, Kolmogorov-Smirnov tests, and histograms statistically and graphically explored these PROMIS data.
**Results:** PROMIS depression scores ranged from 34.2 to 84.4 with a median of 48.2. The histogram indicated a bimodal distribution with a dominant peak at 34.2 and a secondary peak at 48 (Figure1). Twenty percent of all patients scored 34.2. A one-sample Kolmogorov-Smirnov test rejected the null hypothesis of a normal distribution ($p<0.001$). In these same patients PROMIS physical function CAT scores averaged 41 (range: 15.4 to 73.3) (Figure2). PROMIS pain interference scores ranged from 38.7 to 83.8 with a mean of 61 (Figure3). The histogram indicated a unimodal, normal-appearing distribution for both PROMIS function and pain scores.

**Conclusions:** In a specialty orthopedic practice, patients with PROMIS depression scores present in a bimodal distribution with a substantial floor effect. As PROMIS scores for function and pain interference were distributed as expected in this population, it is likely that the unexpectedly low scores on the depression CAT indicates patients' reluctance to admit depressive symptoms for risk of stigmatization. PROMIS physical function and pain interference scores were distributed normally but as expected for patients presenting with musculoskeletal conditions indicated average function that was 1 standard deviation below the general population and pain that was 1 standard deviation greater than the general population.

## 67. Variability of PROMIS Domains across Orthopedic Conditions

[1]AZ Dardas, BA, [1]M. Kelly, [1]A. Chamberlain, MD, MSc, [1]J. Guattery, MS, [1]RP Calfee, MD MSc

**Presenting Author: Agnes Dardas, BA;**
azdardas@go.wustl.edu

**Institutional Affiliation:** Department of Orthopedic Surgery, Washington University School of Medicine

**Objective:** To determine whether PROMIS Depression, Physical Function, and Pain Interference scores varied among patients presenting for orthopaedic care for either upper extremity, lower extremity, or spine condition.
**Methods:** This cross-sectional evaluation analyzed 8,532 consecutive initial outpatient clinic visits of adult patients, ages 18 or older, presenting to a tertiary orthopaedic practice in the United States for an upper extremity, lower extremity, or spine condition from 10/2015-4/2016. All patients prospectively completed electronic PROMIS Depression, Physical Function, and Pain Interference modules as routine clinical care. Kruskal-Wallis Chi-square analyses with post-hoc comparisons using the Mann-Whitney U test with a Bonferroni alpha-correction of 0.017, statistically assessed each CAT's scores according to location of the musculoskeletal condition.
**Results:** Depression, Physical Function, and Pain Interference scores significantly differed between all 3 groups (Depression: $p<0.001$; Physical Function: $p<0.001$; Pain Interference: $p<0.001$). Post-hoc comparisons revealed differences between all 3 patient groups for each of the 3 domains. Patients with upper extremity conditions had the lowest depression (median 46.1) and pain interference scores (median 59.9) and the highest physical function scores (median: 44.9), whereas spine patients had the highest depression (median: 53.3) and pain interference scores (median: 63.5) and the lowest physical function scores (median: 36.9) (Table1). Assuming moderate effect size change (0.5) to be clinically relevant, patients with lower extremity conditions and spine conditions had worse physical function than patients with upper extremity conditions (EF=0.63, 0.80 respectively), and patients with spine conditions indicate greater depression than patients with upper or lower extremity conditions (EF=0.72, 0.51 respectively)
**Conclusions:** PROMIS Depression, Physical Function, and Pain Interference scores statistically significantly differed by location of orthopedic conditions. Further studies comparing these potential differences in health between these patient groups should include additional validated health measures to determine if these apparent health disparities generalize outside of PROMIS scores.

## 68. Patient-Reported Musculoskeletal Function: Comparing PROMIS Upper-Extremity to PROMIS General Physical Function Scores

RP Calfee, MD, MSc, J. Guattery, MS, AZ Dardas, BA, M. Kelly, MD, A. Chamberlain, MD, MSc,

**Presenting Author: Ryan Calfee, MD;**
calfeer@wudosis.wustl.edu

**Institutional Affiliation:** Department of Orthopedic Surgery, University School of Medicine

**Objective:** This study tested the null hypothesis that patients with symptomatic upper-extremity conditions would report comparable physical function on the PROMIS physical function CAT and the recently developed PROMIS upper-extremity CAT.
**Methods:** All adult patients presenting to four upper-extremity surgeons at a single tertiary center from 10/2015 to 4/2016 completed both the electronic PROMIS physical function and PROMIS upper-extremity CATs during initial outpatient appointments. These data were extracted from the electronic health record for this cross-sectional evaluation. Univariate descriptive analyses explored each module's scores in this patient cohort. Bivariate Pearson correlation analysis defined the directional relationship between the modules while two tailed paired t test was performed to examine for absolute differences between the module scores in each patient.
**Results:** 1130 consecutive patients presenting for specialty treatment of symptomatic upper-extremity conditions completed both PROMIS modules. PROMIS physical function scores indicated musculoskeletal function 0.5 standard deviations below population norm (44, SD 10). PROMIS upper-extremity scores in the same patients averaged nearly 1.5 standard deviations below the population norm (35, SD 10). PROMIS physical function scores were highly correlated with the upper-extremity scores ($r_p=0.69$, $P<0.001$) (Figure 1). The difference between PROMIS scores on physical function and upper-extremity function was statistically significant (mean

difference 9, 95% CI 8.6-9.5, P<0.001) and presumed to be clinically relevant with an effect size of 0.9 (Figure 2).
**Conclusions:** PROMIS upper-extremity function scores systematically indicate poorer function when compared to PROMIS physical function scores in patients with upper-extremity conditions. Although PROMIS physical function scores correlate strongly with PROMIS upper-extremity scores, general physical function scores minimize the impact of isolated upper extremity conditions. Longitudinal data are needed to determine the comparative responsiveness of PROMIS upper-extremity and PROMIS physical function modules during treatment for upper extremity conditions

### 69. An Upper Extremity Physical Functioning Item Bank

Aaron J. Kaat, PhD[1], Benjamin Schalet, PhD[1], Nan Rothrock, PhD[1], Mark Vrahas, MD[2], and Richard Gershon, PhD[1]

**Presenting Author**: **Aaron Kaat, PhD;**
aaron.kaat@northwestern.edu

**Institutional Affiliations:** [1]Department of Medical Social Sciences; Northwestern University, Chicago, IL, [2]Department of Orthopedics; Cedars-Sinai Medical Center, Los Angeles, CA

**Objective:** PROMIS Physical Functioning (PF) reflects multiple facets of PF, including upper extremity (UE), mobility, and axial (neck/back) functioning. However, when administered via a computerized adaptive test (CAT), UE items are rarely selected. Although there is a UE-CAT, it has several limitations. This study aimed to create a separate UE item bank from the PF v2.0 items.
**Methods:** Expert panelists reviewed the PF item bank and identified 49 "sufficiently UE" items for consideration. All 49 candidate items were then administered to a new sample (n=600) to evaluate unidimensionality, monotonicity, local dependence, and convergent validity with the Flexilevel Scale of Shoulder Function (FLEX-SF).
Problematic items were eliminated. Then all items were calibrated using the new sample and the PROMIS Wave 1 and PASTOR data (n=11,975). CAT simulations were conducted to examine item bank usage and identify items for a short form.
**Results:** Preliminary analysis suggests that three items should be removed for exhibiting poor item fit and/or local dependence. The remaining items were calibrated using a multiple group estimation centered on the USA general population (using the Wave 1 centering data). The new item bank reflects primarily impaired functioning. The additional data suggests that the UE item bank converges well with the FLEX-SF ($r$ = .69), and although it correlates with the generic PF bank ($r$ = .71), it provides unique information for individuals with UE limitations. A 7-item short form was developed for administration without CAT availability.
**Conclusions:** The UE-PF item bank and short form will greatly benefit PROMIS clinical and research teams. It is recommended in addition to or instead of the generic PF item bank for individuals with known or suspected UE limitations. The bank has adequate reliability and converges well with the FLEX-SF. Future studies should investigate its responsiveness to treatment.

### 70. Comparing IRT pattern Scoring and IRT Summed Score Conversion Tables for PROMIS Short Forms

Eisuke Segawa, PhD[1], Benjamin D. Schalet, PhD[1], David Cella, PhD[1]

**Presenting Author: Eisuke Segawa, PhD;**
eisuke.segawa@northwestern.edu

**Institutional Affiliations:** [1]Northwestern University, Chicago, IL

**Objective**: We compared two scoring methods for PROMIS short forms: IRT pattern scoring (Bock & Mislevy, 1982), and IRT summed-score to T-score conversion tables (Lord & Wingersky, 1984). The IRT summed-score conversion tables are found in PROMIS scoring guides, while pattern scoring is enabled in most online administration platforms and the Assessment Center (AC) scoring service. We considered how well the scoring tables worked in cases when participants did not answer all the items in the short form and scores are in the normal or the severe range. We introduce modified scoring tables – which account for 1, 2, 3, or 4 missing items – in order to eliminate the bias that may be found in the scoring tables when some items are missing.
**Methods:** We examined the 8-item short forms in seven-domain PROMIS profiles. First, the mean squared errors (MSE) of the two scoring approaches were computed using simulated data (10,000 subjects from normal with mean 50 and SD 10) in order to compare their accuracy. Secondly, simulated subjects with extreme anxiety (10,000 subjects with T-score 80) and 4-items missing were scored by three methods: the scoring table, modified scoring table, and AC (pattern scoring). We examined the bias of both tables, using the AC as the criterion. Missing items were imputed using extrapolation in the scoring table; the modified scoring tables do not require extrapolation.
**Results:** With the exception of sleep disturbance, the scoring table MSEs were only slightly (from 0.4% to 3.5%) larger than AC MSEs. For the extreme anxiety, the scoring table, modified scoring table, and AC means were 79.5, 77.7 and 77.8, respectively. Treating the AC as the "gold standard," the scoring table showed 1.7 (= 79.5 - 77.8) bias while the modified scoring table showed minimal bias.
**Conclusion:** Use of IRT pattern scoring, as implemented in the AC Scoring service and elsewhere in PROMIS online delivery is recommended because it provides accurate scoring without bias. We continue to recommend the standard IRT summed-score to T-score table when little or no data are missing; indeed, it was almost as accurate as AC in all domains except sleep. However, when substantial amounts of data are missing (e.g., 50% of the items) and scores are extreme, the bias of the scoring table should not

be neglected and a modified table is recommended.

### 71. Differences in PROMIS Responses in People with RA Recruited Online vs an Academic Medical Center

Clifton O. Bingham, MD[1] Anna Kristina Gutierrez, MD[1], Michelle Jones, BSc[1], W. Benjamin Nowell, PhD[2], Seth Ginsberg, BS[2], Jeffrey Curtis, MD, MPH[3], and Susan J. Bartlett, PhD[1,4]

**Presenting Author**: **Susan Bartlett, PhD; susan.bartlett@mcgill.ca**

**Institutional Affiliations:** [1]Johns Hopkins University, [2]Global Healthy Living Foundation, [3]University of Alabama at Birmingham, [4]McGill University

**Aims:** Patients are increasingly recruited online to provide insight regarding their lived experiences and preferences for treatment and services. We compared physical, social and emotional health of people with rheumatoid arthritis (RA) recruited from an online patient community with those seen in an academic arthritis clinic.
**Methods:** Participants were RA patients recruited from 2 sources: a single academic rheumatology practice, and www.creakyjoints.org, an online arthritis patient community. Online participants were invited through a notice on the website and were screened for probable RA using a modified version of the CSQ, a validated tool used in epidemiologic studies. Patients provided socio-demographic and RA characteristics and completed PROMIS Short Forms: Physical Function (PF), Pain Interference (PI), Fatigue, Participation, Depression, Anxiety, and Sleep Disturbance. These domains were previously identified as important to international RA patients.
**Results:** Compared to clinic patients (n=52), online participants (n=200) had more education (69% vs. 92% >High School), shorter disease duration (15 [11] vs. 10[10] yrs.) and were more likely to be disabled due to RA (8 [15%] vs. 63 [32%]), but did not differ significantly ($p<.05$) by age, sex, or minority status. Compared with the US population, both groups reported significantly lower PF and greater PI and fatigue; online participants reported higher levels of depression, anxiety, and sleep disturbance. Compared with clinic patients, online participants also reported significantly greater impairments in PF (42 [12]) vs. 34[6]), PI (56[10] vs. 65[7]), fatigue (55[14] vs. 66[8]), Participation (49[11] vs. 39[7]), depression (50[12] vs. 58[10]), anxiety (50[11] vs. 58[9]) and sleep disturbance (51[10] vs. 58[9])(all p's < .0001).
**Conclusions:** These results suggest that, while patients recruited through online arthritis communities may differ in some demographic features from those in an academic medical center, they experience significantly worse physical, social and emotional health.

### 72. Validation of PROMIS Emotional Distress Short Forms in a Cervical Cancer Survivor Population

Lari Wenzel, PhD, Kathryn Osann, PhD, Susie Hsieh, PhD, Edward Nelson, PhD

**Presenting Author** – **Lari Wenzel, PhD**; Lwenzel@uci.edu

**Institutional Affiliations**- University of California, Irvine

**Objective:** To evaluate the PROMIS Emotional Distress Depression and Anxiety Short Forms for assessing distress endpoints in a cervical cancer survivorship trial.
**Methods:** A 15 item questionnaire measuring emotional distress was developed by the PROMIS initiative, and used in a cervical cancer biobehavioral randomized trial. It was administered to patients (N=204) prior to randomization, four months post enrollment, and nine months post enrollment. The depression (8 items) and anxiety (7 items) short forms were evaluated in patients participating in this study over all three time points, for internal consistency, construct validity, and responsiveness to change over time. Data were collapsed between arms for analyses of internal consistency and construct validity. To assess responsiveness to change over time, data were analyzed by treatment arm which included either psychosocial telephone counseling (PTC) or usual care (UC). The sensitivity of the short forms to change over time, including responsiveness to PTC, was examined with paired t-test for the change in depression and anxiety scores from baseline to the 4 month assessment among the patients who completed both baseline and 4 month assessment in both PTC and control arms. The sensitivity to change between study arms was examined with the two-sample t-test by comparing the changes in short form scores from baseline to the four month assessment, between the two arms.
**Results:** Internal consistency coefficients were ≥0.95, at baseline, 4 months, and 9 months respectively for depression and anxiety. The average inter-item correlation was 0.72 at the 4 month assessment for both depression and anxiety. The depression short form T-score was correlated with legacy scales ranging from 0.35-0.70, and the anxiety short form ranging from 0.29-0.64. Strongest correlations were with the Brief Symptom Inventory (BSI) and Perceived Stress Scale (PSS) for the Depression T-score (0.70 and 0.67) and 0.62 and 0.64 for the Anxiety T-Score. With respect to sensitivity to change over time at 4 months, in comparison with UC, PTC patients reported significantly greater improvement in depression (3.13 for PTC vs. 0.59 for UC; $p=0.014$). PTC patients also showed greater improvement in anxiety but the difference did not reach statistical significance (2.97 for PTC vs 0.89 for UC; $p=0.068$). Change over time at 4 months was also significantly greater in PTC compared to UC for the BSI Depression Standard Score ($p=0.041$) and non-significantly greater for the BSI Anxiety Standard Score ($p=0.103$).

Correlations for change over time between the PROMIS measures and BSI Standard scores were 0.41 for both depression and anxiety.
**Conclusion:** The PROMIS depression and anxiety short forms reliably and validly assess cervical cancer-specific emotional distress, they capture salient features of distress in this population, and they perform as well or better than legacy measures in this gynecologic cancer survivor population.

### 73. PROMIS-29 Health Profile and Related Short Forms Capture the Experiences of People Living with Rheumatoid Arthritis

Susan J. Bartlett, PhD[1, 2], Anna Kristina Gutierrez[2], Michelle Jones, BSc[2], Clifton O. Bingham, MD[2]

**Presenting Author: Susan J. Bartlett, PhD;** susan.bartlett@mcgill.ca

**Institutional Affiliations:** [1]McGill University and [2]Johns Hopkins University

**Aims:** PROMIS short forms (SFs) and the PROMIS-29 Profile were developed to assess symptoms across chronic diseases. We explored their relevance and importance in people with RA, and whether ratings reflected global vs. disease-specific symptoms.
**Methods:** Participants were RA patients in an academic rheumatology practice. Socio-demographic and clinical characteristics along with SFs for Physical Function (PF; 20a) Pain Interference (PI; 8a), Fatigue (7a, 8a), Ability to Participate in Social Roles and Activities (PSRA; 8a), Depression (8a), and the PROMIS-29 Profile. Participants rated the relevance and importance of items/domains to understanding their lived experience, and whether responses reflected their overall state or that attributable to RA (e.g., “How much of your fatigue is due to your RA?”).
**Results:** Participants were 52 adults who were mostly female (87%), white (89%), and well- educated (69% had some college), with a mean (SD) age of 53 (14) yrs. and well-established disease (15 [11] yrs.). Compared with the US population, participants reported significantly worse (>.4 SD) PF, PI, and Fatigue. Scores were similar among SF assessing the same domain, with the largest differences evident for PF. Most (>75%) considered these domains “very important” to understanding their experience; the majority (>88%) rated items “moderately” to “completely” reflecting their experiences. About half considered both their overall status and the specific impact of RA when rating their PF (56%), PI (52%) and PSRA (48%). Almost all (>88%) stated they would not have responded differently if asked to rate their symptom only in relation to their RA.
**Conclusions:** These results suggest that people with RA view PROMIS SFs as relevant and important to understanding their experiences. PROMIS-29 offers additional information that complements traditional measures routinely collected in RA care and clinical trials. Researchers may wish to use SFs when greater precision is desired.

### 74. Remote Collection of Patient Reported Outcomes Following Carpal Tunnel Release: A Randomized Trial of Telephone, Mail and Email

Andrew Tyser, MD[1], Justin J. Schwartzenberger, MD[1], Rebecca Johnson, MPH[1], Angela P. Presson[1], Ragheed Al-Dulaimi[1], Adam Lyle[1], Andrew Lyall O’Farrell[1]

**Presenting Author: Rebecca Johnson, MPH**; r.johnson@hsc.utah.edu

**Institutional Affiliation:** [1]University of Utah, Salt Lake City, UT

**Objective:** Obtaining remote patient reported outcomes (PROs) is limited by low patient response rates and resource-intensive collection methods. We hypothesized that an email-delivered web-based data collection tool would outperform the traditional methods of telephone and standard mail for collecting long-term Levine-Katz scores following carpal tunnel release.
**Methods:** We conducted a randomized trial of 969 patients who underwent carpal tunnel release at a tertiary medical center within the past 5 years. Participants were randomized to the PRO collection methods of mail, telephone, and email. The primary outcome was response rate at one year after surgery. Secondary analyses included data completeness, the effect of time from surgery and patient demographics on response rates, mode effects, and patient modality preference.
**Results:** At one year out from surgery the response rate was 64% for telephone, and 42% for both mail and email ($p=0.005$). 99% of telephone surveys were complete compared with 88% and 83% for mail and email, respectively ($p<0.001$). There was no significant difference in the response rate as a function of time since surgery (1-5 years) or in the Levine-Katz score among the modalities. Overall, respondents were more likely to be older, retired, and utilizers of our electronic medical record patient portal (all $p<0.05$). Response rates were lowest among young patients, disabled patients and full-time workers (all $p<0.05$).
**Conclusions:** A higher response rate and increased survey completeness was achieved by telephone contact methods compared to standard mailings or web-based methods for patient reported outcome collection after carpal tunnel release one to five years after surgery. A web-based method demonstrated response rates equivalent to standard mail, was the most preferred modality, and offered logistical advantages such as automation and immediate integration with outcome databases.

## 75. Treatments Should Target Physical Function, Mood, and Fatigue to Enhance Participation in People with Rheumatoid Arthritis

Susan J. Bartlett, PhD[1,2], Michelle Jones, BSc[2], Clifton O. Bingham, MD[2]

**Presenting Author: Clifton Bingham, MD;** cbingha2@jhmi.edu

**Institutional Affiliations:** [1]McGill University and [2]Johns Hopkins University

**Objective:** The goal of rheumatoid arthritis (RA) treatment is to maximize health-related quality of life (HRQL) by controlling symptoms and joint damage and improving function to help sustain participation in social and life activities. While the impact of RA on physical function has been extensively studied, relatively little is known about how physical function and other RA symptoms impact participation. We hypothesized that after controlling for RA disease activity, low physical function and mood, and higher pain and fatigue would be associated with reduced participation.
**Methods:** RA patients enrolled in an observational study at an academic center completed PROMIS measures assessing physical function, fatigue, depression, and participation; pain was assessed using a 100 mm VAS. Variance inflation factors were examined to evaluate collinearity among variables, and multiple regression was used to evaluate the relationship among variables using SPSS V23.0.
**Results:** The 171 participants were mostly female (82%) and white (83%) with mean (SD) age of 56 (13) years; 24% had ≤ high school, 29% had RA ≤ 5 years with 13% ≤ 2 years, and 22% were disabled. Mean (SD) age was 55 (13) and swollen joints were 2.3 (3.4); most had well controlled disease (CDAI remission n=56; 32% or Low Disease Activity n=67; 38%); 39 (22%) had moderate and 14 (8%) had high disease activity.
Mean (SD) PROMIS participation was 50.2 (9.1); 31% scored ≤ 45. Mean (SD) physical function was 43.5 (9.0), fatigue 53.9 (10.1), depression 49.1 (8.9), and pain 31/100 (28). After controlling age, physical function (B=.47), depression (B=-.28), and fatigue (B=-.16) were significant ($p<.02$) independent predictors of participation ($F=39.4$, 6 and 164 df; $p = 000$; adjusted $r2=0.58$). Contrary to our hypothesis, after controlling for other factors, neither disease activity nor pain was not associated with participation (p's≥.74).
**Conclusions:** Our results suggest that suboptimal participation is common in people with RA. Physical function is the most robust predictor of participation, followed by fatigue and low mood. RA treatments and interventions that also improve fatigue and mood in addition to physical function may increase participation in social and life situations, restoring a sense of normalcy and ultimately improving HRQL. Funding PCORI IP2-PI0000737 and SC14-1402-10818

## 76. Obtaining the Best Possible Estimates of Health-Related Quality of Life in Patients with COPD Using Computerized Adaptive Testing Based on Three PROMIS® Domains

Muirne Paap, PhD[1,2], Karel Kroeze[2], Caroline Terwee, PhD[3], Job van der Palen, PhD[2, 4]

**Presenting Author: Muirne C.S. Paap, PhD**; m.c.s.paap@cemo.uio.no

**Institutional Affiliations:** [1]CEMO, University of Oslo, Oslo, Norway: [2]Department of Research Methodology, Measurement, and Data-Analysis, Behavioural, Management and Social sciences, University of Twente, Enschede, The Netherlands, [3]Department of Epidemiology and Biostatistics and EMGO+Institute for Health and Care Research, VU University Medical Centre, Amsterdam, The Netherlands, [4]Medical School Twente, Medisch Spectrum Twente, Enschede

**Objectives**: 1) To develop a multidimensional item bank to measure health-related quality of life (HRQoL) in patients with Chronic Obstructive Pulmonary Disease (COPD); 2) to investigate the added value of taking into account multidimensionality when developing a HRQoL-CAT; 3) to compare the performance of fixed versus variable length CAT in this setting.
**Methods**: Based on literature review, and interviews with patients and healthcare professionals, four important domains were identified: three generic COPD-relevant domains (Fatigue, Physical function, and Ability to participate in social roles and activities) from the Patient Reported Outcomes Measurement Information System (PROMIS®), and a COPD-specific domain. The final item bank consisted of 148 PROMIS items and 46 COPD-specific items, and was calibrated using the multidimensional Graded Response Model. CAT simulations were run to compare fixed to variable-length CATs and multi- to unidimensional CATs. Outcome variables included bias, accuracy, and item exposure.
**Results**: The correlations between the domains were high: .76-.86. MCAT outperformed UCAT in terms of bias and accuracy; this difference was largest for short fixed-length CATs. The findings will be illustrated with plots at the conference.
**Conclusions**: Variable length CATs should be the preferred choice for PROMIS CATs; if relatively short fixed-length tests are used, MCAT should be favored over UCAT.

## 77. Tailored symptom reporting and PRO assessment using a home-based cancer symptom management system

Michael A. Kallen, PhD, MPH[1]; Niina Haas, MA[2;] DerShung Yang, PhD[2]

**Presenting Author**: **Niina Haas, MA**;
niina.haas@brightoutcome.com

Institutional Affiliations: [1]Northwestern University;
[2] BrightOutcome Inc.

**Objectives:** Cancer patients frequently receive treatment as outpatients and self-manage symptoms at home. Symptoms are prevalent and can severely impact HRQOL. Despite discharge information and counseling on treatment side-effects and symptom management, stress, low self-efficacy, or low health literacy often negates such help. We employed SymptomCareAnywhere (SCA) to support patient self-management at home.
**Methods**: Patients reported symptom absence, presence, and severity online or by phone via SCA. Sixteen symptoms were reportable; adaptive PROMIS measures quantified moderate-to-severe pain, fatigue, depression, and anxiety. Symptom status exceeding severity thresholds triggered alerts for follow-up. We computed frequency, percentage, and mean report type (no, mild, moderate-to-severe symptoms), contents (number of and specific symptoms), method (phone, internet), and burden (seconds to complete a report) at overall and patient levels.
**Results:** N=645 reports were submitted by N=59 mostly chemotherapy patients. While 19.4% indicated no symptom, 33.0% indicated mild symptom(s) and 47.6% indicated at least one moderate-to-severe symptom. Overall, patients reported N=1,518 symptoms, 43.7% of which were moderate-to-severe. 32.5% of the reports communicated one symptom, 25.0% communicated two symptoms, and 16.5% communicated three symptoms. Most patients (49.1%) used a phone for reporting; 44.1% used the internet. Internet vs. phone users did not significantly differ by age (internet-63.5 years, phone-60.0 years) or by gender (internet-61.5% female vs. phone-58.6% female). On average, patients required 86.8 seconds to complete a report; “mild” reports averaged 45.6 seconds, and “at least one moderate-to-severe” reports (involving PROMIS adaptive measures) averaged 105.1 seconds. The three most frequently reported moderate-to-severe symptoms were fatigue (19.1%), pain (15.2%), and numbness (14.8%). Anxiety and depression symptoms made up 4.7% and 3.5%, respectively, of moderate-to-severe symptom reports.
**Conclusions:** SCA functions as an effective and practical distance-based symptom communication tool, enabling frequent but as-needed symptom status reporting involving adaptive PROMIS measures according to real-time patient experience, supporting successful symptom management.

## 79. Descriptive, Psychometric, And Feasibility Summaries of PROMIS Physical Function Instruments for the TOIMIA database

Kaisa Kokko, MSc, PT[1], Heidi Anttila, PhD, PT[1], Heli Valkeinen, PhD, PT[1]

**Presenting Author: Kaisa Kokko**, **MSc;**
kaisa.kokko@thl.fi

**Institutional Affiliations:** [1]National Institute for Health and Welfare (THL); Welfare department; Ageing, Disability and Functioning Unit, Helsinki, Finland

**Objective:** The Finnish database of functioning instruments (TOIMIA, www.toimia.fi) is maintained by the National Institute for Health and Welfare. Currently, it includes evaluations and descriptions over 80 instruments and 20 recommendations. Health care professionals can find reliable information in one place which will harmonize practices. Valid and competent assessment of functioning helps planning and allocating actions and resources adequately. PROMIS Physical Function adult item bank 1.2 is being translated into Finnish. We aimed 1) to describe PROMIS measurement system in general and all the PROMIS Physical Function instruments, 2) to extract the psychometric properties and 3) to make a feasibility recommendation to be published in the TOIMIA database.
**Methods:** Descriptive data about PROMIS was collected from the HealthMeasures website and relevant studies. A literature search to identify studies of psychometric properties on all PROMIS adult Physical Function instruments (item bank, CAT and short forms) was conducted in 13 databases in March 2016. One reviewer screened the titles and abstracts for inclusion criteria, and selected the articles for data extraction after reading the full texts. Data about the validity, reliability, responsiveness and feasibility as categorized by the COSMIN taxonomy were extracted.
**Results:** A description of PROMIS measurement, a summary of the psychometric properties and a feasibility recommendation based on the psychometric data were prepared and sent for audition in all five TOIMIA expert groups. For the psychometric data extraction we identified 38 studies, and 27 studies were included in the study.
**Conclusion:** The data extraction provides evidence that the PROMIS Physical Function instruments can be recommended to use in normal population and different chronic conditions. The collected summaries in Finnish provide useful information for clinicians and researchers in Finland to make them familiar with PROMIS measurement system. Subsequently, data of other PROMIS instruments will be collected to the TOIMIA database.

## 80. Patient-Reported Outcomes from a Phase 3 Study of Baricitinib in Patients with Early Rheumatoid Arthritis

Michael Schiff, MD[1], Tsutomu Takeuchi, PhD[2], Carol Gaich, PharmD[3], Amy M. DeLozier, MPH[3], Douglas Schlichting, RN, MSc[3], Wen-Ling Kuo, PhD[3], Patrick Durez, MD[4], Tara Carmack, MSc[5], Ji Eon Won, MD[3], Roy Fleischmann, MD[6]

**Presenting Author**: **Esbjörn Larsson, MD, PhD**; larsson_esbjoern@lilly.com

**Institutional Affiliations:**
[1]University of Colorado School of Medicine, Denver, United States; [2]Keio University, Tokyo, Japan; [3]Eli Lilly and Company, Indianapolis, United States; [4]Université Catholique de Louvain, Brussels, Belgium; [5]Quintiles, Durham, United States; [6]Metroplex Clinical Research Center, Dallas, United States

**Objective:** Baricitinib (bari), an oral JAK1 and JAK2 inhibitor, was efficacious in a Ph 3 study (RA-BEGIN) in RA patients (pts) who had limited or no exposure to methotrexate (MTX) and who were naïve to other csDMARDs and bDMARDs1. The objective was to evaluate patient-reported outcomes (PROs) from RA-BEGIN.
**Methods:** Pts were randomized to MTX QW, bari 4 mg QD, or bari 4 mg QD+MTX QW. PROs (physical function [HAQ-DI]; Pt's Global Assessment of Disease Activity [PtGDA]; pt assessment of pain; fatigue [FACIT-F]; quality of life [SF-36 PCS; MCS]) and the Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA) questionnaire were collected on electronic tablets during study visits. Bari vs. MTX and bari+MTX vs. MTX were assessed with ANCOVA and logistic regression models.
**Results:** 584 pts were randomized. Mean baseline PROs for MTX, bari, and bari+MTX, respectively, were HAQ-DI:1.67, 1.64, 1.58; PtGDA:65.6, 65.0, 63.1; pt assessment of pain:65.2, 64.1, 62.6. Compared to MTX, bari and bari+MTX were superior in physical function, PtGDA, pain, and fatigue at Wks 24 and 52. Statistically significant improvements in all components of the WPAI-RA (absenteeism, presenteeism, work productivity loss and activity impairment) were seen in the bari and bari+MTX pts vs. MTX at Wk 24; statistically significant improvements were seen for work loss for bari+MTX vs. MTX at Wk 52 and for activity impairment for bari and bari+MTX vs MTX at Wk 52.
**Conclusions**: In this Ph 3 study of pts with early active RA, bari alone or with MTX was associated with significant improvements at 24 and 52 wks compared to MTX in most PROs.

**References:**
[1] Fleischmann et al. Arthritis Rheumatol 2015; 67(S10)

## 81. Internal Consistency of the Norwegian Translation of the PROMIS-57 Profile Instrument

Jan E. Nordvik, PhD[1]; Stein A. Rimehaug, PT[1], Ingvlid Grimstad, Cand.mag[1]; Torgeir Moberget, PhD[1, 2], Marie Berg, PhD[1]; Jennifer Moore, PT, DHS[1, 2]; Erik Resaland, MA[1]; Allen Heinemann, PhD[3]

**Presenting Author: Jan Egil Nordvik, PhD;** janegil.nordvik@sunnaas.no

**Institutional Affiliations:** [1]Regional Center of Knowledge Translation in Rehabilitation, Sunnaas Rehabilitation Hospital, Norway, [2]NORMENT, Oslo University Hospital, Norway, [3]Center for Rehabilitation Outcomes Research, Rehabilitation Institute of Chicago, USA

**Objective:** This project aimed to determine the internal consistency of a provisional version of the Norwegian translation of PROMIS-57 profile instrument applied in an inpatient rehabilitation population. PROMIS-57 includes seven subscales: Physical function (PF), Anxiety (A), Depression (D), Fatigue (FA), Sleep disorder (SL), Social function (SF) and Pain (PI). Each subscale consists of eight items, with the exception of the pain scale which include a ninth additional question. This instrument was chosen as a candidate measure for a planned minimum data set for the rehabilitation services in the health region of South-Eastern Norway, due to its relevance across physical conditions and diagnoses, widespread international use, availability of normative data, and relevance to the ICF model.
**Methods**: PROMIS-57 was translated to Norwegian and a provisional version was distributed to inpatients at 10 rehabilitation wards and centers, between October 2015 and July 2016. The patients filled in the questionnaire both at admission and discharge. Data were analyzed to determine internal consistency (Cronbach's α) of each subscale.
**Results:** Data were collected from 613 subjects undergoing inpatient rehabilitation with a primary diagnosis of a neurologic or musculoskeletal injury or disease that limited physical functioning. Cronbach's Alpha ranged from 0,913 to 0,964 at admission and 0,925 to 0,965 at discharge (see Table 1).
**Conclusion:** The Norwegian translation of the PROMIS-57 profile demonstrated excellent internal consistency across the subscales at admission and discharge for inpatient rehabilitation. These results indicate that the Norwegian translation of PROMIS-57 is an acceptable measure for the minimum data set.

## 82. The Clinical Spectrum of PROMIS Physical Function Scores over Time in Patients with Operative Lumbar Pathology

Stephen Pehler, MD[1], Yue Zhang, PhD[1], Ashley Neese, BS[1], Brandon D. Lawrence, MD[1], W. Ryan Spiker, MD[1], Mathieu D. Squires, BS[1], Darrel S. Brodke, MD[1]

**Presenting Author: Ashley Neese, BS**; ashley.neese@hsc.utah.edu

**Institutional Affiliations:** [1]University of Utah, Salt Lake City, UT

**Objective:** With a continually increasing focus on Value in today's healthcare economy, Patient Reported Outcome Measures (PROMs) provide critical information for assessing the effectiveness of care. PROMIS physical function (PF) domain scores have been shown to be valid and accurate, with low floor and ceiling effects in the spine patient population. The purpose of this study is to understand how PF CAT scores vary over time with treatment. The scope of this study analyzes scores of patients with lumbar pathology in the preoperative period through post-operative time points with treatment.
**Methods:** A retrospective review was conducted of patients' PRO scores at a university spine center, from October, 2013 to April, 2015. All patients had a primary complaint consistent with lumbar pathology and underwent surgery - lumbar decompression +/- fusion. Patients were excluded if they were younger than 18, or not able to complete the questionnaire.
**Results:** A total of 870 unique patients were identified, representing 5659 PROMIS PF CAT scores. A linear mixed effect model and linear mixed quantile regression model were fitted, respectively. The predicted PF CAT scores (± 95% confidence intervals) were mapped over time from the preoperative period to 12 months post op for all lumbar spine patients and separately for the fusion and decompression alone patient groups. Age-adjusted percentile rank of scores was also mapped over time for the same patient groups. Subgroup analysis comparing diabetic and non-diabetic patients showed no change in the age-adjusted percentile rank scores post-operatively.
**Conclusions:** To our knowledge, this is the first study to describe the overall trend of PF CAT scores over time following treatment, as well as the age-matched percentiles of PROMIS scores. This data can potentially assist in educating patients on their expected progress and improvements in function with treatment. This data also provides a unique perspective for patients to visualize their progress in the treatment.