PROMIS HEALTH ORGANIZATION

PHO INAUGURAL CONFERENCE

Philadelphia Marriott
Philadelphia, PA

MAY 15, 2015

ABSTRACT BOOK

Citation Example: Arnold, G., Baranowski, R., Lynn, L. (2015) Use of PROMIS short form to validate a new patient/caregiver questionnaire focusing on palliative care. Poster presented at the PHO Inaugural Conference, Philadelphia, PA., p. 11.

ORAL PRESENTATONS (listed alphabetically by presenting author)				
Presenting Author	**Title**	**SESSION #**	**ROOM #**	**PAGE #**
Susan J. Bartlett, PhD	PROMIS-29 en français: Coordinating Translation Efforts in Canada and France for a Universal Approach	3	302/303	11
Susan J. Bartlett, PhD	How Well do Generic Patient PROMIS Instruments Capture Health Status in Rheumatoid Arthritis?	1	302/303	12
Brandon D. Becker	Estimating Pediatric PROMIS Item Bank Item Parameters Representative of the US Population	1	304/305	13
Clifton Bingham, MD	Using PROMIS in PCOR: Lessons from PCORI's Pilot Projects	1	Liberty	15
Jason B. Colditz, MEd	Social Media Use and Anxiety as Measured by the PROMIS Anxiety Short Form	3	304/305	18
Karon Cook, PhD	Uncle Sam Wants PROMIS: Leveraging PROMIS to Build and Pilot the DoD's Pain Assessment Screening Tool and Registry (PASTOR)	3	Liberty	20
Heidi M Crane, MD MPH	Lessons from implementing and collecting PROs for HIV care and research: results from 10,000 patients	2	Liberty	22
Paul K Crane, MD MPH	Measurement precision and differential classification of at-risk alcohol use with AUDIT-C vs. PROMIS in HIV	3	304/305	23
Amy Eisenstein, PhD	Title: Standardizing and Personalizing Patient-Centered Rheumatoid Arthritis Treatment Targets	1	Liberty	24
Christopher B. Forrest, MD, PhD	Profile of subjective wellbeing among children in the US	2	302/303	26
Carol M. Greco, PhD	New Patient-Reported Measures of Contextual Factors in Healing	1	304/305	33

Presenting Author	Title	SESSION #	ROOM #	PAGE #
Ann L. Gruber-Baldini	Calibration and Validation of the PROMIS® Self-Efficacy for Managing Chronic Conditions Measures	1	304/305	35
Elizabeth A. Hahn, MA	Health Literacy and Patient-Reported Outcomes In Underserved English- and Spanish-Speaking Diabetes Patients	2	302/303	36
John D. Hartman, MA	Comparison of PROMIS-29 and EQ-5D-3L Quality-Adjusted Life Years	3	302/303	37
Roxanne E. Jensen, PHD	Evaluation of Responsiveness of 8 PROMIS Measures in a Community-Based Sample of Cancer Patients	2	304/305	40
Michael Kallen, PhD	Grooming a CAT: Customizing CAT Administration Parameters to Increase Response Efficiencies in Specific Research and Clinical Settings	2	Liberty	41
Irene Katzan MD	Predicting EQ-5D Utility Index from the PROMIS-10 Global Form – Ready for PrimeTime?	3	302/303	45
Dinesh Khanna, MD, MS	The PROMIS-29 in systemic sclerosis: associations with clinical characteristics	1	302/303	46
Sherri LaVella, PhD, MPH, MBA	Assessing Physical and Mental Health of Veterans Receiving VA Health Care Using a PROMIS Profile	3	304/305	49
Yvonne Lee, MD, MMSc	PROMIS Pain, Fatigue, Sleep, Fatigue, Anxiety and Depression in Patients with Active Rheumatoid Arthritis	3	Liberty	50
Constance A. Mara, PhD	Establishing Clinical Meaning and Severity Cut-points on Patient Reported Outcome Measures in Juvenile Idiopathic Arthritis	1	302/303	53
Ulrike Ravens-Sieberer, PhD, MPH	Translation of eight Pediatric PROMIS® item banks into Spanish and German	2	302/303	59
Dennis Revicki, PhD	Developing The Fibrodysplasia Ossificans Progressiva-Physical Function Questionnaire For Adults: Intelligent Test Design Using A Promis Item Bank	1	302/303	60

Presenting Author	Title	SESSION #	ROOM #	PAGE #
Ana Maria Rodriquez, PhD MSc PT	The Future of PRO Development: Using Technology to Advance Patient- Centered Measure Development Pt. 1 & 2	1	Liberty	62
Matthias Rose, MD, PhD	Simple web-based IRT score estimation for common metrics: http://www.common-metrics.org	2	Liberty	64
Nan Rothrock, PhD	Validation of PROMIS Mobility Computer Adaptive Test and Short Form in Patients with Orthopedic Trauma	2	304/305	64
Nan Rothrock, PhD	Validation of Patient-Reported OutcomesMeasurement Information System (PROMIS) Computer Adaptive Tests in Lumbar Stenosis Patients	2	304/305	65
Shannon K. Runge, MA	Choosing between the PROMIS Global and EQ-5D for Comparative Effectiveness Research: Are they really different?	3	302/303	66
Karen Schifferdecker, PhD, MPH	Creating Condition-Specific Patient Reported Outcome Measures	1	304/305	67
Priscilla M. Schulz, MSW, LCSW-C	PROMIS®29: A tool for improving mental healthcare and well-being of US Peace Corps Volunteers	3	304/305	68
George J. Stukenborg, PhD	Collection and Use of PROMIS measures for Palliative Care Cancer Patients	3	Liberty	73
George J. Stukenborg, PhD	PROMIS assessment for patients undergoing colorectal surgery in enhanced recovery protocol	2	304/305	74
Gemma Vilagut, MSc	The Spanish PROMIS® Depression measures are valid and responsive in a clinical sample of mental health patients	2	302/303	78
James H. Willig, MD, MSPH	“Basic PROMIS Panel” a proposal for a unified visual interface to summarize multiple PROMIS domains	2	Liberty	80
James H. Willig, MD, MSPH	PRO implementation in clinical practice: Failure as a steppingstone to success	3	Liberty	81

Poster Presentations (listed alphabetically by presenting author)		
Presenting Author	**Title**	**Page**
Dagmar Amtmann, PhD	Comparison Of The SF-36 and PROMIS Symptom Profiles Of People With Chronic Pain	10
Gerald K. Arnold, PhD, MPH	Use Of PROMIS Short Form To Validate A New Patient/Caregiver Questionnaire Focusing On Palliative Care	10
Heather Becker, PhD	Do PROMIS Cognitive Abilities Scores Among People With Multiple Sclerosis Improve Following A Cognitive Intervention?	14
Katherine B. Bevans, PhD	Qualitative Development Of Pediatric Oral Health Item Banks	14
Nina P. Bobowski, MPH	Sarcoma Survivors' Patient-Reported Late-Effects Of Their Treatment	16
Jayme C. Burket, PhD	PF, PI, And Mobility CAT Validation Against The FAOS For Common Foot And Ankle Pathologies	17
Jason B. Colditz, Med	Is Higher Social Media Use Associated With Reduced Social Isolation?	19
Helena Correia, BA	International Use Of PROMIS®: Multilingual Translation Efforts And Standards	21
Kathryn E. Flynn, PhD	Health-Related Quality Of Life In Couples Seeking Treatment For Infertility	25
Rob Fredericksen, Phd MPH	Development And Content Validation Of A PROMIS Sexual Risk Measure	26
Denise C. Fyffe, PhD	Assessing The Relationship Between Socio-demographics And The Interpretation Of Spanish PROMIS® Physical Function Items	28
Frances B Garfield, PhD	Using PROMIS® Instruments To Evaluate A Community Based Program To Empower Women	28
Aileen Gariepy, MD, MPH	Impact Of Unintended Pregnancy On Women's Quality Of Life	29
Aileen Gariepy, MD, MPH	Clinical Validation Of PROMIS Global Short Form In Pregnancy	30
Elizabeth Gilger, NP	Implementing A PROMIS To Improve The Quality Of Life In Pediatric Patients With CNS Tumors	31

Comparison of the SF-36 and PROMIS symptom profiles of people with chronic pain

Dagmar Amtmann, PhD, Rozanne Wilson, PhD, Alyssa Bamer, MPH, Rana Salem, MA

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

Institutional Affiliations: [all] University of Washington, Department of Rehabilitation Medicine

Objective: To compare the symptom profiles based on The Patient Reported Outcomes Measurement Information System 29 Profile v1 (P29) and the SF-36 Health Survey v2 (SF-36) in a sample of adults living with chronic pain.

Methods: In a cross-sectional study of pain, P29 and SF-36 were co-administered to a sample of people living with chronic pain. Eligibility criteria included presence of daily pain for more than three months and at least moderate self-reported pain intensity in the past week. For both instruments, a score of 50 (SD=10) represents the mean of representative samples of the U.S. general population (the norms), except for PROMIS satisfaction with participation in social roles, which is centered on the mean of a reference clinical sample. The SF-36 norms were weighed for age and gender based on the 1998 US census while the PROMIS norms were matched to the 2000 US census on gender, age, and race/ethnicity. Pearson's correlations and the magnitude of difference from the norms were used to examine the relationship between P29 and SF-36 scores.

Results: A total of 219 participants responded to the mailed survey (back pain N=67, osteoarthritis N=68, diabetic neuropathy N=84). Participants reported worse health than the norm across both instruments for all domains. On the P29 the differences from the norm ranged between 0.40 (depression) to 1.30 (pain interference) SD; SF-36 differences from the norm ranged from 0.53 (mental health) to 1.54 (bodily pain (BP)) SD. All correlations between SF-36 and P29 domain scores were in the expected direction. Correlations between similar domains, i.e., P29 physical function (PF) with SF-36 PF and P29 satisfaction with social roles with SF-36 social functioning, were 0.78 and 0.60 respectively. The correlation between SF-36 BP and (a) P29 pain interference was -0.71, (b) P29 pain intensity was -0.67 and (c) P29 PF was 0.60. SF- 36 vitality was correlated with P29 fatigue at - 0.76 and social roles at 0.56. P29 anxiety and depression were -0.80 and -0.84 correlated with the SF-36 mental health score. SF-36 emotional role functioning was strongly correlated with P29 anxiety and depression (-0.70 both). None of the correlations of SF-36 general health with P29 domains exceeded 0.44.

Conclusions: P29 and SF-36 profiles provide somewhat different pictures about samples measured. The difference between the norm scores on the SF-36 indicated slightly worse health than the differences on the P29 domains. Compared to the norm the SF-36 scores suggested that the sample was slightly sicker than did the P29 scores. The differences were small in magnitude, but were consistent across all domains. One explanation is that the PROMIS norms included relatively sicker people than the SF-36 norms. Advantages of P29 include more specific clinically relevant domains and inclusion of sleep disturbance. The inclusion of the PROMIS Global would provide similar composite scores. SF-36 advantages include ease of interpretation because on all domains a higher score is better, although on BP that is somewhat counter-intuitive. In addition, the SF-36 mental and physical composite scores may be informative for epidemiological studies.

Use of PROMIS Short Form to Validate a New Patient/Caregiver Questionnaire Focusing on Palliative Care

Gerald K. Arnold, PhD, MPH; Rebecca A. Baranowski, MEd, MS; Lorna A. Lynn, MD

Presenting Author: **Gerald K. Arnold, PhD, MPH;** garnold@abim.org

Institutional Affiliations: [all] American Board of Internal Medicine

Objective: This study was designed to validate a new patient/caregiver questionnaire (PCQ) developed for assessment of physicians treating chronically or seriously ill patients.
Methods: The Global Health Scale (GHS) of PROMIS® was administered as validation criterion for the PCQ scores. Physical and mental health scores were calibrated into two T-scale measures with means and standard deviations of 50 and 10, respectively. The scored PCQ scales were correlated (Spearman correlations) with GHS scores. Salient validity coefficients were correlations with absolute values exceeding $r > 0.1$. Domain structure was assessed in three steps: 1) Factor number was determined using a disjoint cluster method, repeated with 1000 bootstrap samples; 2) Maximum-likelihood factor analysis tested fit; and 3) confirmatory factor analysis estimated factor loadings.
Results: Factor analysis of 1009 responses yielded four domains: 1) physician communication skills, 2) Pain, 3) Dyspnea, and 4) Emotional distress. Salient correlations with patients' overall rating of their physician included perceived mental health ($r = 0.23$, $p < .0001$) and being active socially ($r = .19$, $p = .005$) but not perceived physical health ($r = .05$, $p = .16$) or overall health status ($r = .12$, $p < .56$). Validity coefficients were moderate ($r = -.25$ to $-.42$, $p < .0001$) and negative; this was expected because larger GHS scores imply better health, while larger PCQ scores represent greater morbidity. Both the GHS and the PCQ had pain scales which were highly correlated ($r=.75$, $p<.0001$), suggesting they measured a similar construct.
Conclusions: The research explored physician skills related to providing palliative care for chronic or serious illnesses. The short form GHS served as a validating criterion for the PCQ scales. The PCQ assesses physician–patient communication, patient-reported outcomes, relationships among these outcomes, and patient experiences with their physicians.

PROMIS-29 en français: Coordinating Translation Efforts in Canada and France for a Universal Approach

Susan J. Bartlett[1,2], Marie-Eve Carrier[3], Helena Correia[4], Linda Kwakkenbos[1,3], Brett Thombs[1,3], SPIN Investigators

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

Institutional Affiliations: [1]McGill University, Montreal Canada; [2]McGill University Health Centers, Montreal Canada;[3]Lady Davis Institute, Jewish General Hospital;[4]Northwestern University;

Objective: Increasingly, research involves patients who complete outcomes in different languages. This occurs in countries with more than one common language, such as Canada, as well as in international collaborations, which are utilized frequently in rare diseases such as scleroderma. Our goal was to develop a French translation of the PROMIS-29 questionnaire for use in clinical research, care, and population monitoring applications.

Methods: As part of a larger process to translate health measures for an international initiative, the Scleroderma Patient-Centered Intervention Network (SPIN), we followed PROMIS guidelines and a universal approach (one translation for multiple French-speaking countries) to translate the PROMIS-29 into French.

Working with the PROMIS Statistical Center, we identified 15 items (of 29) requiring translation. Two forward translations were conducted (1 in France, 1 in Québec) and then reconciled, and the reconciled version was back-translated. Expert reviewers independently reviewed the translation and all preceding steps to recommend the most appropriate translation for each item. Expert reviewers were clinicians associated with SPIN in Montréal and Paris. A quality review was performed by the PROMIS Statistical Center to address consistency between items and with other translations. Linguistic validation of the translated items was done by conducting cognitive debriefing interviews with French patients from France and Québec. Scleroderma patients were recruited through the Jewish General Hospital (Québec) and the French Scleroderma Patient Association (France). Ethical approval was obtained.
Results: Translations were reviewed by 4 clinicians, 1 PROMIS Statistical Center expert and 4 project staff. Overall, forward and backward translation resulted in French items that were acceptable and consistent with English meaning to reviewers. Debriefing interviews were conducted with seven francophone women from Québec and France, either by phone or videoconference. Patients reported that all questions and responses were easy to understand and none were offensive. Minor wording changes were suggested by patients to improve the wording of Q22 (I have trouble doing all of the family activities that I want to do) and Q24 (I have trouble doing all of the activities with friends that I want to do). It was suggested that items use "prendre part à" to avoid using the verb "faire" twice in the same sentence. Alternate forms of Q26 (How much did pain interfere with work around the home?) and Q28 (How much did pain interfere with your household chores?) were tested and minor changes were made based on patient preferences.
Conclusions: The French version of the PROMIS-29 questionnaire has been rigorously translated using a universal approach and adds to the growing list of PROMIS measures available through Assessment Center. Cross-linguistic measurement equivalence should be evaluated using quantitative data, which is an important prerequisite for pooling data obtained in multiple languages.

How Well Do Generic Patient PROMIS Instruments Capture Health Status in Rheumatoid Arthritis?

Bartlett SJ [1, 2], Orbai AM [2], Jones M [2], Bingham CO [2]

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

Institutional Affiliations: [1] McGill University, Montreal Canada; [2] Johns Hopkins School of Medicine

Objective: PROMIS offers precise, reliable generic measurement of physical, mental and social health across chronic conditions. However, little is known about the validity and performance of PROMIS instruments in rheumatoid arthritis (RA), where high levels of pain and fatigue and low physical function and mood are common.
Methods: Data were from the baseline visit of the first 125 RA patients enrolled in an ongoing study to evaluate the impact of systematically integrating broader assessment of patient-valued symptoms and impacts into ongoing arthritis care. Patients completed PROMIS computerized adaptive tests (CATs) assessing pain, fatigue, physical function, mood, sleep and social roles/activities using a tablet computer linked to Assessment Center in the waiting room immediately prior to the visit. Legacy measures (100 mm VAS for pain and fatigue, MHAQ) were also obtained as well as traditional clinical indicators of disease activity. Relationships between PROMIS and legacy measures were assessed using correlation and regression, and across CDAI disease activity levels with ANOVA.

Results: Patients were mostly female (79%) and white (86%) with a mean (SD) age of 56 (13) and disease duration of 12 (9) yr; 10% were diagnosed <= 2 yr. PROMIS CATs included an average 63 (12) items requiring 12 (5-32) minutes to complete. PROMIS CATs for pain (intensity, impact), fatigue, physical function, anxiety, depression, anger, sleep (disturbance, impairment) and social roles/activities (participation, satisfaction) correlated highly with pain VAS (rho's=.82-.83), fatigue VAS (rho=.86) and mHAQ (rho=-.74) (p's<.001). A dose-response relationship was evident in PROMIS measures across remission, low and moderate disease activity levels (except anger). Floor effects were common in legacy measures (16%, 11% and 43% for pain VAS, fatigue VAS and MHAQ, respectively) and were not common in PROMIS instruments.
Conclusions: These data contribute preliminary evidence of convergent and known groups validity and demonstrate generic PROMIS instruments can reliably assess health status, symptoms and impacts at health visits in people with RA. PROMIS CATs can be completed relatively quickly with results available immediately, and appear to address some of the well-recognized limitations (non-linearity, floor effects) of existing legacy measures in RA.

Estimating Pediatric PROMIS Item Bank Item Parameters Representative of the US Population

Brandon D. Becker, MPH[1], Adam Carle, PhD[2], Katherine B. Bevans, PhD[1,3], Carole A. Tucker, PT, PhD[4], Rachel E. Teneralli, MS[1], Christopher B. Forrest, MD, PhD[1,3]

Presenting Author: **Brandon D. Becker**; beckerb1@email.chop.edu

Institutional Affiliations [1] Children's Hospital of Philadelphia; [2] Cincinnati Children's Hospital and Medical Center; [3] University of Pennsylvania School of Medicine; [4] Temple University

Objective: To investigate a method to improve the scoring accuracy of 22 PROMIS pediatric measures by linking each measure's item parameters to a metric established in a representative sample of non-institutionalized US children aged 5-17.
Methods: Sample 1: We administered child- and parent-report item banks measuring children's stress experiences, family relationships, physical activity, and subjective wellbeing. We recruited convenience samples ranging in size from approximately 900-1,000 for parents of children aged 5-17 and 1,800-2,000 for children aged 8-17 from schools, pediatric clinics, and an internet panel. We estimated item parameters using the Graded Response Model (GRM) and developed 8 item short forms (SF-8s) for each bank. Sample 2: We recruited 5,206 parents of children aged 5-17 and 4,005 children aged 8-17 from a probability-based internet research panel to obtain a pediatric sample representative of the non-institutionalized US population. Participants completed 1 of 4 randomly assigned questionnaire sets, each comprised of 4-6 SF-8s. Data collection continued until age and gender quotas were met for each form. Using raking methods, the sample was weighted according to US population prevalence for child age, gender, race, ethnicity, geographic region, and household income. We used a fixed-parameter method to link Sample 1's full bank item parameters to Sample 2's SF-8 item parameters. Once linked, the parameters would result in a distribution of scores with a mean and standard deviation representative of the US population (T-score metric M=50, SD=10). To achieve this, we first estimated short form IRT parameters in the representative sample (Sample 2). Then in Sample 1, we fixed each measure's short form item parameter values to the values that resulted from estimating short form IRT parameters in Sample 2 and freely

estimated the remaining parameters for each measure.
Results: We compared linked and non-linked (original) item parameters, item characteristic curves (ICCs), test characteristic curves (TCCs), and EAP-based IRT scores. Generally, the correlations between linked and non-linked scores were large (e.g., r>0.98). However, examination of the ICCs and TCCs showed that differences in expected scores did result across linked and non-linked parameters for most measures. Generally, the non-linked parameters tended to result in larger expected scores and the difference occurred mainly at low trait levels. However, for psychological stress and physical stress the non-linked parameters resulted in smaller expected scores. For psychological stress, the differences occurred across the trait, while for physical stress, the differences occurred mainly at high trait levels. Finally, we observed little difference between linked and non-linked parameters for the physical activity banks.
Conclusions: For the 22 linked pediatric PROMIS item banks, EAP-based scoring will generate child health outcome values that can be interpreted relative to the US population. Look-up tables will produce similarly interpretable estimates. Two potential limitations of this approach are 1) the shifting of item parameters that results in differential item ordering or 2) adjustments in the distance between discrimination parameters when comparing the full item bank calibration parameters in Sample 1 to the linked parameters.

Do PROMIS Cognitive Abilities Scores among People with Multiple Sclerosis Improve Following a Cognitive Intervention?

Heather Becker, Ph.D., Alexa Stuifbergen, R.N., Ph.D., FAAN, Ashley Henneghan, R.N., MSN

Presenting Author: **Heather Becker**, Ph.D; heatherbecker@mail.utexas.edu

Institutional Affiliations: [all] The University of Texas at Austin School of Nursing

Objective: Because cognitive limitations have been observed in 50-75% of people with multiple sclerosis, we tested the efficacy of an intervention designed to build cognitive abilities. We used the PROMIS V1.0 8a Cognitive Abilities Scale to compare ratings of cognitive abilities before and after the 8-week intervention.
Methods: Participants from three Texas cities were recruited through local neurologists and other community contact. We used a randomized clinical trial to compare change over time between those who received the MAPSS-MS intervention with those who were referred to standard of care cognitive resources. Results: The tri-ethnic predominantly female sample (n=82) had an average age of 49 years. They had been diagnosed with M.S. an average of 13 years. Results of paired t-tests yielded a statistically significant increase in PROMIS Cognitive Abilities Scores for the intervention group ($t=2.55$, $p<.05$,$df=1/35$) while there was no significant change for the comparison group ($t=-.47$, NS, $df=1/25$).
Conclusions: The 8-item Cognitive Abilities Scale detected change over time in preliminary analysis of the efficacy of a cognitive intervention for people with multiple sclerosis. Short yet psychometrically sound measures are particularly needed in research about people with multiple sclerosis, whose fatigue makes it difficult for them to complete lengthy batteries.

Qualitative development of pediatric oral health item banks

Jean Hee Moon [1], Katherine B. Bevans [1,2], Ramya Pratiwadi, Rachel Teneralli [1], Joanie Jean [1], Christopher B. Forrest

[1,2]

Presenting Author: **Katherine B. Bevans**; bevans@email.chop.edu

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

Objectives: To describe the qualitative development of child- and parent/proxy-report outcomes measures (PROMs) of pediatric oral health. The PROMs assess children's oral pain and wellbeing.
Methods: We developed item pools in accordance with Patient Reported Outcome Measurement Information System® (PROMIS) standards. A pediatric oral health conceptual framework was specified through a literature review and interviews with 13 content experts, 28 children (aged 8-17), and 25 parents (of children aged 5-17). Initial item pools were informed by a systematic review of existing oral health PROMs. Items were written and iteratively refined on the basis of child cognitive interviews.
Results: We developed a pediatric oral health conceptual framework that differentiated core dimensions of oral health (e.g., oral health conditions, oral functioning) from their influences (e.g., healthcare, hygiene) and effects (e.g., social/mental health). Experts identified dimensions that are directly or indirectly targeted by oral healthcare and those that are best assessed by child- and/or parent/proxy-report (rather than by clinical exam). Children and parents identified patient-centered oral health terminology (e.g., "wiggly teeth" or "a hole in your tooth").The initial oral pain item pool contained 62 items that represented pain quality, pain location, temperature, bleeding/swelling, and pain impact. Nine poorly understood items were deleted from the item pool. The item pool is written at the 1.6 grade level. The initial wellbeing item pool contained 42 items that represented satisfaction with appearance, perceived overall attractiveness, satisfaction with reflection, desire for enhancements, halitosis, and wellbeing impact. All items were well understood. The item pool is written at the 1.6 grade level.
Conclusions: The final oral health item pools are comprised of 53 pain items and 42 wellbeing items. Next, we will generate unidimensional calibrated item banks, short-forms, and computerized adaptive test algorithms using child- and parent-report data collected from an internet panel. A cross-sectional validation study will be conducted to assess the degree to which the PROMs differentiate between groups known to differ on oral health outcomes.

Using PROMIS in PCOR: Lessons from PCORI's Pilot Projects

Clifton O. Bingham, III, MD [1], Peter A. Merkel, MD, MPH [2],Thelma Mielenz, PhD [3], Paul A. Pilkonis, PhD [4], Lauren Edmundson [5], Raj Sabharwal, MPH [5]

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

Institutional Affiliations: [1] Johns Hopkins University, Baltimore, MD [2] University of Pennsylvania, Philadelphia, PA [3] Columbia University, New York, NY [4] University of Pittsburgh School of Medicine, Pittsburgh, PA [5] AcademyHealth, Washington, DC

Objective: The goal of this project was to investigate the utility, feasibility, and scope of using Patient-Reported Outcome Measurement Information System (PROMIS) measures in patient-centered outcomes research (PCOR). Four pilot projects funded by the Patient-Centered Outcomes Research Institute (PCORI) that used PROMIS in their studies provided an opportunity for investigation.
Methods: The four PCORI pilot projects provided their experiences using PROMIS

measures, computerized adaptive testing (CAT), and the Assessment Center in their unique PCOR studies via a self-report survey. The survey included questions about reasons for choosing to use specific PROMIS measures, methods for implementing PROMIS in their particular setting, challenges encountered, and identified solutions. The information was analyzed and synthesized to identify common themes.
Results: Despite differences in the specific research questions, patient populations, and investigators' experiences in the pilot projects, these studies revealed that PROMIS can be used successfully in PCOR. The experiences of the study teams highlighted the need to understand patient populations; specifically, researchers should consider what types of data are important to patients and how the PRO data can optimally be shared with patients to improve care. For example, several pilot projects tested different ways to present PROMIS data to determine which format was best understood and preferred. Successful use of PROMIS also requires advanced planning regarding implementation, including logistics and technology. The pilot projects identified successful strategies to make data collection feasible in a variety of settings, minimizing the burden and optimizing efficiencies of data collection for patients, researchers, clinicians, and clinic staff. For instance, several pilot projects reported working closely with clinic staff and seeking their input regularly. Additionally, challenges in using PROMIS CATs were overcome through collaboration with PROMIS Assessment Center staff and by making accommodations for patients, such as providing styluses to patients with limited dexterity.
Conclusions: While the pilot projects faced some challenges in implementing PROMIS, their overall successes demonstrate that PROMIS is a valuable tool for PCOR. These collective experiences and lessons
learned can be used to guide and encourage future researchers, and expand the use of PROMIS in PCOR.

Sarcoma Survivors' Patient-Reported Late-effects of their Treatment

Nina P. Bobowski, MPH, Laurence H. Baker, DO, Sarah Hawley, PhD

Presenting Author: **Nina P. Bobowski**, MPH; npbo@med.umich.edu

Institutional Affiliations: University of Michigan Health System, Ann Arbor, MI

Objective: To utilize PROMIS questionnaires as a screening and diagnostic measure to collect patient self-reported data at baseline and over time to tailor clinical and educational recommendations in our Sarcoma Survivorship Program.
Methods: We prospectively evaluate sarcoma survivors with the PROMIS questionnaires on Anxiety, Depression, Mobility, Pain Interference, Sleep Disturbance, and Physical Function. Patients complete these questionnaires before their first clinic visit and then every six months thereafter. During a survivorship clinic visit, patients are seen by medical oncologist, Dr. Baker and cardio-oncologist, Dr. Monika Leja. In addition to Baker's and Leja's comprehensive health exams and the patient's medical and operative records, imaging, relevant pathology reports, and other relevant medical records and tests, the PROMIS measures help to address the healthcare needs of these survivors no longer in active treatment including monitoring long-term and late physical effects of cancer and its treatment, management of psychological and comorbid medical conditions. The Survivorship Program is focused on risk-based screening and health promotion for survivors. At the completion of each visit, patients receive a summary of their prior sarcoma therapy as well as a written plan for survivorship. Our model of care is patient-centered so that services provided are appropriate to the level of care required by the individual survivor. The PROMIS questionnaires provide patient self- reported assessments that

enable our program to provide appropriate survivorship plans based on the individual patient's needs and capabilities. The Program's annual survivorship clinic visits and the regular PROMIS questionnaires will allow us to monitor and recognize potential physical and mental conditions due to the disease or treatment over time as well as allow us to diagnose and treat any conditions earlier and with better outcomes.
Results: Research shows that many cancer survivors have deficient health-related knowledge and are not engaging in recommended health promotion and screening practices that could improve their long-term outcomes. The well-being of the survivor requires a level of integrated coordination requiring a cultural shift in approach away from a disease-only focus. In addition to routine surveillance for recurrence of cancer, survivors' follow-up management requires proactive care, which includes systematic planning for cancer prevention and patient-centered surveillance based on the survivor's personal risk, cancer therapy, genetic predispositions, lifestyle behaviors, physical and mental abilities, and other comorbid health conditions.
Conclusions: Most of what we know about survivorship focuses largely on the period between diagnosis and 2 years after treatment (the early survivorship phase). However, most late effects of cancer treatment have much longer latency periods and tend to occur during the extended survivorship years. Our limited knowledge and many questions about the health status, functioning, and quality of life for this understudied survivor group underscores the need for continued research in this growing population. The PROMIS questionnaires provide reliable assessments of self-reported health that aid our research to address the specific gaps in our knowledge such as the incidence of and risk factors for late and long-term effects of sarcoma and its treatment and appropriate follow up care and surveillance for survivors.

PF, PI, and Mobility CAT validation against the FAOS for Common Foot and Ankle Pathologies

Jayme C. Burket, PhD; Stephen T. Greenfield, MD; Huong T. Do, MA; Scott J. Ellis, MD

Presenting Author: **Jayme C. Burket**; burketj@hss.edu

Institutional Affiliations: [all] Hospital for Special Surgery, New York, NY

Objective. Validate the physical function (PF), pain interference (PI), and Mobility PROMIS CATs using the Foot and Ankle Outcome Score (FAOS), a well-performing and validated "legacy" instrument, in patients with the six most common foot and ankle pathologies.
Methods*:* Patients (n=131) with osteochondral defects of the talus, hallux rigidus, hallux valgus, adult acquired flatfoot, ankle instability, or ankle arthritis were consented and enrolled at their initial surgical visit. Eligible patients were those indicated for surgery for a single, isolated condition. PF, PI, and Mobility CATs, and the FAOS were administered at the preoperative visit and again one week later prior to surgery. The order of administration was randomized to minimize bias due to questionnaire ordering and fatigue, and the time to complete each test was recorded. Test-retest reliability was assessed with intraclass correlation coefficients (ICC) and 95% confidence intervals (SAS 9.3). Precision was assessed with standard error of measurements (SEMs), and the minimum detectable change (MDC) was calculated for each instrument. Convergent validity was assessed with Spearman's correlation coefficients. Results are presented as means and 95% confidence intervals.
Results: ICCs of the PF (0.86 [0.81, 0.90]) and Mobility (0.85 [0.80, 0.89]) CATs were similar to the FAOS subscales (ICCs range: 0.83 to 0.91), whereas that of the PI CAT (0.75 [0.67,0.82]) was still acceptable but

slightly lower, likely due in part to the small range of pain scores observed preoperatively. CAT SEMs ranged from 2.62 to 3.38, whereas FAOS SEMs ranged from 7.20 to 12.08, indicating substantially poorer precision. Consequently, the MDCs of the CATs (range: 7.26-9.37) were smaller than that of the FAOS (range: 19.95- 33.50). The Mobility CAT was most strongly correlated with the Activities of Daily Living subscale (Spearman's: 0.79). Additionally, the PF CAT had a Spearman's correlation of 0.70 with this subscale. The correlation between the PI CAT and the FAOS pain subscale was -0.68. The average time to complete each CAT was under 1 minute, whereas the FAOS took nearly 6 1/2 minutes.

Conclusions: The PF, PI, and Mobility CATs demonstrated comparable test-retest reliability with substantially better precision than the FAOS in preoperative patients with the six most common foot and ankle conditions. Further evaluation of these instruments should include additional psychometric evaluation, including Rasch modeling to evaluate unidmensionality and person and item reliability, post-operative convergent validity, and responsiveness for each foot and ankle condition. With their superior efficiency and enhanced precision, the PF, PI, and Mobility CATs show great potential to reduce respondent burden in foot and ankle outcomes research, as patients will spend less time completing these outcomes instruments, and fewer patients will be required to achieve adequate study power.

Social Media Use and Anxiety as Measured by the PROMIS Anxiety Short Form

Jason B. Colditz, MEd , Ariel Shensa, MA, Jacquelynn R. Jones, BS, Liu yi Lin, BA, Paul A. Pilkonis, PhD, Bruce L. Rollman, MD, MPH, Brian A. Primack, MD, PhD

Presenting Author: Jason B. Colditz, MEd; colditzjb@upmc.edu

Institutional Affiliation: [all] University of Pittsburgh School of Medicine, Pittsburgh, PA

Objective: To determine associations between social media use and Patient-Reported Outcomes Measurement Information System (PROMIS) 4-item Anxiety Short Form scores in a representative sample of US young adults.

Methods: We surveyed a nationally representative sample of US young adults between the ages of 18 to 30. Data collection was assisted by a survey research company that recruited participants via random-digit dialing and address-based sampling frames, representing over 97% of the US population. Participants responded to online surveys over a 1-month period in October/November of 2014. We assessed our independent variable, social media use, in three different ways, including (1) self-reported time spent with online social networks per day, (2) average number of social network site (SNS) visits per week, and (3) responses on a global frequency scale that was adopted from the Pew Internet Research Questionnaire. Our primary dependent variable was self-report anxiety as measured by the 4-item PROMIS short form. We used ordered logistic regression with sample weights to assess associations between SNS use and anxiety while controlling for relevant socio-demographic, personal, and environmental factors. All logistic regression models satisfied the proportional odds assumption.

Results: Our sample of 1,781 participants was 61.8% female, 64.1% Caucasian, 10.0%

African American, 16.6% Hispanic and 9.3% Mixed or other race. Our fully adjusted multivariate models included all covariates, including age, sex, race, ethnicity, relationship status, living situation, household income, and education level. T-Scores for anxiety (*M*=51.43, *SD*=9.73) were similar to established PROMIS norms, although they demonstrated a strong floor effect, with roughly 1/3 of participants having the lowest possible T-Score of 40.3. Therefore, anxiety T-Scores were converted to tertiles prior to analysis to compensate for the skewed distribution and non-normality (Shapiro-Wilk *W*=0.99, *p*<.001). Multivariate analyses demonstrated that, compared to those in the lowest quartile for total SNS use per day (0-30 minutes of use), participants in the highest quartile (2 or more hours of use) had increased odds of having greater anxiety (AOR=1.86, 95% CI=1.28-2.69). In addition, compared with those in the lowest quartiles, those in the highest quartiles of SNS checks per week (AOR=2.49, 95% CI=1.70-3.67) and the global frequency scale (AOR=2.34, 95% CI=1.80-3.03) reported greater anxiety. All associations between independent variables and anxiety demonstrated strong, linear, dose- response trends (*p*≤.001) that was robust to sensitivity analyses.

Conclusions: We identified a strong, positive association between anxiety and SNS use among a representative sample of US young adults. This association persisted across three different methods of operationalizing SNS use and after adjusting for socio-demographic covariates. This study highlights an important association, given the increasing prevalence of SNS use and the substantial morbidity associated with anxiety. It also suggests that it may be valuable to use SNS to identify individuals at risk for anxiety or mood disorders.

Is Higher Social Media Use Associated with Reduced Social Isolation?

Brian A. Primack [1], Jason B. Colditz, MEd [1], Ariel Shensa, MA [1], Daniel Rosen, PhD [2], Carol Cheng [1], Jamie Harris [3], David Brent, MD, PhD [1,3]

Presenting Author: **Jason B. Colditz**, MEd; colditzjb@upmc.edu

Institutional Affiliations: [1] University of Pittsburgh School of Medicine, Pittsburgh, PA [2] University of Pittsburgh School of Social Work, Pittsburgh, PA [3] Western Psychiatric Institute and Clinic, Pittsburgh, PA

Objective: Social isolation is associated with increased morbidity and mortality from physical and mental health conditions. Recent increases in use of social networking sites (SNS) such as Facebook may provide opportunities for alleviation of social isolation. However, these associations have not been tested empirically. Therefore, we aimed to assess multivariable associations between social media use and social isolation in a nationally-representative sample of US young adults.

Methods: We surveyed a nationally-representative sample of US young adults between the ages of 18 to 30. We assessed social media use in three different ways, including self-reported time spent on SNS per day, average number SNS visits per week, and responses on a global frequency scale adopted from the Pew Internet Research Questionnaire. Each of these latter two scales prompted participants regarding the most commonly used 11 SNS, such as Facebook, Twitter, Instagram, Tumblr, and Reddit, and combined response data. Our primary dependent variable was Social Isolation as measured by the 4-item Patient-Reported Outcomes Measurement Information System (PROMIS) short form. Based on the distribution of data, we collapsed this dependent variable into tertiles for analysis. We used ordered logistic regression with sample weights to

assess associations between SNS use and social isolation while controlling for relevant socio-demographic, personal, and environmental factors.
Results: Our sample of 1,781 participants was 62% female, 64% Caucasian, 10% African-American, 17% Hispanic and 9% of mixed or other race. Fully-adjusted multivariable models included all covariates, including age, sex, race, ethnicity, relationship status, living situation, household income, and education level. These multivariable analyses demonstrated that, compared to those in the lowest quartile for total SNS use per day, participants in the highest quartile had twice the odds of having greater social isolation (AOR=2.00, 95% CI=1.41-2.83). Similarly, compared with those in the lowest quartiles, those in the highest quartiles of SNS visits per week (AOR=3.38, 95% CI=2.26-5.06) and the global frequency scale (AOR=3.25, 95% CI=2.21-4.80) reported substantially greater social isolation. All associations demonstrated strong, linear, dose-response trends ($p<.001$), and results were robust to all sensitivity analyses.
Conclusions: We found robust, linear associations between SNS use and increased social isolation, even after adjusting for all covariates. Therefore, this study suggests that young adults with high SNS use are more, and not less, likely to be socially isolated. While the reason for these counter-intuitive results is not clear, it is possible that socially isolated young adults are substantially more likely to seek out community on SNS. It is also possible that increased SNS use paradoxically leads to an increased feeling of social isolation. This may be because (1) time online displaces more authentic social experiences and/or (2) certain characteristics of the online milieu may lead to feelings of being "left out." Regardless of the direction of this link, however, this study highlights an important association, given the increasing prevalence of SNS use and the substantial morbidity associated with social isolation.

Uncle Sam Wants PROMIS: Leveraging PROMIS to Build and Pilot the DoD's Pain Assessment Screening Tool and Registry (PASTOR)

Karon Cook, PhD; [1], Michael Kallen, PhD; [1], Chester Buckenmaier, MD; [2] Richard Gershon, PhD[1]

Presenting Author: **Karon Cook, PhD**; karon.cook@northwestern.edu

Institutional Affiliations: [1] Northwestern University; [2] Defense and Veterans Center for Integrative Pain Management

Objectives: In 2010 the Army Pain Management Task Force (PMTF) delineated recommendations for improving pain care in the military health system. One of these recommendations was implementation of a comprehensive electronic patient assessment tool that could be standardized across DoD and VHA clinics. The resulting tool, Pain Assessment Screening Tool and Outcome Registry (PASTOR), includes both PROMIS and military-specific measures. In addition to being a data collection tool, PASTOR is also a clinical reporting tool for improving patient-clinician encounters, as well as a registry of pain data and its correlates that, in the future, will be mined in "Big Data" analyses to inform the effectiveness of alternative pain treatment strategies and to support "personalized medicine". This presentation describes PASTOR's development and implementation and present results from retrospective data analysis of the first year of piloting.
Methods: Identification of measures to include in PASTOR was accomplished with input from multiple stakeholders across clinical types and specialties. In 2011 a large cadre of military health specialists, researchers, and scientific advisors met to discuss how PASTOR should be constituted. Input from this large stakeholders' meeting was refined, and additional input was obtained from clinicians and outcomes researchers. Many of the variables

stakeholders identified as important to include in a registry were patient-reported. The decision was made to use PROMIS measures when available for these outcomes. Stakeholders' input for a clinical report was solicited, a sample clinical report was developed and circulated, and modifications were made based on feedback. An initial version of PASTOR was integrated into a standalone AC-Lite application for piloting at Walter Reed Army Medical Center and at Madigan Army Medical Center. Based on additional feedback from clinicians and patients using the pilot, PASTOR measures are now being implemented within a military health system's electronic health record system.

Results: Twelve PROMIS measures were selected for inclusion in PASTOR (Pain Intensity, Pain Interference, Pain Quality, Physical Function, Fatigue, Sleep-related Impairment, Headache, Depression, Anxiety, Anger, Alcohol Use, and (when it becomes available) Abuse/Misuse of Prescription Medication. A provider's summary report enhances the clinical management of patients with pain. This report is a summary of an individual patient's status based on that patient's self-report. Patients' self-reported goals are included on the report along with their self-rated progress towards these goals. Graphs and tables show opioid equivalents, pain maps, symptoms and outcomes, and progression towards goals over time in addition to concerning results from the most recent evaluation (e.g., screened positive for depression, alcohol misuse, PTSD). The report is a powerful tool to efficiently focus the limited clinical encounter time on specific treatment issues that the patient has identified as relevant. In the first year of the PASTOR pilot, >400 individuals completed PROMIS measures and >150 completed up to 4 follow-up assessments. A retrospective study is underway to evaluate the validity and response burden of PROMIS measures. Results of these analyses will be presented.

International Use of PROMIS®: Multilingual Translation Efforts and Standards

Helena Correia, David Cella

Presenting Author: **Helena Correia, BA**; helena-correia@northwestern.edu

Institutional Affiliations: [all] Northwestern University

Objective: Produce accurate and culturally appropriate multilingual translations of PROMIS measures, to enable their use with non-English-speaking populations and in international comparative health status research.

Methods: PROMIS measures have been translated through an iterative process of independent forward translations, reconciliation, back-translation, bilingual expert review (mainly linguistic but also from healthcare professionals where needed), and pre-testing with cognitive debriefing (i.e., linguistic validation). Harmonization across languages (i.e., assessing translation accuracy and the range of variation between languages) and a universal approach to translation (i.e., aiming to produce one language version for multiple countries instead of country-specific versions of the same language) have guided the process. Each item has been cognitively debriefed with at least 5 native speakers of the target language.

Results: PROMIS measures have been translated into over 30 languages and efforts are underway to translate some instruments into almost as many languages. Cognitive debriefing results have usually indicated that the items were well understood by the target population and conceptually equivalent to the English source. Translations have been revised after cognitive debriefing as needed, when difficulty has been reported or if participants' comments revealed misunderstanding of an item's intended meaning. The process of harmonization across languages has delimited the range of

acceptable variation between languages and ensured the equivalence of the translated versions vis a vis the English source. Using a universal approach to translation has prevented the creation of unnecessary language versions and reduced the bias introduced by multiple versions of the same language produced at different times. Country-specific versions of selected items have been created only when a universal version was not possible (due to irreconcilable differences). This presentation inventories available and in-progress translations of adult and pediatric item banks and short forms by domain.
Conclusions: The linguistic validation methodology used to translate PROMIS instruments meets the prevailing standard for release of translated PRO instruments. Translated PROMIS instruments are conceptually equivalent to the English source and can be used in research and clinical practice.

Lessons from implementing and collecting PROs for HIV care and research: results from 10,000 patients

Rob E Fredericksen, PhD [1], Dan Drozd, MD [1], JAC Delaney, PhD [1], Robin Nance, MPH [1], WC Matthews, MD [2], Ken Mayer, MD [3], Greer Burkholder, MD [4], James Willig, MD [4], Sonia Napravnik, PhD [5], Joseph Eron, MD [5], Katerina Christopoulos, MD [6], Elvin Geng MD [6], Michael Mugavero, MD [4], Justin McReynolds, MS [1], Greg Barnes, MS [1], William Lober, MD [1], Mary E McCaul, MD [7], Heidi Hutton, PhD [7], Geetanjali Chander, MD [7], Richard Moore, MD [7], Mari M Kitahata, MD, MPH [1], Paul K Crane, MD, MPH [1], Michael Saag, MD [4], Heidi M Crane, MD, MPH [1]

Presenting Author: **Heidi M Crane**, MD MPH, hcrane@uw.edu

Institutional Affiliations: 1] University of Washington, Seattle, WA [2] University of California, San Diego, San Diego, CA [3] Harvard University, Boston, MA [4] University of Alabama at Birmingham, AL [5] University of North Carolina, Chapel Hill, NC [6] University of California, San Francisco, San Francisco, CA [7] Johns Hopkins University, Baltimore, MD

Objective: Information from patient-reported outcomes (PROs) can enhance patient-provider communication, improve clinical outcomes, and facilitate clinical research. Providers often underdiagnose at-risk behaviors and outcomes, including depression, suicidal ideation, alcohol and drug use, and poor medication adherence. However, there are barriers to collecting PROs in clinical settings. Recent technological advances may help overcome these barriers. We examined the feasibility, benefits, and PRO results of implementing a web-based application on tablet PCs with touch screens to collect PROs in seven busy, multi-provider, outpatient HIV clinical care settings across the CNICS cohort.
Methods: English and Spanish-speaking patients from seven sites across the U.S. presenting for routine clinical care visits completed a clinical assessment of PROs which contained 62 to 113 items depending on patient responses. The assessment included instruments measuring adherence to antiretroviral therapy (ART), drug/alcohol/tobacco use, depression, anxiety, symptom burden, health-related quality-of-life, physical activity, and body morphology changes. Among a subset of ~1500 patients, we examined provider awareness of at-risk behavior as measured by same-day clinical documentation measured in the 8 months before and after initiating delivery of PRO results to providers as part of clinic visits. We interviewed subsets of patients and providers to assess impact and preferences including formats and structures for delivery of assessment results. We used Plan-Do-Study-Act methods for quality

improvement to address concerns raised regarding clinic flow, technology, scheduling, and delivery of assessment results, Results were delivered to providers prior to same-day clinical care visits to enhance care and also incorporated into the CNICS Data Repository to facilitate clinical research.
Results: To date, >10,000 CNICS patients have completed the assessment at routine visits every 4-6 months totaling 45,078 assessments including 10,372 assessments in 2014 alone. The mean age of patients who have completed the assessment is 45 years. Median completion time is ~11 minutes. Feedback from the PROs increased provider documentation of depression, inadequate adherence substance use, at-risk alcohol use but not sexual risk behavior. Among the 89% on ART, mean adherence was 2.5% higher among those over 60, compared with those >30, or 30-40 years old (p values <0.05). Moderate or severe depression was reported by 24% of patients overall, and 26% among those with at risk alcohol use, 29% among those currently using marijuana, 30% among current smokers, 32% among those with hepatitis C virus co-infection, 38% among current cocaine/crack or amphetamine/ crystal users, and 46% of current heroin users.
Conclusions: It is feasible to integrate a web-based PRO clinical assessment into busy multi-provider HIV primary care clinics. We found a high prevalence of inadequate medication adherence, depression, at-risk alcohol use, and substance use. PRO data improved provider recognition of key patient behaviors and outcomes for many domains. Successful implementation of PROs required careful tailoring of approaches at each site to minimize disruption to clinical flow. PROs prior to visits enables improved patient care and provides useful data for cohort research.

Measurement precision and differential classification of at-risk alcohol use with AUDIT-C vs. PROMIS in HIV

Paul K Crane, MD, MPH [1], Laura Gibbons, PhD [1], Rob Fredericksen, PhD [1], Mary E McCaul, MD [2], Geetanjali Chander, MD [2], Heidi Hutton, PhD [2], William Lober, MD [1], W Chris Matthews, MD [3], Ken Mayer, MD [4], Greer Burkholder, MD [5], Michael Mugavero, MD [5], James Willig, MD [5], Michael Saag, MD [5], Mari Kitahata, MD, MPH [1], Todd Edwards, PhD [1], Donald Patrick, PhD [1], Heidi M Crane [1]

Presenting Author: **Paul K Crane, MD MPH,** pcrane@uw.edu

Institutional Affiliations:[1] University of Washington, Seattle, WA [2] Johns Hopkins University, Baltimore, MD [3] University of California, San Diego, San Diego, CA [4] Harvard University, Boston, MA [5] University of Alabama at Birmingham, AL

Objective: Performance of the PROMIS alcohol use short form for identifying at-risk alcohol use among people living with HIV (PLWH) is unknown. We sought to compare at-risk alcohol use defined by PROMIS to definitions from the AUDIT-C.
Methods: Patients from 4 CNIS cohort sites: Seattle, Boston, Birmingham, and San Diego completed tablet-based assessments at the time of clinic appointments from 4/13-2/14 that included the AUDIT-C and PROMIS alcohol use short form. Patients who completed both instruments and reported using alcohol at least monthly were included. We used three at-risk alcohol definitions derived from the AUDIT-C: 1) scores ≥5 for men and ≥4 for women (high AUDIT-C); 2) scores ≥4 for men and ≥3 for women (low AUDIT-C); 3) at least monthly binge drinking (binge). We used item response theory (IRT) to calibrate the AUDIT-C to the PROMIS metric. We used a receiver operator characteristic (ROC) curve analysis with the higher AUDIT-C to identify at threshold for PROMIS at-risk

alcohol use. We used that threshold for PROMIS short form IRT scores. We examined plots of standard error of measurement against IRT-based alcohol scores focusing especially on the region close to the at-risk alcohol threshold.
Results: There were 2,219 PLWH who drank at least monthly. Of these, 658 (30%) had at-risk alcohol use based on high AUDIT-C, 1,022 (46%) using low AUDIT-C, 448 (20%) using binge, and 557 (25%) using PROMIS. In all, 1,103 (50%) were not identified as having at-risk alcohol use by any definition, and 287 (13%) had at risk alcohol use by all of the definitions, leaving 829 (37%) who were differentially classified across definitions. Compared to high AUDIT-C, PROMIS agreed for 79% and disagreed for 21%. For low AUDIT-C, these values were 71% and 29%, and for binge, 81% and 19%. PROMIS had greater measurement precision across the range including scores close to at-risk thresholds. Mean completion times were 18 vs. 42 seconds for AUDIT-C vs. PROMIS.
Conclusions: At-risk alcohol use ranged between 20% and 46% in PLWH who drank at least monthly depending on the specific definition used. PROMIS had high but not perfect agreement with all three AUDIT-C-derived definitions of at-risk alcohol use. PROMIS with more items has greater measurement precision near the clinically relevant threshold where precisions may be most valuable. However PROMIS also took more than twice as long to administer which may be relevant when considering use in clinical care settings. Additional analyses are underway using PROMIS alcohol use computer adaptive tests (CAT) instead of the short form and other alcohol instruments commonly used in clinical care.

Standardizing and Personalizing Patient-Centered Rheumatoid Arthritis Treatment Targets

Amy R. Eisenstein PhD. 1,2], Marissa Marshak MSW.,[3] Jennifer Beaumont MS.,[2] George
J. Greene PhD, [2] Eric Ruderman MD.,[3] and David Cella PhD.[2]

Presenting Author: **Amy Eisenstein**; amy.eisenstein@cje.net

Institutional Affiliations: [1]- CJE SeniorLife, Chicago, Illinois [2]- Department of Medical Social Sciences, Northwestern University, Chicago, Illinois [3] – Department of Rheumatology, Northwestern University, Chicago, Illinois

Objective: The goal of this study is to individualize and standardize patient reported endpoints for treatment targets in patients with Rheumatoid Arthritis (RA). Using PROMIS item banks, we developed patient-centered targets in five domains: pain, fatigue, depression, physical function, and social function. Targets were allowed to be individualized and yet they retained valid measurement. The current paper will explore implementation and baseline experience in personalizing treatment targets for RA patients.
Methods: Patients with documented RA were recruited from a Rheumatology clinic at Northwestern Memorial Hospital, in an effort to include roughly 50% in remission or with low disease activity and 50% with moderate or high disease activity, using their CDAI (Clinical Disease Activity Index) scores. Patients completed a baseline assessment while in the clinic including clinical questionnaires, PROMIS CAT's, open-ended items about their quality of life, prioritization of PROMIS domains, and selection of 5 items that they felt were most important within their most highly prioritized domain. Their responses from all five PROMIS domains were then graphed and given to their rheumatologist for future

discussion and entered into their Electronic Health Records.
Results: Sixty-six patients have completed the first assessment, with an average age of 55 (SD=11.9), 47% (n=31) exhibited high disease activity (CDAI m=13.14; SD=10.35). Respondents scored worse than average for the general US population in the PROMIS domains of fatigue (m=55.72), pain (m=57.35), and physical function (m=42.65). Baseline scores on depression and social functioning were closer to average in the general US population. When selecting prioritization areas for treatment targets, 38% of patients selected physical function, followed by 35% of participants selecting pain, 18% selecting fatigue, 6% selecting depression, and 3% selecting social function. Participants spoke openly through their process of prioritizing these five domains, and the researchers have been able to qualitatively capture thoughts and themes related to goal setting for RA treatment.
Conclusions: Most patients selected physical function as the domain of most importance for their personal treatment goal setting, with pain a close second choice. Results of this individualized PROMIS approach to RA treatment goal setting are unique. With further monitoring and follow-up of the approach described above, we will be able to examine the benefits of this patient-centered approach to treatment goal setting.

Health-related quality of life in couples seeking treatment for infertility

Kathryn E. Flynn, PhD, Elizabeth A. Duthie, PhD

Presenting Author: **Kathryn E. Flynn, PhD**; kflynn@mcw.edu

Institutional Affiliations: Medical College of Wisconsin, Milwaukee, WI

Objective: Infertility affects 17% of couples trying to conceive, and 12% of women of childbearing age have used infertility services. We collected data on self-reported health status using PROMIS measures within the context of a mixed-methods, single-center longitudinal study of couples' decision making during the 12 months after making an initial appointment with a Reproductive Endocrinologist and Infertility specialist.
Methods: Both partners of 37 couples completed semi-structured interviews and surveys at 6 time points: before an initial appointment with a reproductive specialist, one week after the appointment, and then again at roughly 2, 4, 6, and 12 months. Surveys included the PROMIS-29 with a modified set of questions for Physical Function, given that the sample was expected to be significantly younger than the US population. Surveys were administered using REDCap and scored using the Assessment Center Scoring Service. We calculated unadjusted means and standard deviations (SD) for pregnancy candidates (37 women) and their supporting partners (35 men, 2 women).
Results: For pregnancy candidates, physical function was 57 points (SD 5) at baseline (before the initial appointment) and remained there for women who did not become pregnant but declined over time to 49 (7) for women who did become pregnant. For supporting partners, physical function remained steady across all time points (56-57 points with SD of 5-7). For pregnancy candidates, anxiety was 52 (8) at baseline, 51 (8) after the appointment, and then 53-54 for the remaining time points. It was higher for the 3 women who were pregnant at 2 months at 59 (7) and the 9 women who were pregnant at 4 months at 58 (8). For partners, anxiety was 49 (8) at baseline, 46 (7) after the appointment, and then 47-49 for the remaining time points. Depression ranged from 46-49 (6-8) for pregnancy candidates and 44-47 (6-9) for partners. Fatigue increased over time for pregnancy candidates, from 45 at baseline to 49 (6-8)

at 12 months for women who did not become pregnant and to 54 (8) for women who became pregnant. For supporting partners fatigue ranged from 44-48 (8-11). For pregnancy candidates, sleep disturbance increased over time from 46-50 (6-8), and for supporting partners it was steady at 47-48 (7-9). Social role satisfaction ranged from 53-56 (6-8) for pregnancy candidates and 52-55 (6-8) for supporting partners. Finally, pain interference ranged from 44-46 (6-7) for pregnancy candidates and 46-44 (5-6) for supporting partners.

Conclusions: In this small, single-center study of couples seeking treatment for infertility, some domains of health remained steady over time while others showed changes over the 12 months after making an initial appointment with a reproductive specialist. Fatigue increased and physical function decreased with pregnancy. Anxiety was higher for pregnancy candidates compared to supporting partners at all time points and lowest for both groups immediately following the initial appointment.

Profile of subjective wellbeing among children in the US

Christopher B. Forrest, MD, PhD [1,2], Ulrike Ravens- Sieberer, PhD [3], Janine Devine, PhD[3], Brandon D. Becker, MPH[1], JeanHee Moon, PhD [1]

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

Institutional Affiliations: [1] Children's Hospital of Philadelphia; [2] University of Pennsylvania School of Medicine [3] University of Hamburg - Eppendorf

Objectives: The purpose of this study was to investigate differences, by age, in PROMIS pediatric short form scores and other measures of well-being among a representative sample of US children aged 8-17.

Methods: A probability based internet survey company was used to collect data from 4,005 children aged 8-17 years. Respondents were weighted by form using child age, sex, race, ethnicity, geographic region and household income to the March 2013 Current Population Survey Supplement with a post stratification adjustment calculated using iterative proportional fitting. Each PROMIS short form was scored with item response theory calibration parameters using Bayesian Expected A Posteriori (EAP) estimation procedure. All non-PROMIS measures were scored using published recommendations. Average scores for each measure were plotted by two-year age intervals.

Results: Levels of life satisfaction, positive affect, meaning and purpose, hope, and peer relationship quality decreased as age increased. Inversely, levels of depressive symptoms, fatigue, psychological stress, somatization and perceived stress increased as age increased

Conclusions: These results from a nationally representative sample of US children aged 8-17 demonstrate, in cross-section, a decline in positive markers of well-being that is simultaneously accompanied by an increase is negative markers of well-being as children age from middle childhood to late adolescence. These trends may reflect biologic and cognitive changes that occur as children age. Future research should monitor longitudinal changes in PROMIS measures of well-being using age-based normative scores to further understand developmental changes in pediatric well-being.

Development and Content Validation of a PROMIS Sexual Risk Measure

Authors: Fredericksen RJ, Edwards TE, Dant L, Yang F, Walcott M, Paez E, Brown S, Loo S, Gutierrez C, Jampel J, Church A, Willig J,

Mugavero M, Mayer K, Crane PK, Crane HM, Patrick D.

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

Institutional Affiliation: [all] University of Washington

Objective: To develop a clinically relevant measure of sexual risk behavior for use in clinical care using mixed qualitative and quantitative methods. The qualitative methods are described here.
Methods: We identified validated sexual risk measures designed for use in clinical care and placed them into an item pool. Three reviewers with expertise in item development winnowed item pool content using PROMIS Qualitative Item Review (QIR) criteria. Item pool content areas informed development of an interview guide for exploring how patients in primary care conceptualize sexual risk behavior. We conducted in-depth interviews with 89 patients in Birmingham, AL, Boston, MA, San Diego, CA, and Seattle, WA. 80 were patients living with HIV, 9 were HIV-uninfected. One-third of patients were cis-gender females; 7% were transgender females. 51% were men who have sex with men (MSM); of these, 24% had female partners. Sexually active patients with 2+ partners in the past six months were eligible; for female and transgender patients, eligibility criteria were expanded to fewer partners and a 1-year time frame. Of the 89, equal thirds were considered high, moderate, or low risk based on responses to condom use behavior items. We used open-coding to categorize item pool content into thematic areas: knowledge, attitudes/beliefs, sexual behavior, prevention practices, loss of control (of sexual behavior), and sex as a means to an end. We used Dedoose software to create excerpts from transcribed interviews. Two researchers coded the excerpts to ensure inter-rater reliability, reconciling codes when differences existed. Coded excerpts were then matched to item pool content. Unmatched excerpts were analyzed for item development potential; excerpts containing truly new concepts were developed into items by the research team and refined or omitted using the QIR process. Provider input (n=7) from multiple sites regarding clinical relevance of instrument content further winnowed the items. We then conducted two rounds of cognitive interviews with patients (n=37) to assess comprehension and interpretation of items.
Results: An algorithmic measure with a maximum of 10 items was developed, with skip patterns that maximize personal relevance and reduce patient burden. Areas identified for potential new item development included perception of partner viral load, use of PrEP, gender of partner with categories to include transgender, partner type (casual vs. primary), concurrent sexual partner and group sex behavior, direction of anal sex (insertive vs. receptive), context of meeting partners, reasons for sexual risk behavior, perception of partner fidelity, and concern over STI exposure. Items deemed by providers as most clinically relevant were primarily behavior-based. Cognitive interviews revealed the items to be well-understood and tolerated by patients, with a need for more inclusive language for transgender patients.
Conclusions: A brief, clinically relevant sexual risk measure is warranted in order to maximize personal relevance and reduce patient burden. This measure should be meaningful across gender and sexual orientation categories; indicative of clinically relevant sexual decision-making factors such as perceived viral load; and should employ skip patterns based on endorsement of gender category, sexual orientation, number of partners, HIV serostatus, and HIV partner status.

Assessing the relationship between sociodemographics and the interpretation of Spanish PROMIS® Physical Function items

Denise C. Fyffe, PhD [1], Helena Correia, BA [2], Sylvia Paz, PhD [3], Susan Magasi, PhD [4], Alan Jette, PT, PhD [5] David S. Tulsky, PhD [6]

Presenting Author: **Denise C. Fyffe, PhD;** dfyffe@kesslerfoundation.org

Institution al Affiliations [1] Kessler Foundation [2] Northwestern University [3] UCLA [4] University of Illinois at Chicago [5] Boston University [6] University of Delaware

Objective: The PROMIS® Physical Function item bank was translated into Spanish using a standard linguistic validation methodology and a universal approach to translation (i.e., one Spanish version appropriate for multiple countries). Psychometric analysis based on the PROMIS Spanish calibration sample, revealed that 50 of 114 items exhibited language differential item functioning (DIF). The goal of this qualitative pilot study, is to ascertain whether responses to the Spanish PROMIS® Physical Function items are associated with sociodemographic characteristics of Spanish- speaking Hispanic/Latino respondents.

Methods: Cognitive interviews were conducted with Spanish-speakers from either Mexico or Puerto Rico, residing in New Jersey, who self- identified as Hispanic/Latino. The sample was stratified by age and educational attainment. Cognitive interviews included standard administration of sets of PROMIS® Physical Function items followed by cognitive probes for selected items, using the think-aloud method. Particular attention was given to items that had presented DIF in the calibration study.

Results: Twenty-five cognitive interviews were conducted with participants who self-identified as Hispanic/Latino. Participants averaged 43 years of age (SD = 17.7), 69.2% were women and 42.2% had less than high school education. Physical Function scores were significantly higher among younger participants in comparison with older participants. Variation in response patterns and item interpretation were observed across age and education level.

Conclusions: Preliminary qualitative analyses suggest an association between PROMIS Physical Function responses and item interpretation, and key demographic characteristics among Spanish-speaking Hispanics/Latinos. Future qualitative research comparing English-speaking and non-Hispanic samples in the US will provide additional insights about cultural and linguistic factors that may affect response differences observed among Spanish-speaking Hispanic/Latinos to the PROMIS Physical health scores.

Using PROMIS® Instruments to Evaluate a Community Based Program to Empower Women in Transition

Frances B Garfield, Ph.D; Mary Jaynes, BA, MA LMHC

Presenting Author: **Frances B Garfield, PhD**; fbg@zeus.bwh.harvard.edu

Institutional Affiliation: WE CAN

Objective: To measure changes in selected PROMIS® scales during the WE CAN PathMaker program to determine whether the instruments are appropriate measures to assess interventions in the community.

Methods: The Pathmaker program matches women in transition to volunteer mentors. They meet individually one or more times a month from September to May and attend monthly workshops together. Initial program evaluation consisted of a qualitative study and a pilot study. The qualitative and pilot findings led to the addition of six PROMIS® scales to the self report instruments given to the PathMakers in the current program. The

scales were administered during the PathMakers' first meeting in August before the matchup of PathMakers and mentors. The scales were also given to program participants during their January meeting, which was halfway through the program. Demographic and socioeconomic variables were collected at intake.
Results: The 19 women in the program had an average age of 47.3, (range = 32-68) and 15 years of schooling. The women's financial situations were more diverse. 63% were employed or self-employed, 21%were unemployed or underemployed, and 11% were retired or disabled. Almost 32% owned their own home or apartment, 32% rented, and 21% lived with friends or family. Before the start of the program, the average scores for the scales were: Perceived Stress 33.8, Emotional Support 27.5, Perceived Rejection SF18.6, Self Efficacy 32.1, General Life Satisfaction 34.1, Meaning and Purpose SF 57.7. Halfway through the program, average ratings on the instruments were: An analysis of variance found that changes in the participants' ratings on the scales were /were not significant.
Conclusions: The PROMIS® instruments are appropriate/not appropriate for assessing interventions in community based programs.

Impact of Unintended Pregnancy on Women's Quality of Life

Aileen Gariepy, MD, Eleanor Schwarz, MD, Lisbet Lundsberg, PhD

Presenting Author: **Aileen Gariepy, MD, MPH**, aileen.gariepy@yale.edu

Institutional Affiliations: Dept of Obs, Gyn, & Reproductive Sciences, Yale University

Objective: It is estimated that 51% of US pregnancies are unintended. We sought to compare quality of life scores for women with intended and unintended pregnancies.
Methods: We recruited women newly diagnosed as pregnant from a walk-in pregnancy testing clinic for a prospective, mixed methods study. Women were eligible if aged 15 to 44, fluent in English or Spanish, and were pregnant at less than 20 weeks gestational age. Participants completed a self- administered PROMIS Global Short Form as well as multiple measures of pregnancy intention, wantedness, and timing. PROMIS mean T-scores were calculated and compared to measures of pregnancy intention using the Wilcoxon test.
Results: We enrolled 87 participants from June 2014 through January 2015 at an urban, Medicaid- funded clinic. Women were young (mean age 26.5 years, range 18-44 years) with first trimester pregnancies (mean gestational age 8 weeks, SD 3 weeks). Most participants self-identified as Hispanic (57.6%) or Black (34.1%), English or Spanish speaking (57% and 43% respectively), most were unmarried (79.1%), and had at least one child already (82%). Few (31%) had more than a high school education (69%), and 40% reported working outside the home (full-time or part-time). An "intended" pregnancy was reported by 43% of participants; 49% said they "wanted to have a baby," 55% reported that they had become pregnant at the "right time," 78% reported the pregnancy was "desired," and 85% reported they felt "happy" when they found out about pregnancy. Compared to women who intended pregnancy, those who did not had significantly lower Global Mental Health T-scores (49.4 vs. 53.6, p=0.04), but similar physical health scores. Compared to women who "wanted to have a baby," those who did not had lower Global Mental (48.4 vs. 54.1, p=0.0075) and Physical Health T-scores (46.2 vs. 52.1, p=0.0065). Compared to women responding that the timing of this pregnancy was "the right time," those saying this was "ok, but not right time" or "wrong time," had significantly lower mental (47.7 vs. 54.1, p=0.0023) and physical health T-scores (45.1 vs. 52.4, p=0.0004). Compared to women who reported pregnancy was "desired," those who did not had significantly lower mental (45.4 vs. 52.9,

p=0.0022) and physical health T-scores (44.6 vs 50.4, p=0.0245). Compared to women who were "somewhat" or "very" happy to learn of pregnancy, those who were not had significantly lower mental (44.6 vs. 52.4, p=0.0024) and physical health T- scores (47.7 vs. 54.1, p=0.0023). Measures of pregnancy wantedness, timing, happiness, and desirability were significantly correlated with PROMIS Global Mental (r=0.30, r=0.34, r=0.30, r=0.33, respectively) and Physical Health scores(r=0.34, r=0.42, r=0.24, r=0.27, respectively). Although pregnancy intention (r=0.22, p=0.15) was correlated with mental health scores, it was not correlated with physical health scores.

Conclusions: Women's feelings about their pregnancy affects their quality of life, as measured by PROMIS Global Mental and Physical Health scores. PROMIS Global Mental and Physical Health scores were significantly correlated with measures of pregnancy wantedness, timing, happiness, and desirability. Although pregnancy intention was correlated with mental health scores, it was not correlated with physical health scores.

Clinical Validation of PROMIS Global Short Form in Pregnancy

Aileen Gariepy, MD, Eleanor Schwarz, MD, Lisbet Lundsberg, PhD

Presenting Author: **Aileen Gariepy, MD, MPH**, aileen.gariepy@yale.edu

Institutional Affiliations: Dept of Ob, Gyn, & Repro Sciences, Yale University

Objective: Pregnancy is a unique health state that may significantly affect women's quality of life, in both positive and negative ways. Despite being a transient health state, approximately 3.3 million US women are pregnant at any given time. Given this prevalence, robust measurement tools for determining quality of life during pregnancy are needed. We sought to compare the PROMIS Global Short Form to a legacy measure (SF-12® Health Survey) for evaluating quality of life in pregnant women.

Methods: We recruited women newly diagnosed as pregnant from a walk-in pregnancy testing clinic for a prospective, mixed methods study. Women were eligible if aged 15 to 44 years, fluent in English or Spanish, and were pregnant at less than 20 weeks gestational age. Participants completed a self-administered PROMIS Global Short Form, a legacy measurement tool (SF-12® Health Survey), and a validated pregnancy-specific screening tool, the Edinburgh Depression Scale (EDS). We examined the correlations between PROMIS Global Mental Health Score and Results:We enrolled 87 participants from June 2014 through January 2015 at an urban, Medicaid-funded clinic. Women were young (mean age 26.5 years, range 18-44 years) and most (92%) were in the first trimester of pregnancy (mean gestational age 8 weeks, SD 3 weeks, range 3-18 weeks). Most participants self-identified as Hispanic (57.6%) or Black (34.1%), English or Spanish speaking (57% and 43%, respectively), most were unmarried (79.1%), and had at least one child already (82%). Few (31%) had more than a high school education, and 40% reported working outside the home (full-time or part-time).An "intended" pregnancy was reported by 43% of participants; 49% said they "wanted to have a baby," 55% reported that they had become pregnant at the "right time," 78% reported the pregnancy was "desired," and 85% reported they felt "happy" when they found out about pregnancy. Compared to women who intended pregnancy, those who did not had significantly lower Global Mental Health T-scores (49.4 vs. 53.6, p=0.04), but similar physical health scores. Compared to women who "wanted to have a baby," those who did not had lower Global Mental (48.4 vs. 54.1, p=0.0075) and Physical Health T-scores (46.2 vs. 52.1, p=0.0065). Compared to women

responding that the timing of this pregnancy was “the right time,” those saying this was “ok, but not right time” or “wrong time,” had significantly lower mental (47.7 vs. 54.1, p=0.0023) and physical health T-scores (45.1 vs. 52.4, p=0.0004). Compared to women who reported pregnancy was “desired,” those who did not had significantly lower mental (45.4 vs. 52.9, p=0.0022) and physical health T-scores (44.6 vs 50.4, p=0.0245). Compared to women who were “somewhat” or “very” happy to learn of pregnancy, those who were not had significantly lower mental (44.6 vs. 52.4, p=0.0024) and physical health T- scores (47.7 vs. 54.1, p=0.0023). Measures of pregnancy wantedness, timing, happiness, and desirability were significantly correlated with PROMIS Global Mental (r=0.30, r=0.34, r=0.30, r=0.33, respectively) and Physical Health scores(r=0.34, r=0.42, r=0.24, r=0.27, respectively). Although pregnancy intention (r=0.22, p=0.15) was correlated with mental health scores, it was not correlated with physical health scores.

Implementing a PROMIS to Improve the Quality of Life in Pediatric Patients with CNS Tumors

Mariko DeWire, MD, Maureen Gallagher, NP, Elizabeth Gilger, NP, Julie Plummer, BS, Maureen Donnelly, LISW

Presenting Author: **Elizabeth Gilger, NP**; Elizabeth.Gilger@cchmc.org

Institutional Affiliations: Cincinnati Children’s Hospital Medical Center, Cincinnati

Objective: While science and protocol directed therapies continue to search for cures for neuro-oncology patients, work must also continue in enhancing patient quality of life throughout their treatment trajectory. The objective of this pilot study is to evaluate the feasibility of measuring patient reported outcomes in pediatric and young adults diagnosed with brain tumors receiving new therapies.

Methods: Eligible patients are identified during scheduled weekly clinical planning meeting which is attended by clinical and research staff. Potential patients are approached to participate on a separate day following the development of the treatment plan and signing of the treatment and survey consents. This initial interview with the PROMIS questionnaire by study staff, administered on paper, computer, or tablet (iPAD), occurs either during a planned clinic visit within 10 days of proceeding with surgery, starting chemotherapy or radiotherapy. Evaluations occur at baseline and then every 3 months for up to two years. Data collected through the implementation of this toolkit is reviewed in “real time” by the study team. Additionally, they are discussed during weekly psychosocial patient reviews by a multidisciplinary team comprised of physicians, social workers, nurse practitioners, nursing, psychology, school intervention specialists, chaplains, child life specialists, and representatives from palliative care. The T-score of PROMIS measure is analyzed using the NIH PROMISwebsite https://www.assessmentcenter.net/Manuals.aspx.

Results: Five patients currently are enrolled. Primary diagnosis of all patients portended poor outcomes or increased morbidity due to disease. Ages of patients range from 6 years to 25 years of age. Evaluation and implementation of multidisciplinary interventions was feasible. However, the variety of questionnaire modalities have brought up unexpected yet, valuable information. Moreover, ttunanticipated distress in the parents and not as prevalent in the patient was noted when reviewing the question “do you think you will die?” in the pediatric questionnaire. Due to patient and family relocation, school related question were somewhat difficult for parents to answer when taken at face

value.
Conclusions: The unique collaboration of clinical research and multidisciplinary patient care has made it possible to further develop the current evaluation practice of neuro-oncology patients. By creating a process to administer the designated PROMIS measures to children, their parents, and adult patients, Neuro-oncology care providers have the opportunity to systematically assess and improve the patient's quality of life in actual time.

The reliability of Pain Intensity and Pain Interference (Polish version) PROMIS short forms: a preliminary report

Glinkowski, Wojciech [1,2,3], Żukowska Agnieszka[1,2]

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

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

Objective: This pilot study was designed to assess the reliability of two short forms of PROMIS item banks – Pain Intensity SF 3a and Pain Interference SF6 in Polish language translation.
Methods: Patients are suffering musculoskeletal disorders were randomly recruited for this study. The patients were diagnosed mostly with osteoarthritis or rheumatoid arthritis. The study group consisted of 13 women and two men. Six patients were after total hip or knee replacement; five patients were awaiting total joint replacement surgery. Two patients had metal implants removed, and two were after revision surgery of joint replacements. Short forms of Pain Intensity and Pain Interference in Polish language were completed by 15 adults aged 34-83. The group received "paper and pencil" versions of Pain Intensity and Pain Interference short forms to complete both short forms day after day. Internal consistency was tested by Cronbach's alpha; test-retest reliability was tested by the intra-class correlation coefficient (ICC).
Results: The results showed almost perfect reliability of assessed Items of this specific population. The Cronbach's alpha with standardized variables was 1,00 and the 95% lower confidence limit was 1.00 for all items for preoperative cases. Similarly, Intraclass Correlation Coefficient was calculated 1,00 for this groupPostoperative cases results:
The consistency assessed for the same raters for all subjects, two-way model with a mean of intraclass correlation (95% Confidence Interval) was 0,96 for single measures and 0,98 for average measures.
Conclusions: The questionnaire may be used as a valid and reliable measure for assessing pain intensity and pain interference in Polish patients. This pilot study shows that all items are internally consistent. Further study will be continued in order to collect a larger study group and to perform the statistical tests to assess the validity, reliability and consistency of Polish versions of other PROMIS Item Banks.

Health-Related Quality of Life in Childhood Cancer Survivors at Risk for Cardiovascular Late Effects

Rebecca Williamson, MPH [1], Shannon L Golden, MA [2], Ann C. Mertens, PhD [1,3], Janet A. Tooze, PhD, MPH [4], Nancy Avis, PhD [4], Sharon M. Castellino, MD, MSc [2,4]

Presenting Author: **Shannon L. Golden, MA**; sgolden@wakehealth.edu

Institutional Affiliations: [1] Aflac Cancer and Blood Disorders Center at Children's Healthcare of Atlanta, Atlanta, GA [2] Department of Pediatrics at Wake Forest School of Medicine, Winston-Salem, NC [3] Department of Pediatrics at Emory University, Atlanta, GA [4] Division of Public Health Sciences at Wake Forest School of Medicine, Winston-Salem, NC

Objective: Due to advances in treatment, the overall survival rate for childhood cancer exceeds 80%. However, childhood cancer and its subsequent treatment predispose survivors to a higher risk of certain life- threatening and debilitating conditions called late effects. By age 45, it is estimated 80.5% of survivors will have experienced a serious/disabling or life-threatening late effect. Specifically, research shows that patients treated with anthracycline therapy and/or thoracic radiation are at an increased risk of cardiovascular late effects, such as heart valve disorders, congestive heart failure, and cardiomyopathy. The purpose of this analysis is to characterize health-related quality of life (HRQoL) among young adult survivors of childhood cancer who are at risk for cardiovascular late effects but have preserved left ventricle (LV) ejection fraction and are asymptomatic with regards to cardiovascular symptoms.

Methods: Pediatric cancer survivors currently aged 18-40 years who received anthracycline therapy with or without thoracic radiation and were ≤20 years off therapy and age-, sex-, BMI, and race/ethnicity-matched healthy controls were recruited to participate in the Resonance Imaging Trial for Heart Biomarkers in Adolescent/Young Adult Cancer Survivors (RITHM). All enrolled participants completed the PROMIS® Global short form. Global physical and mental health scores were computed and converted to US general population standardized t-scores. Comparisons between survivors and controls were made using Fisher's exact and t-tests. Survivors with global scores >1 standard deviation below the mean of the healthy matched controls were classified as poor health status. Comparisons among survivors of poor health status were made using Fisher's exact methods.

Results: Overall, 53 cancer survivors and 19 healthy controls completed the PROMIS Global short form at a mean (±SD) age of 22.6±3.5 years. Survivors were a median of 7.0 years from the end of all cancer treatment. There were no differences in education, prevalence of smoking, or physical activity between survivors and controls. Overall, controls reported high HRQoL: mean (±SD) physical t-score: 58.5±5.2 and mean mental t-score: 58.5±6.4. Compared to controls, cancer survivors reported significantly poorer physical (53.9±7.8; p=0.02) and mental health (53.2±8.4; p=0.02). Among survivors, 43.4% reported poor physical health, with a higher proportion of those diagnosed at age ≥10 years having poor physical HRQoL (50.0% vs. 18.2%, p=0.09). Poor mental health was noted in 37.7% of survivors, with females (54.2% vs. 24.1%, p=0.05) and those treated with both thoracic radiation and anthracycline therapy (55.0% vs. 27.3%, p=0.08) more likely classified as poor.

Conclusions: Healthy young adult controls report high HRQoL. Compared to matched controls, young adult pediatric cancer survivors at risk for cardiovascular morbidity who are within the first 15 years following treatment report lower physical and mental HRQoL. The PROMIS Global short form can be integrated into research with young adult cancer survivors. However, until age-adjusted norms are available for the young adult population, age- matched controls should be used in HRQoL analyses applying the PROMIS Global short forms.

New Patient-Reported Measures of Contextual Factors in Healing

Carol M. Greco, Ph.D., Lan Yu, Ph.D., Mary Lou Klem, Ph.D., MLIS, Nathan Dodds, B.S., Kelly Johnston, MPH., Christine McFarland,

B.A., Suzanne Lawrence, M.A., Jason Colditz, M.S., Catherine J. Mailhoefer, M.A., Natalia E. Morone, MD, MSc, Ronald Glick, MD, Michael J. Schneider, DC, Ph.D., Neal D.Ryan MD, Paul Pilkonis, Ph.D.

Presenting Author: **Carol M. Greco,** PhD; grecocm@upmc.edu

Institutional Affiliations: University of Pittsburgh

Objective: In healthcare, nonspecific factors such the patient provider relationship, treatment expectations, and patients' attitudes are important elements of the treatment experience. Such contextual features of care can influence treatment outcomes. However, no precise, concise, and highly generalizable instruments to measure these factors exist. We report on the development, calibration, and initial validation of a set of item banks, titled the Healing Encounters and Attitudes Lists (HEAL), which assess nonspecific factors in treatment from the patient's perspective.

Methods: We used PROMIS® methodology to develop patient report measures of nonspecific factors that may influence healing, including both environmental, contextual factors such as patient-provider relationship and treatment expectations, and intrapersonal influences (e.g., optimism, self-efficacy, spirituality). Model development was informed by themes identified in 6 patient focus groups and interviews with 22 clinicians, as well as literature review. Comprehensive literature searches using 9 databases yielded over 500 existing instruments and resulted in an initial item pool of several thousand items. After iterative qualitative item analysis, including cognitive interviewing, 296 items were included in field testing. The calibration sample included 1657 respondents, 1400 from the general population (obtained through an internet panel) and 257 from local conventional medicine and integrative medicine clinics. Data analysis included exploratory and confirmatory factor analysis (EFA, CFA), and item response theory (IRT). To explore initial validity, 6 weeks after completing the item banks 221/257 of the local patient sample rated their Clinical Global Impression (CGI) of change, completed legacy questionnaires and the PROMIS29.

Results: EFA and CFA retained 250 items which were further refined to 168 items based on IRT analyses. The final HEAL item banks are: Patient-Provider Connection (57 items), perceptions of the Healthcare Environment (25 items), Treatment Expectancy (27 items), Positive Outlook (27 items), and Spirituality (26 items). In addition, a 6-item short form, Attitudes toward Complementary and Alternative Medicine (CAM), was retained from a heterogeneous group of Health and Wellness Attitudes items. Short forms (6 – 7 items each) were also developed from each item bank, in case CAT administration is not feasible. Patients' IRT theta scores for all factors except Spirituality were significantly but modestly correlated with CGI, with Treatment Expectations showing the highest correlation with CGI (Spearman's Rho=.358, $p<0.01$). HEAL thetas were modestly associated with legacy instruments assessing similar constructs. Further validation studies of HEAL in patients receiving a new treatment for chronic back and/or neck pain are currently underway.

Conclusions: The project has successfully created item banks to measure several nonspecific factors in healing. The banks are applicable across a broad range of treatments and clinical conditions. Test information curves indicate that the HEAL item banks provide substantial information across a broad range of each construct. HEAL item banks and the CAM attitudes short form showed initial evidence of predictive and concurrent validity, suggesting that they are suitable for measuring nonspecific factors that can influence treatment outcome.

Calibration and Validation of the PROMIS® Self-Efficacy for Managing Chronic Conditions Measures

Ann L. Gruber-Baldini, PhD [1], Craig Velozo, PhD, OTR [2], Sergio Romero, PhD [3], Lisa M. Shulman, MD [])

Presenting Author: **Ann L. Gruber-Baldini**; abaldin@epi.umaryland.edu

Institutional Affiliations: [1] University of Maryland School of Medicine, Baltimore, MD. [2] Medical University of South Carolina, Charleston, SC [3] University of Florida, Gainesville, FL

Objective: Calibration of the PROMIS® Self-efficacy for Managing Chronic Conditions Measures.

Methods: The PROMIS® Self-efficacy for Managing Chronic Conditions item pool was created to include 5 subdomains with a total of 143 items. The subdomains are: Self-Efficacy for: Managing Daily Activities (36 items), Managing Medications and Treatments (27 items), Managing Symptoms (28 items), Managing Emotions (28 items), and Managing Social Interactions (24 items). The pool was calibrated in 1087 subjects from 2 data sources: 837 patients seen at the University of Maryland Neurology Ambulatory Center, with epilepsy (n=171), multiple sclerosis (MS, n=166), neuropathy (n=163), Parkinson disease (PD, n=170), and stroke (n=167); and in 250 subjects recruited from an online internet sample of adults with general chronic conditions. Scores were compared with one legacy scale: the Stanford Self-Efficacy for Managing Chronic Disease 6-Item Scale and with 5 PROMIS® short forms: Global Health (Physical and Mental), Physical Function, Fatigue, Depression, and Anxiety. Physician-rated severity (no, mild, moderate, severe) and disability scales (Modified Rankin, Barthel Index) were available for the neurologic center samples only.

Results: The sample was 57% female, mean age=53.8 (SD=14.7, range 18-89), 76% white, 21% African American, 3% other race, with 6% Hispanic and 76% having higher than high school education. Full item banks, as well as 8-item and 4-item short forms were created for the 5 subdomains. Only 3 items were eliminated for lack of fit or local dependence (1 in Managing Medications and Treatment, 1 in Managing Emotions, 1 in Managing Social Interactions). All measures had good internal consistency (Chronbach's alpha>0.95 for full banks, and alpha>0.85 for short forms) and correlated well with the Stanford Self-efficacy legacy measure (ranging from r=0.56 for Medications and Treatments to r=0.75 for Symptoms). Significant correlations were seen between the Self-efficacy measures and other PROMIS short forms (all r>0.38), with the highest correlations between Self-efficacy for Managing Daily Activities with Physical Functioning (r=0.78), and Self-efficacy for Managing Emotions with General Mental Health (r=0.70). For the other self-efficacy scales, the highest correlations were with General Mental Health (Managing Medications and Treatments r=0.52, Managing Social Interactions r=0.63, and Managing Symptoms r=0.63). Among the neurologic samples, physician-rated disease severity and disability (Barthel Index, Modified Rankin) showed the highest correlations with Self-efficacy for Managing Daily Activities (severity r=-0.43; Rankin r=-0.47; Barthel r=0.46), and the lowest correlations with Self-efficacy for Managing Emotions (severity r=-0.19; Rankin r=-0.18; Barthel r=0.12) and with Managing Medications (severity r=-0.16; Rankin r=-0.19; Barthel r=0.22).

Conclusions: The newly developed PROMIS® Self-efficacy for Managing Chronic Conditions measures include 5 scales that have been calibrated across diverse chronic conditions and show good internal consistency and cross-sectional validity.

Health Literacy and Patient-Reported Outcomes In Underserved English- and Spanish-Speaking Diabetes Patients

Elizabeth A. Hahn, MA [1], James L. Burns, MS [1], Elizabeth A. Jacobs, MD [2], Pamela S. Ganschow, MD [3], Sofia F. Garcia, PhD [1], Joshua P. Rutsohn [1], David W. Baker [1]

Presenting Author: **Elizabeth A. Hahn**, MA; e-hahn@northwestern.edu

Institutional Affiliations: [1] Northwestern University Feinberg School of Medicine, Chicago, IL [2] University of Wisconsin School of Medicine and Public Health, Madison, WI [3] Rush University Medical Center, Chicago, IL

Objective: To examine associations between patient characteristics, health behaviors and health outcomes, and to explore the role of health literacy as a potential mediator of diabetes outcomes.
Methods: Convenience sampling was used to enroll English- and Spanish-speaking adults with type 2 diabetes being treated with oral medication or insulin. Participants used a bilingual multimedia Talking Touchscreen (TT) / Pantalla Parlanchina (PP) kiosk to complete the following patient-reported outcomes (PROs) instruments in a safety net general medicine clinic: diabetes knowledge, health beliefs, self-efficacy, diabetes self-care, decision-making preference, satisfaction with communication, health literacy, and health status (PROMIS Global Health instrument). Dependent variables were diabetes self-care behaviors, health status and satisfaction with communication. Independent variables included sociodemographic and clinical characteristics, health literacy, health beliefs and self-efficacy. The Behavioral Model for Vulnerable Populations guided multivariable regression and mediation testing.
Results: English-speaking (n=146) and Spanish-speaking (n=149) groups differed on many characteristics. The Spanish-speaking group had more women, lower education, fewer with health insurance, more homemakers, and less prior computer experience. The majority of the Spanish-speaking participants (84%) specified their ethnicity as Mexican, Mexican-American or Chicano. The majority of the English speaking group was non-Hispanic Black (65%). Many participants felt their income was not adequate to meet their needs, and that their physical condition or medical treatment caused them financial difficulties. Spanish-speaking participants had lower health literacy, and poorer physical, mental and overall health, compared to English-speaking participants. Health literacy was negatively associated with social support for diet, diet and medication barriers, and age, and positively associated with BMI, knowledge, and talking with health care professionals to get diabetes information. Health literacy was not associated with diabetes self-care, health status and satisfaction with communication, and it did not mediate the effects of other factors on these outcomes. Diabetes self-efficacy was significantly associated with health behaviors and outcomes. Even though many study participants had never used a computer before, they were able to complete multiple questionnaires on their own using the TT/PP.
Conclusions: In contrast to expectations, health literacy was not associated with diabetes self-care, health status and satisfaction with communication. The association between Spanish language preference and poorer health was not mediated by this group's lower health literacy. Increasing health-related self-efficacy might be an important clinical strategy for improving outcomes in underserved patients with type 2 diabetes.

Comparison of PROMIS-29 and EQ-5D-3L Quality-Adjusted Life Years

John D. Hartman, MA, Benjamin M. Craig, PhD

Presenting Author: **John D. Hartman, MA**; john.hartman@moffitt.org

Institutional Affiliations: [all] Moffitt Cancer Center and University of South Florida, Tampa, FL

Objective: Using quality-adjusted life years (QALYs) based on patient reported outcomes (PRO) as a summary measure of health-related quality of life (HRQoL) is an essential component to any economic evaluation comparing treatment outcomes. While multiple papers have compared across PRO instruments, no paper has compared the PROMIS and EQ-5D QALYs.
Methods: As part of a nationally representative survey, 2,623 US adults completed the PROMIS-29 (PROMIS) and the 3 level version of the EQ-5D (EQ-5D-3L). Their responses were summarized on a QALY scale using published estimates and compared with each other as well as with general self-reported health. Regression analysis was used to map between PROMIS and EQ-5D QALYs.
Results: The findings confirm that the PROMIS QALYs are much lower than the EQ-5D QALYs. Compared to the EQ-5D, PROMIS QALYs are more continuously distributed (i.e., fewer gaps) and more correlated with self-reported health on either a visual analogue scale (Pearsons: .625 vs. .600) or 5-point scale (log-likelihood: -3201.78 vs -3259.02). Specifically, the respondents reported 1670 and 74 distinct QALY values for the PROMIS and EQ-5D, respectively; therefore, the PROMIS QALYs map well to EQ-5D QALYs, but the EQ-5D QALYs poorly map to the PROMIS QALY. This is largely because 43% of respondents reported no problems on the 5 EQ-5D domains (compared to 4% for the 7 PROMIS-29 domains).
Conclusions: While both instruments are widely used, PROMIS QALYs have a stronger relationship with self-reported health than the EQ-5D QALYs and can predict EQ-5D QALYs (not vice versa). Future research is needed to compare the PROMIS QALYs to the recently released 5 level version of the EQ-5D.

Acceptance of PROMs in eliciting and measuring progress toward goals: A qualitative study

Maureen Henry, JD, PhD; Erin Giovanetti, PhD; Jessica Briefer French, MHSA; Erica Anderson, MA; Judy Ng, PhD.

Presenting Author: **Maureen Henry,** henry@ncqa.org

Institutional Affiliation: NCQA

Objective: To evaluate patient and care manager perspectives on the use of PROMIS-29 and the COLLAGE Lifestyle Survey 1) to identify areas of concern to address with providers, 2) to help identify targets for individuals' goals, and 3) as patient-reported outcome measures (PROMS). The study also evaluated reactions to the PROMIS-29 including level of difficulty and missing domains.
Methods: Semi-structured interviews were conducted with 6 care managers, 4 care management teams, and eleven individuals with long-term service and support (LTSS) needs. A low-inference coding template based on the questionnaire and study aims was developed and was systematically applied to the transcripts by two researchers; open coding was used to capture additional relevant data not addressed in the template development. Themes were identified.
Results: *Utility:* Identifying individuals' needs and targets for individuals' goals: Care managers and individuals viewed an interactive conversation as superior to

completing a structured instrument as a mechanism for identifying individuals' needs and eliciting individuals' goals, although care managers expressed some enthusiasm for using a structured questionnaire to identify domains that might be missed in a conversation. *Tracking individuals' responses over time:* With prompting, both care managers and individuals acknowledged the value in tracking individuals' responses over time as a measure of moving toward or away from goals. Some care managers and individuals noted that comparing scores over time could be discouraging, and that scores might be more reflective of outside forces than of the quality of interventions.
Population measurement: No care managers spontaneously identified population measurement as a use for scores on either instrument. When prompted, however, many agreed that an instrument could be used to measure the wellbeing or progress toward goals of the population served by the organization. *Level of Difficulty:* Judgments about the difficulty to complete the PROMIS-29 differed across populations. Physical, intellectual, and sensory disabilities, dementia, and serious mental illness were cited by care managers as conditions that would increase the difficulty. Some care managers suggested that challenges could be ameliorated if they could read and/or explain the questions or fill the form out for the person. Both care managers and individuals noted that selecting one number on a Likert scale can be challenging. When asked directly, individuals said the form was not difficult. *Missing Domains*: Among both individuals and care managers, there was a desire to address issues relevant to the individual with more specificity. Both expressed the desire for questions focusing on the positive in the individuals' life, such as what is going well. Finally, both groups identified the need for more questions about the nature and quality of the individuals' relationships with others.
Conclusions: Care managers and individuals acknowledged potential in using a structured instrument like the PROMIS-29 to measure change over time, but conversation was favored as the mechanism to identifying individuals' needs and goals. Care managers also saw the value in using PROMs to assess wellbeing and progress toward goals at the population level. Difficulties with completing the form would need to be addressed.

Collecting Outcomes in a High-throughput Outpatient Clinic: Significance and Practicality of PROMIS CAT Implementation

Patrick J. Hesketh, BS, Evan R. Bannister, BS, Taylor M. Camm, Annamarie D. Horan, PhD, and Samir Mehta, MD

Presenting Author: **Patrick J. Hesketh, BS**; Patrick.Hesketh@uphs.upenn.edu

Institutional Affiliations: University of Pennsylvania Department of Orthopaedic Surgery

Objective: The collection of valid and informative patient reported outcomes is a growing part of modern healthcare infrastructure. Our goals were to assess the feasibility of acquiring 4 PROMIS CAT instruments (Depression, Pain Behavior, Pain interference and Physical Function) in a busy Orthopaedic Trauma Clinic and correlate the results with age and injury criteria.
Methods: Facility WiFi upgrade was required and performed to ensure adequate access the PROMIS CAT interface and 2 iPads were purchased and loaded with appropriate applications for the initiative. A dedicated research assistant was hired and deployed to the Penn Orthopaedic Surgery Trauma Outpatient Clinic to capture PROMIS CAT instruments in a manner that did not interfere with clinic flow. PDFs of scoring results were temporarily saved to the iPad and transferred to secure servers. Scores

and patient demographics were captured into a RedCap case report form (n = 494). Statistical analysis was performed using IBM SPSS Version 20 on participants who completed all four PROMIS CAT instruments (n = 403). Non-parametric tests were used because the data did not meet criteria for ANOVA.
Results: Subjects were approached cross-sectionally in the outpatient clinics for 4 Orthopaedic Surgeons. Subject Age was broken down into 7 groups (Mean age = 47 ± 16 years). Scores for Pain Behavior, Pain Interference, and Physical Function were significantly different across age groups ($p<<0.01$). There was a significant difference in Physical Function scores across body regions ($p < 0.05$) with better scores observed for the Upper Extremities, Hand and Wrist than for the lower half of the body. All participant scores were significantly different compared to the general population, age, and gender relative to PROMIS CAT normative data ($p << 0.01$)
Conclusions: We were able to capture 4 complete PROMIS CAT instruments from 82% of participants in a busy Orthopaedic Trauma clinic without disruption in clinic flow. That the Physical Function Scores were significantly worse for lower body region injuries than upper body injuries was consistent with existing literature. Given that Orthopaedic Trauma patients enrolled cross-sectionally are much worse for all scores compared to the general population, participant age, and participant gender, additional longitudinal study is required to determine whether PROMIS CAT scores eventually normalize through the course of healing. The PROMIS CAT system is potentially very useful to discern the need for additional services when clinical and radiologic findings are inconsistent with patient reported outcomes.

Challenges and lessons learned in developing a patient-reported outcomes clinical research program
Man Hung, PhD

Presenting Author: **Man Hung, PhD**; man.hung@hsc.utah.edu

Institutional Affiliation: University of Utah, Salt Lake City, UT

Objective: Collecting large-scale patient-reported outcomes (PRO) data for clinical care and research is a fundamental step toward improving treatment outcomes for patients. This presentation describes the challenges and lessons learned in developing a prospective PRO data collection infrastructure and in conducting PRO research during the past six years at a tertiary orthopaedic center in the United States.
Methods: Prospective PRO data collection was started in 2009 using the Patient-Reported Outcomes Information System (PROMIS) instruments in addition to some existing instruments. These instruments were completed by patients on clinic visits and more recently prior to clinic visits via automatic email reminders.
Results: Initially, there was a large amount of resistance to incorporate the PROMIS tools in the clinical settings. Today, this resistance has been dramatically decreased resulting from the ability to remove certain barriers and to demonstrate value and efficiency gain.
Conclusions: For a PRO clinical research program to be successful, it needs to provide culture, leadership, analytic and information technology competencies. Various stakeholders must be involved.

Evaluation of Responsiveness of 8 PROMIS Measures in a Community- Based Sample of Cancer Patients

Roxanne E. Jensen [1], Arnold L. Potosky [1], Tania Lobo [1], Donald Stull [2], Ron D. Hays [3], David Eton [4], Ashley Wilder Smith [5], Elizabeth Hahn, [6] Carol M. Moinpour [7]

Presenting Author **Roxanne E. Jensen**; rj222@georgetown.edu

Institutional Affiliations:[1] Georgetown University, Washington, DC [2] RTI-Health Solutions, Washington, DC [3] UCLA Department of Medicine, Los Angeles, CA [4] Mayo Clinic, Rochester, MN [5] Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD [6] Northwestern University Feinberg School of Medicine, Chicago, IL [7] Fred Hutchinson Cancer Research Center, Seattle, WA

Objective: To document the responsiveness of 8 PROMIS® measures in a large population-based sample of cancer patients.

Methods: We recruited cancer patients in the Measuring Your Health (MY-Health) study through 4 SEER cancer registries in three states between 2010-2013, oversampling Hispanic, Black, and Asian race/ethnicity and younger age. Participants were diagnosed with colorectal, lung, Non-Hodgkin's Lymphoma, breast, gynecologic or prostate cancers. Enrolled participants completed a paper baseline and 6-month follow-up assessment for 8 PROMIS Measures (English, Spanish, or mandarin Chinese language translations) : Physical Function (11-item), Fatigue (14-item), Pain Interference (11-item), Anxiety (11- item), Depression (10-item), Ability to Participate in Social Roles (10- item), Cognitive Function (8-item), and Sleep Disturbance (8-item). We selected PROMIS items in each measure based on either their inclusion in commonly used validated short forms, or their frequent selection in the online PROMIS Computer Adaptive Testing (CAT) format. Anchors to evaluate change were single patient-reported items for each domain describing change since the first assessment six months earlier (i.e., a global transition rating). Unadjusted change scores were evaluated within each anchor response option. Mean change was also compared using other clinical groupings (e.g., performance status, cancer status).

Results: Overall, 2968 of 5513 cancer patients (53%) completed baseline (median = 9.5 months post-diagnosis) and 6-month follow-up surveys. The percentage of participants who reported improvement ranged from 35% (sleep disturbance) to 54% (pain interference); the percentage reporting decline ranged from 8% (social function) to 13% (fatigue). We identified small to medium effect sizes (ES) for participants who reported being "a little" worse (Cohen's d range =0.31 – 0.56), and medium to large ES for participants who reported being "a lot" worse (d range = 0.53 – 0.72). We found ES for improvement to be smaller than ES for declines. ES for patients who indicated "a little" improvement were negligible (d range = -0.01 – 0.14), while those for patients who indicated "a lot" of improvement were small (d range = 0.22 – 0.45). Measures were responsive to 6-month change measured by both ECOG performance status (d range = 0.34 – 0.71, all $p<0.0001$), and by cancer status. Patients who indicated active cancer at follow-up (n=443) reported significantly worse symptoms and function across all PROMIS measures vs. those reporting being cancer-free. ES ranged from 0.20 (sleep disturbance) to 0.32 (pain interference), all $p<0.001$. Anxiety increased the most among patients who self-reported a cancer recurrence (n=77) vs. patients who reported no change (n=1,849) in cancer status (mean difference: 4.1 points, d = -0.51, $p< 0.0001$).

Conclusions: The 8 PROMIS domains examined in the MY-Heath study were shown to be sensitive to both negative and positive patient perceived change and

changes in cancer status between baseline and 6 –month surveys. Anchor-based declines were associated with larger changes in PROMIS scores than were improvements over the 6-month period. These results provide useful information to inform the design of future research studies and clinical initiatives.

Grooming a CAT: Customizing CAT Administration Parameters to Increase Response Efficiencies in Specific Research and Clinical Settings

Michael Kallen, PhD, MPH; Karon Cook, PhD; Richard Gershon, PhD

Presenting Author: **Michael Kallen**; michael.kallen@northwestern.edu

Institutional Affiliations: [all] Northwestern University

Objectives: We studied assessment histories from CAT administrations of PROMIS item banks to determine (1) the efficiencies of "standard" CAT administration parameters in limiting the number of items presented per case to achieve established score precision levels and (2) the potential for refining CAT parameters to increase item-presentation efficiencies while maintaining score precision requirements. In the PROMIS 1 Wave 2 Back Pain and Depression Study (UW), CATs were administered to N=417 cases being assessed across 11 PROMIS domains at baseline and up to two follow-up occasions. The following CAT administration parameters were employed: (a) start each CAT with a pre-identified item of moderate difficulty (unique per domain; presented as the first item to each case assessed in that domain); (b) administer a minimum of 4 items to each case; (c) stop each CAT when the SE of a case's estimated theta declines to < 0.3 (OR) when a maximum of 12 items is presented.

Methods: Original CAT. For this study, 12,622 individual CAT administrations were analyzed. CATs ranged in number of items presented from 4 to 12 items, with more than two-thirds of cases (72.5%) experiencing 4-item CATs. The second and third most frequently occurring CATs were 5-item CATs (n=1102; 8.7%) and 12-item CATs (n=964; 7.6%). A total of 64,062 items were presented in these 12,622 CATs, averaging 5.1 items per CAT. Customized CAT. Three new CAT stopping rules were then introduced, each with potential to increase item-presentation efficiency and maintain required score precision. Rule 1: Stop if a case responds to the first 2 items presented using each item's "lowest" response category (indicating non/clinically unimportant status). Rule 2: Administer a minimum of 2 items per case. Rule 3: Stop if the change in SE estimate from previous to current item administration is positive but < 0.01.

Results: Applying Rules 1-3 to the originally analyzed 12,622 CAT data, the flow of cases through CAT administrations of the 11 PROMIS domains assessed was examined from the perspective of "how many cases remain in under assessment status after presenting X items." Implementing the 3 new stopping rules reduced the total number of items presented by 25,643 items to 38,419 items (60.0% of the original number of items presented). After 4 items were administered, only n = 1,824 cases (14.5%) were still under assessment (compared to n = 3,477 (27.5%) in the original CATs). On average, cases needed to complete 3.0 items per CAT (vs. 5.1).

Conclusions: Each of these 3 new rules was introduced to address specific noted inefficiencies in the CAT administration process: Rule 1, for cases not having or possessing only a low/clinically unimportant level of the assessed construct; Rule 2, to allow the SE < 0.3 stopping criterion to come into effect earlier in the item presentation process; Rule 3, for cases experiencing poor domain item bank measurement, typically "floor" or "ceiling"

cases. By customizing a CAT's administration parameters to better suit the needs of a specific research study or clinical practice, increased efficiencies can be anticipated.

Enhancing Communication, Improving Outcomes: PRO Assessment and Reporting via a Home-based Cancer Symptom Management System

Michael Kallen, PhD, MPH; 1 Niina Haas, MA;2 DerShung Yang, PhD2

Presenting Author: Michael Kallen; michael.kallen@northwestern.edu
Institutional Affiliations: 1Northwestern University; 2BrightOutcome

Objectives: Cancer symptoms due to disease progression or treatment side-effects are widely prevalent and can severely impact patients' health-related quality of life. Since most cancer patients receive outpatient treatment, they must then self-manage their subsequent cancer-related symptoms at home. Although patients may receive information and even counseling from their care team regarding anticipated treatment side-effects and symptom management, they may not be able to fully assimilate information and counseling at the time it is provided, due to stress, low literacy, or lack of self-efficacy. We developed a home-based cancer symptom management system, SymptomCareAnywhere (SCA), with the goals of promoting patient self-management, facilitating patient-provider communication, and, ultimately, improving patient outcomes while reducing healthcare costs.
Methods: Intervention patients reported symptoms online or by phone via SCA, which includes 16 core patient-reported cancer symptoms (Reeve et al., 2014). PROMIS Item Banks were used to assess the additional impacts of pain, fatigue, anxiety, depression, and insomnia. Tailored patient education content was offered, based on specific patient symptom severity values or changes. Symptom severity exceeding set thresholds triggered email alerts to Research Nurses for triage and follow up. In addition, a questionnaire measuring patient-physician communication about symptoms was presented. The *Patient Communication About Symptoms* index asks about 16 highly typical symptoms or symptom-related issues experienced by cancer patients: fatigue, nausea/vomit, pain, sleep, mouth sores, fever, numbness, anxiety, depression, hair loss, weight gain, shortness of breath, sexual relationship, physical capabilities, daily routines/work, and leisure/social activities. To indicate if they had discussed a particular symptom with their physician, patients responded using a 3-point option set: 0 (no discussion), 1 (brief discussion), or 2 (longer discussion).
Results: Preliminary findings indicate that the Control Arm (usual care; n=58) and the Intervention Arm (access to SymptomCareAnywhere; n=32) did not statistically significantly differ at baseline in the number of symptoms discussed with their physician (Control=6.09 symptoms; Intervention=7.21 symptoms). However, by 4-weeks follow up, Intervention Arm patients were able to decrease the number of symptoms they needed to discuss with their physician by an average of 2.95 symptoms, while Control Arm patients increased the number of symptoms they needed to discuss with their physician by an average of 0.52 symptoms (ANOVA; p=.016). This symptom discussion need difference continued to be observed at 8-weeks follow up, at which time Intervention Arm patients were able to decrease the number of symptoms they needed to discuss with their physician by an average of 3.70 symptoms, while Control Arm patients were only able to decrease the number of symptoms they needed to discuss with their physician by an average of 0.36 symptoms (ANOVA; p=.035).
Conclusions: Initial results suggest SymptomCareAnywhere, a home-based cancer symptom management system, may function as a bona fide communication tool,

increasing the frequency of distance-based communications, enabling the more successful management of cancer-related symptoms, and allowing face-to-face patient-physician communications to focus on the most difficult to manage symptoms.

CHOIR CAT: Integrated Multi-Objective Probabilistic Computer Adaptive Testing Platform for NIH PROMIS and Toolbox

Ming-Chih Kao, PhD, MD [1], Karon Cook, PhD Garrick Olson, BS [1],Teresa Pacht, BS [1], Beth Darnall, PhD [1],Susan C. Weber, PhD [1], Sean Mackey, MD, PhD [1],

Presenting Author : **Ming-Chih Kao, PhD, MD** mckao@stanford.edu

[1] Stanford Hospital and Clinics, [2] Northwestern University

Objectives: In the study of biopsychosocial model of diseases, psychometrics is the rate-limiting step. Item- response theory (IRT) and computer adaptive testing (CAT), mainly developed in the field of Education, provides the framework for catalyzing this rate-limiting step. The NIH PROMIS and NIH Toolbox efforts in particular offer simultaneous, multidimensional psychometric measurements for assessment and tracking of diseases. Learning from the field of Education and its experiences with CAT, we develop a multi-objective, integrated and probabilistic CAT algorithm to support the assessment and tracking of health outcomes using polytymous items.
Methods: We develop a modular system of efficient psychometric algorithms on CHOIR (Collaborative Health Outcome Information Registry) using IRT and CAT. This system, called CHOIR CAT, is designed as a multi-feature computation engine for the NIH funded psychometric item banks NIH PROMIS and NIH Toolbox. Implemented on open source MEAN stack with D3.js visualization, the system's features include initialization (individualized or patient population priors), item selection (expected Kullback-Leibler, minimum expected posterior variance), advanced item selection (alpha-stratification, exposure control, content balancing, probabilistic constrained optimization), stopping rule (predicted standard error reduction, percentile width, hybrid), and estimation (expected a posteriori, maximum likelihood, maximum a posterior). The multi-objective item selection process is integrated using linear programming.
Results: Item banks and linkages are obtained from Northwestern Access Center and PROsetta Stone. Performance in 4,466 measurements in the Registry are analyzed. We find that basic CAT provided significant reduction in burden (mean number of items ± SD, fold reduction): Anger (6.24+/-1.21, 4.6-fold vs BPAQ), Anxiety (4.93+/-0.97, 1.4-fold vs GAD-7), Depression (4.97+/-1.07, 1.8-fold vs PHQ-9), Fatigue (4.78+/-0.76, 8.4-fold vs FACIT-F), Physical Function (4.11+/- 0.48, 4.9-fold vs HAQ-DI), Pain Interference (4.19+/-0.71, 1.7-fold vs BPI), Sleep Disturbance (4.95+/-1.41, 2.4-fold vs SDQ), Sleep-Related Impairment (4.54+/-1.24, 1.8-fold vs ESS). Altogether, the 132 classic instrument items may be alternatively assessed by 38.7 +/- 7.9 items, for 2.8 to 4.3 fold reduction in patient burden.
Conclusions: In conclusion, using IRT and advanced CAT, CHOIR CAT leverages the powers of NIH PROMIS and Toolbox, and enable true systems biology approach to the study and management of pain.

Collaborative Health Outcomes Information Registry (CHOIR): Open Source Learning Health Systems Platform

Ming-Chih Kao, PhD, MD [1], Sean Mackey, MD, PhD [1] , Karon Cook, PhD [2] , Garrick Olson, BS [3], Teresa Pacht, BS [3], Beth Darnall, PhD [3], Susan C. Weber, PhD [3]

Presenting Author: **Ming-Chih Kao, PhD, MD**; mckao@stanford.edu

[1] Stanford Hospital and Clinics, Dept of Anesthesiology, Division of Pain Medicine
[2] Northwestern University, Medical Social Sciences, [3] Stanford School of Medicine, Stanford Center for Clinical Informatics

Objectives: Attendees will be able to identify defining features of learning healthcare systems, recognize the roles Registry systems play in this setting, and compare and contrast Registry systems versus Electronic Medical Records in their patient care activities.2. Attendees will be able to distinguish classic psychometric instruments from item-response theory based instruments from the NIH, including NIH PROMIS and NIH Toolbox, and explain the differences in characteristics. 3. Attendees will be able to evaluate the impact of Registry systems on patient care and research activities based on the experience of CHOIR.
Methods: In answer to the call from the IOM, we developed the Collaborative Health Outcomes Information Registry (CHOIR), an open-source web application to assess patients and to support clinic staff with integrating the pain registry into the clinic workflow. On the back-end, patient assessment included integrated NIH PROMIS (Patient Reported Outcome Information Measurement Systems) item-response theory computer adaptive testing engine. On the front- end, assessments are designed for use on mobile devices with touch interfaces (smart phones and tablets), while also supporting desktop web browsers. Key technologies used include Java, Oracle database, Google Web Toolkit, jQuery Mobile, AngularJS, and Bootstrap.
Results: Since roll-out in August 2012 and the subsequent slow ramp-up, over 6,000 unique patients have completed surveys, with over 150,000 NIH PROMIS assessments including Global Health (Physical and Mental), Mood (Depression, Anxiety, Anger), Function (Fatigue, Physical Function), Sleep (Sleep Disturbance, Sleep-Related Impairment), Social (Emotional Support, Instrumental Support, Satisfaction with Roles and Activities, Social Isolation, and Ability to Participate in Social Activities). Surveys were completed at home via email link, or at the Pain Clinic, using computers, iPads, Android tablets, and Chrome notebooks.
Conclusions: In conclusion, we have created an open source, extensible platform CHOIR (Collaborative Health Outcomes Information Registry) that enables rapid definition and deployment of data capture tools. This represents a successful partnership between the NIH and Stanford with funding from most of the NIH Institute Directors. Future works include the expansion of survey items, into additional disease areas, dissemination of code, as well as networked registry build-out.

Feasibility and Validity of PROMIS© in systemic lupus erythematosus (SLE)

Shanthini Kasturi, MD, Jayme Burket Koltsov, PhD, Jessica Berman, MD, Kyriakos Kirou, MD, Alana Levine, MD, Lisa Sammaritano, MD, and Lisa A. Mandl, MD, MPH

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

Institutional Affiliations: Hospital for Special Surgery, Weill Cornell Medical School, New York, NY

Objective: To assess the feasibility of administering PROMIS computer adaptive tests (CATs) serially to SLE outpatients, as well as correlating PROMIS CATs with legacy patient reported outcome measures (PROMs), SLE disease activity, and organ damage.
Methods: Adult SLE patients were recruited from a tertiary SLE Center. Subjects completed the Short-Form-36 (SF-36),

LupusQoL- US, and selected PROMIS CATs at baseline and at one subsequent time point. SLE disease activity and damage were measured at the same time points with the SELENA-SLEDAI and SLICC-ACR respectively. Spearman's correlation coefficients were calculated comparing PROMIS domains with a) similar domains in the SF-36 and LupusQoL-US b) SLE disease activity and c) organ damage.

Results: To date, 33 /40 (83%) of SLE patients approached enrolled, and 28 (70%) have completed a follow-up assessment. Subjects were 21-63 years, (mean 40), 85% female, 78% non-white, 38% Hispanic, and 39% were Medicaid. SLE duration ranged from 1-22 years (mean 12). Subjects had a wide range of disease activity (range 0-16; higher is worse) and organ damage (range 0-8; higher is worse), with 38% assessed during a disease flare. PROMIS CATs showed strong correlations with relevant domains in the LupusQoL-US and SF-36. PROMIS physical function scores highly correlated with LupusQoL- US physical (r=0.79), pain (r=0.76), as well as SF-36 physical function (r=0.91), role physical (r=0.71) and social function (r=0.84) domains; there was an inverse correlation with SF-36 bodily pain (r= -0.77). PROMIS fatigue scores inversely correlated with LupusQoL-US physical (r= -0.77), pain (r= -0.71), planning (r= -0.75), and fatigue (r=-0.75) while strongly correlating with the SF-36 bodily pain domain (r=0.70). PROMIS pain behavior and pain interference domains showed strong negative correlations with physical (r = -0.85 and - 0.80) and pain (r= -0.88 and -0.81) domains in the LupusQoL-US as well as SF-36 physical function, role physical, and social function domains (r= -0.72 to -0.88), but a positive correlation with bodily pain (r=0.72 and 0.75). All correlations had p-values <0.001. There were no statistically significant correlations between any PROMIS domain and SLE disease activity or damage. Subjects provided overwhelmingly positive feedback regarding PROMIS, citing the concise questions, ease of administration, and relevance of the assessment to their life.

Conclusions: To our knowledge, this is the first study to assess the feasibility and validity of administering PROMIS CATs to adult SLE outpatients. These data show that PROMIS CATs can be successfully administered to a diverse cohort of SLE patients, and that PROMIS CATs show strong, statistically significant correlations with established general health and SLE-specific PROMs. In this small sample, PROMIS did not correlate with disease activity or damage, suggesting that PROMIS captures patient relevant domains distinct from standard disease activity and damage measures. Valid PROMs which are acceptable to patients 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 disease management in SLE.

Predicting EQ-5D Utility Index from the PROMIS-10 Global Form - Ready for PrimeTime?

Irene Katzan MD, Nicolas Thompson MS, Youran Fan PhD, Sandra Griffith PhD

Presenting author: **Irene Katzan MD**; katzani@ccf.org

Institutional Affiliation: Cleveland Clinic

Objective: PROMIS-10 Global Form is being increasingly utilized to provide a brief assessment of mental and physical health. The ability to generate a utility index from the PROMIS-10 would expand its use to health utility analyses. A model to predict EQ-5D index from the PROMIS-10 is available (Revicki, et al, Qual Life Res (2009) 18:783–791), but there has not been extensive work on its real-world application in disease-specific populations. The objective of this study was to evaluate the performance of the published EQ-5D prediction model in a neurological population.

Methods: EQ-5D and PROMIS-10 Global Form were electronically completed one time by patients seen in epilepsy, sleep, and cerebrovascular clinics at Cleveland Clinic from October 23, 2014 to November 6, 2014. Model 3 of the prediction model published by Revicki was used to generate predicted EQ-5D index values. These were compared to actual EQ-5D index values by examining predicted values stratified by actual EQ-5D index values in the following groups: -0.109 to 0.25, 0.26 to 0.5, 0.51 to 0.75, 0.76 to 1. Data from the cohort were then used to derive new parameter estimates for that prediction model and the analysis was repeated.
Results: There were 769 patients in the study cohort with an average age of 50.9 years (standard deviation (SD) 15.6) , of which 56.6% were female and 79.5% were Caucasian. The median EQ-5D index was 0.82 [interquartile range 0.71, 1]. The average PROMIS physical T score was 44.2 (SD 9.3), the average PROMIS mental T score was 46.2 (SD 9.9). Predicted EQ-5D index values ranged from 0.37 to 0.88 using the published model and from 0.40 to 1.06 using our model. Actual EQ-5D scores in the study cohort ranged from –0.109 to 1 (the theoretical range of EQ-5D index). The lowest and highest possible predicted EQ-5D were 0.33 and 0.88 based on the published model and 0.33 and 1.07 based on our model. The low EQ-5D index values were overestimated and high EQ-5D index values are underestimated in both models. This was less pronounced when using model parameters based on data from our own cohort.
Conclusion: Prediction of EQ5D index from PROMIS10 using a published model did not perform well in a separate cohort comprised of patients with neurological disease. Reestimation of model parameters based upon our own study population improved performance, although it continued to overestimate the EQ-5D index in patients with low health-related quality of life. Re-estimation of parameters in the available model may be required before it can be used in other populations. Further research is needed to optimize predictions of EQ-5D index using the PROMIS-10 Global Form.

The PROMIS-29 in systemic sclerosis: associations with clinical characteristics

Linda Kwakkenbos [1], Brett D. Thombs[1], Susan Bartlett2, Marie Hudson [2], Luc Mouthon[3], Vanessa Malcarne [4], Maureen Sauve [5], Dinesh Khanna [6], SPIN investigators

Presenting Author: **Dinesh Khanna**; khannad@med.umich.edu

Institutional Affiliations: [1]Department of Psychiatry, McGill University, Montreal, Canada; [2]Department of Medicine, McGill University and Jewish General Hospital, Montreal, Canada; [3]Université Paris Descartes, Paris, France; Assistance Publique-Hôpitaux de Paris (AP-HP); [4]Department of Psychology, San Diego State University, San Diego, California, USA; [5]Scleroderma Societies of Canada and Ontario, Hamilton, Ontario, Canada; [6]University of Michigan Scleroderma Program, Ann Arbor, MI

Background: Systemic sclerosis (SSc) is a chronic, multisystem connective tissue disorder that is characterized by vasculopathy and fibrosis of the skin and internal organs. SSc is frequently divided into limited and diffuse cutaneous subsets (lcSSc and dcSSc), with the latter associated with worse outcomes. SSc has far reaching consequences for physical health, as well as emotional and social well-being and health-related quality of life.
Objective: To explore associations of PROMIS-29 domains with clinical variables in subjects with SSc enrolled in a large multinational study.
Methods: Adult SSc subjects were enrolled in the Scleroderma Patient-centered Intervention Network (SPIN) Cohort from 17 centers across Canada, the USA and the UK between April 2014 and February 2015. Baseline medical data are provided by the

enrolling physician, and SPIN Cohort subjects complete outcome measures online every 3 months. Subjects were included in the analyses if they completed at least one PROMIS-29 domain at baseline. For all PROMIS-29 domains, Pearson correlations were calculated with age and continuous clinical variables, and t-tests were conducted for gender and dichotomous clinical variables. The standardized mean effect size [95 Confidence Interval (CI)] was calculated to assess the magnitude of the difference between groups. Correlations and effect sizes were interpreted as small (≤0.3), moderate (between0.3 and 0.5), or large (≥0.5).

Results: In total, 403 subjects were included in the analyses. Mean age was 54.8 years (SD=11.9) and mean time since onset of the first non-Raynaud symptom was 11.7 years (SD=8.9). Most subjects were female (n=348, 86.4%) with lcSSc (n=234, 58.1%). The PROMIS-29 domains were 0.1-0.8 SD below general population. Of disease-related variables, the involvement of the gastrointestinal tract was consistently associated with worse outcomes across PROMIS-29 domains (lower scores for function and roles, higher scores for fatigue, pain interference, pain intensity, anxiety and depression), with moderate to large effect sizes in 7 of 8 domains (Tables 1 and 2). Other clinical variables with decrements in at least 4 domains included: dcSSc vs. lcSSc, and presence of contractures. Other associations are shown in Tables 1 and 2. (available upon request)

Conclusions: SSc is associated with significant impairment in PROMIS-29 domains, with certain disease characteristics more commonly associated than others. These data inform priorities for future patient-centered research.

Construct validity of the PROMIS-29 in systemic sclerosis: preliminary results from the Scleroderma Patient-centered Intervention Network (SPIN) Cohort

Linda Kwakkenbos [1], Brett D. Thombs [1], Susan Bartlett [2] , Marie Hudson [2], Luc Mouthon [3], Vanessa Malcarne [4], Maureen Sauve [5], Dinesh Khanna [6], SPIN investigators

Presenting Author: **Dinesh Khanna;** khannad@med.umich.edu

Institutional Affiliations: [1]Department of Psychiatry, McGill University, Montreal, Canada; [2]Department of Medicine, McGill University and Jewish General Hospital, Montreal, Canada; [3Université Paris Descartes, Paris, France; Assistance Publique-Hôpitaux de Paris (AP-HP); [4]Department of Psychology, San Diego State University, San Diego, California, USA; [5]Scleroderma Societies of Canada and Ontario, Hamilton, Ontario, Canada; [6] University of Michigan Scleroderma Program, Ann Arbor, MI

Background: The Patient-Reported Outcomes Measurement Information System (PROMIS®) initiative is a cooperative research program designed to develop, evaluate, and standardize item banks to measure patient- reported outcomes across medical conditions. The PROMIS-29 measure contains 29 items, which include four items each for domains reflecting physical function, anxiety, depression, fatigue, sleep disturbance, pain interference, and ability to perform social roles, plus a single item on pain intensity. Scores are standardized with a mean of 50 and standard deviation (SD) of 10. Higher scores represent more of the domain being measured (e.g., greater sleep disturbance, greater ability to perform social roles).

Objectives: To examine feasibility and construct validity of the PROMIS29 in patients with systemic sclerosis (SSc) enrolled in a large multinational study.

Methods: English-speaking patients with SSc and ≥18 years of age were enrolled in the Scleroderma Patient- centered Intervention Network (SPIN) Cohort between April 2014 and January 2015 from 17 centers across Canada, the USA and the UK. Baseline medical data are provided by the enrolling physician, and SPIN Cohort patients complete outcome measures online every 3 months. Patients were included in the analyses if they completed at least one PROMIS-29 domain at baseline. Floor and ceiling effects were defined as >15% of patients having the lowest or highest possible domain score, respectively. To examine convergent validity of domains, hypotheses were formulated a-priori about the associations of domains and legacy measures. The magnitude of the correlations was interpreted as small ($|r| \leq 0.3$), moderate ($0.3 < |r| < 0.5$), or large ($|r| \geq 0.5$).

Results: In total, 376 patients were included in analyses. Mean age was 55.1 years (SD=11.9) and mean time since onset of the first non-Raynaud symptom was 14.4 years (SD= 11.8). Most patients were female (n=329, 87.5%) and diagnosed with limited SSc (n=207, 55.1%). Means for the PROMIS-29 domains were: function 42.4 (SD=8.7), anxiety 51.8 (SD=10.1), depression 50.9 (SD=9.3), fatigue 56.2 (SD=11.4), sleep 52.0 (SD=5.2), roles 47.4 (SD=9.5), pain interference 55.9 (SD=9.8), and pain intensity 3.8 (SD=2.7). There was a floor effect for anxiety (33.8%) and depression (37.7%), and ceiling effects for function (19.1%), roles (15.1%) and pain interference (24%). Most hypotheses were confirmed (7 of 9) and all were in the hypothesized direction (Table 1).

Conclusions: Results of our study support the construct validity of the PROMIS-29 in patients with SSc. Future studies should examine the influence of floor- and ceiling effects for some domains, as well as other psychometric properties of the measure.

Evaluation of the comparability of scores from CAT and short-form

Jin-Shei Lai, PhD [1], Jennifer L. Beaumont, MS [1], Stewart Goldman, MD [2]

Presenting Author: **Jin-Shei Lai, PhD**, js-lai@northwestern.edu

Institutional Affiliations: [1] Medical Social Sciences, Northwestern University [2] Lurie Children's Hospital, Chicago.

Objective: Despite the demonstrated negative effects of tumors and treatments, the health-related quality of life (QOL) of childhood brain tumor patients and survivors has not been well-studied. It is partially because their experiences are considered unique compared to the majority of pediatric cancer survivors and also because the functional impact of the tumors and the range of surgical and treatment effects can vary based upon tumor location. PROMIS offers an opportunity to better understand the QOL of pediatric BT patients by comparing how it deviates from that of the US pediatric general population. Two administration modes are available in PROMIS, computerized adaptive testing (CAT) and short-form, and each provides their own advantage. This study aims to evaluate the comparability of scores obtained from CATs and short-forms.

Methods: Brain tumor patients (ages 5-21 years; at any disease and treatment stage) and their parents were recruited from CDH Proton Center and Lurie Children's Hospital, Chicago. Parents of children aged 5-7 years completed proxy version of PROMIS CATs and short-forms (Fatigue, Mobility, Upper Extremity, Depressive Symptoms, Anxiety, and Peer Relationships); children aged 8-21 completed the self-reported version. Pearson correlations were used to evaluate the comparability between CAT and short-form scores by domain.

Results: The target sample size is 450. To date, 108 brain tumor patients (86 aged 8-21 years; 22 aged 5-7 years) completed

baseline assessment using appropriate versions (either proxy or self-report version). For patients (self-report version), mean (median) number of items administered in CAT was 9.4 (12), 8.9 (8.5), 8.4 (7), 10.6 (12), 8.2 (6), 9.4 (12), and 7.6 (6) for Anxiety, Fatigue, Mobility, Upper Extremity, Depressive Symptoms, Anxiety, and Peer Relationships, respectively. For parent (proxy version), mean (median) number of items administered in CAT was 7.1 (5), 7.4 (5), 5.9 (5), 5.2 (5), 6.9 (5), 7.1 (5), and 5.7 (5) for Anxiety, Fatigue, Mobility, Upper Extremity, Depressive Symptoms, Anxiety, and Peer Relationships, respectively. High correlation coefficients between CAT and short-form scores were found across all domains, ranging from 0.92 (Peer Relationships) to 0.98 (Depressive Symptoms) and from 0.94 (Peer Relationships) to 0.99 (Anxiety) for self-report and proxy versions, respectively. However, more skewed distributions were observed in short-form scores than in CAT scores.
Conclusions: This study demonstrated the comparability of scores obtained from CATs and short-forms. Yet, more skewed distributions were found in scores using short-forms. This finding suggests that the dynamic CAT administration is better in measuring patients who are at the extreme ends (either floor or ceiling) of the measurement continuum.

Assessing Physical and Mental Health of Veterans Receiving VA Health Care Using a PROMIS Profile

Sherri L. LaVela, PhD, MPH, MBA [1,2]; Bella Etingen, MA [1]

Presenting Author: **Sherri L. LaVela**, PhD, MPH, MBA; s-lavela@northwestern.edu; sherri.lavela@va.gov

Institutional Affiliations: [1] VA Health Services Research and Development, Department of
Veterans Affairs, Hines, IL, [2] Center for Healthcare Studies, Institute for Public Health and Medicine, General Internal Medicine and Geriatrics Feinberg School of Medicine, Northwestern University, Chicago, IL

Objective: Although patient-reported outcomes are vital indicators of thorough health assessment, Patient-Reported Outcome Measurement Information System (PROMIS) items are not widely or systematically used within Veterans Health Administration (VA). The objective of this study was to assess Veterans' responses to patient-reported outcome (PROs) scales, to describe the relationship among Veteran responses to PROMIS-29 profile scales, and to examine associations with diagnoses recorded in patient charts.
Methods: Cross-sectional mailed national survey of Veterans (n=3221) receiving VA health care. The mailed survey provided demographics, Veteran characteristics, and PRO data using the PROMIS-29 Profile v1.0 scales. PROMIS scales produce standardized t-score based on population norms (m=50, sd =10); all PROMIS scores represent a higher level of the construct being measured by each scale. VA administrative databases provided information on select health conditions (depression, anxiety, sleep disorders, pain disorders). Pearson product-moment correlation coefficients (r) were calculated for PROMIS-29 Profile scale t-scores; bivariate comparisons (t-tests) were conducted for Veterans diagnosed with select health conditions and corresponding mean PROMIS scale scores (e.g., PROMIS depression scores were compared for Veterans with and without depression).
Results: Veteran respondents were mostly male (94.7%), white (79.3%), and were 68 years of age, on average. Correlations among Veteran respondents' PROMIS 29 domain scores were mostly moderate (e.g.,

r=0.64 to .43), however, depression and anxiety (r=0.86, p<.0001), pain interference and pain intensity (r=086, p<.0001), fatigue and anxiety (r=0.66, p<.0001), fatigue and depression (r=0.67, p<.0001), pain interference and physical function (r=-0.65, p<.0001), and pain interference and fatigue scale scores (r=0.65, p<.0001) produced strong correlations; correlations among all scales were significant at the p<.0001 level. Of our Veteran sample, 28.6% had a depression diagnosis, 25.4% had diagnosed anxiety, 8.2% had a pain disorder, and 24.8% had a sleep disorder. Bivariate comparisons showed that Veterans with depression reported significantly higher PROMIS depression scores, on average, compared to Veterans without depression (60.3 vs. 49.6, p<.0001). Likewise, Veterans with anxiety reported significantly higher average PROMIS anxiety scores than Veterans without anxiety (62.7 vs. 50.9, p<.0001). Further, Veterans with a pain disorder reported significantly higher pain interference (65.3 vs. 57.7, p<.0001) and pain intensity (6.4 vs. 4.4, p<.0001) scores (vs. Veterans without a pain disorder). Additionally, Veterans with a sleep disorder (vs. Veterans without a sleep disorder) reported significantly higher sleep disturbance (55.8 vs. 51.2, p<.0001) and fatigue (57.5 vs. 51.8, p<.0001) PROMIS domain scores.

Conclusions: Veteran responses to PROMIS scales have a moderate-to-strong relationship with one another. Additionally, mean differences in PROMIS scale scores measuring constructs congruent with select clinical diagnoses suggest that PROs may be an appropriate indicator of condition symptomology among Veterans. Collectively, these results indicate that the PROMIS-29 Profile is appropriate for use in a Veteran population receiving VA health care. PROs may be a useful tool for VA providers to assess patient's physical and mental health. Additionally, PROMIS items provide measurements normed to the general population, offering the VA a way to easily compare the health status of Veterans to the adult population at large.

PROMIS Pain, Fatigue, Sleep, Fatigue, Anxiety and Depression in Patients with Active Rheumatoid Arthritis

Yvonne C. Lee, MD, MMSc [1], Alyssa Wohlfahrt [1], ChihChin Liu,PhD [1], Y. Emily Lo, MPH [1], Tuhina Neogi, MD, PhD [2], KristinePhillips, MD, PhD [3], Marcy Bolster, MD [4], Clifton Bingham, MD [5]

Presenting Author: **Yvonne Lee, MD, MMS**c; ylee9@partners.org

Institutional Affiliations: [1] Brigham and Women's Hospital, Boston, MA [2] Boston University School of Medicine, Boston, MA [3] University of Michigan Medical School, Ann Arbor, MI[4] Massachusetts General Hospital, Boston, MA

Objectives: Previous studies of the performance of PROMIS in rheumatoid arthritis (RA) have consisted primarily of individuals with well-controlled inflammatory disease. The objective of this study was to examine the performance of PROMIS in individuals with active RA.

Methods: Data are from the baseline visit of the first 82 participants in a multicenter study examining pain in RA patients starting a disease modifying antirheumatic drug (DMARD) for active inflammatory disease. The Clinical Disease Activity Index (CDAI) was calculated. Participants completed a pain numeric rating scale (NRS), the PROMIS pain intensity short form and the PROMIS pain behavior, pain interference, fatigue, sleep disturbance, sleep-related impairment, depression and anxiety computerized adaptive tests. Mean values were compared to the general population norm of 50 ± 10. Associations between PROMIS measures and CDAI were calculated using linear regression models, adjusted for

age, gender, race, serostatus and disease duration.

Results: Participants were mostly women (79.3%) and white (76.6%),with a mean age of 51.3 ± 13.6 years. 51.2% had high disease activity(CDAI > 22). 34.1% had moderate disease activity (10 < CDAI ≤ 22), and 14.6% had low disease activity (CDAI ≤ 10) (Table Available upon request). Across CDAI categories, the mean pain NRS (0-10) increased from 4.8 to 7.0, and mean PROMIS pain intensity scores increased from 50.2 to 54.2. PROMIS pain interference and pain behaviors scores were elevated across all CDAI categories. PROMIS fatigue, sleep disturbance and sleep-related impairment scores were similar to general population norms in the low and moderate CDAI categories but significantly elevated among those with high disease activity. PROMIS depression and anxiety scores did not differ across CDAI categories.

Conclusions: Among RA patients with active inflammatory disease, the range of PROMIS pain intensity scores was limited compared to the range of pain NRS scores. Compared to general population norms, RA patients had similar pain intensity scores but higher pain behaviors and pain interference scores in all disease activity categories. PROMIS fatigue, sleep disturbance and sleep-related impairment exhibited the greatest range across categories of disease activity. PROMIS depression and anxiety were not elevated compared to general population norms and were not associated with disease activity

PROMIS in TCM diagnosis: chance and challenge

Fengbin Liu, MD [1], Xiaoying LEE, MD [2], Zhengkun HOU, MD [1],Xinlin CHEN,PhD[3]

Presenting Author: **Fengbin Liu**,
veinlix@163.com

Institutional Affiliations: [1] the first affiliated hospital of Guangzhou University of Chinese Medicine [2]Jiangmen WuYi hospital of Chinese Medicine, the sixth affiliated hospital of JiNan University [3] Guangzhou University of Chinese Medicine

Objective: To evaluate the feasibility, effectiveness, challenge of PROMIS applying to Traditional Chinese Medicine (TCM) diagnosis.

Methods: Retrospective analyzing the papers on development, application of PRO, quantification, modernization researches in Chinese medicine, and core technology of PROMIS, we evaluate the application of PROMIS in TCM effect assessment, and suggest a diagnostic process with PROMIS in TCM. Results: The system of TCM syndrome diagnosis development is partly based on philosophy. And the main difference for diagnosis and curative effect evaluation between western medicine and Chinese medicine is the basic index objective or subjective. And one of the most important theory in TCM is the eight principle theory (yin and yang, cold and heat, outer and inner, deficiency and excess), which is characterized by unity of opposites. The diagnostic information collected is through inquiry, auscultation, olfaction, inspection and palpation, which decides the syndrome diagnosis and treatment of the patient. How to quantify and make it accurate is the key point for TCM development. With the methodology of PROMIS, the process of syndrome differentiation in Chinese medicine is found to be similar to acquire the patients' potential trait, which is mostly related to the identification of opposing essence. When the exact potential trait is detected by item response theory (IRT) and computerized adaptive test (CAT), the situation of the patients could also be manifested by the capability value θ. With different dimensions value, different syndromes importance weights can also be compared. According to the value of weight, doctors can judge the major syndrome and minor syndrome for the treatment. However, we should also notice that the

evaluation of tongue diagnosis and pulse taking is based on doctors' experience, which needs to be quantified as well. If the methodology of PROMIS were used for smart TCM diagnosis, "PROMIS" should be changed into "CROMIS" (clinical reported outcomes). **Conclusion:** As PROMIS is mainly used in outcomes evaluation in western medicine, the application of PROMIS methodology in Chinese may even go further, not only in the field of effect evaluation, but also diagnostic process, especially hopeful of the quantification of syndrome differentiation.

Incorporating Chinese Medicine Syndrome into Patient-Reported Outcomes Assessment: Systematic Review and Recommendations

Aiping Lu, MD, PhD; [1,2], Chi Zhang, MD, PhD [2], Miao Jiang, MD, PhD [2,3], Zhaoxiang Bian, MD, PhD [1]

Presenting Author: **Aiping Lu, MD, PhD**; aipinglu@hkbu.edu.hk

Institutional Affiliations: [1] School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong [2] Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China [3] Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, US

Objective: Chinese medicine Syndrome (CMS) requires reliance on the patient's subjective report of symptoms. It can provide better understanding and detail regarding the impact of classification, diagnosis, and approach to treatment for disorders. The last decade has witnessed increased attention to the role of CMS patient-reported outcomes (PROs) in Chinese medicine. However, these outcomes have not been systematically examined.

Methods: We systematic review randomized controlled trials (RCTs) using CMS PRO measures, and summarize and integrate information on published CMS PRO reviews. Recommendations are made on better CMS PROs for classification, diagnosis and evaluating responsiveness. The National Library of Medicine's Medline and China National Knowledge Infrastructure (CNKI) were searched (inception to Jan 7, 2015).

Results: One hundred twenty RCTs were identified. Forty-eight studies involved depression, 11 involved insomnia. Fifty-eight dealt with syndrome of stagnation of liver qi, 43 deficiency of spleen and 27 deficiency of kidney. Ninety-five percent of studies used a CMS PRO as a primary endpoint. Although 94% addressed clinical significance of the outcomes, only 29% of the studies used a universally accepted and validated CMS PRO questionnaire. In addition, 6 reviews were retrieved. Overall, all reviews analyzed exhibited a number of methodological drawbacks. A number of CMS scales are in the early stages of development, and all require further validation work.

Conclusions: CMS assessment in clinical practice has been shown to lead to a better understanding of patients' classification and treatment response. Potentially, CMS assessment should be more fully incorporated into outcome evaluation to monitor the progress of a patient's disease and benefit from treatment. Appropriate measures must possess adequate evidence for validity. To facilitate the interpretation of results from such Chinese medicine RCTs, investigators are encouraged to pay more attention to key questions and methodological issues as identified in this study.

Establishing Clinical Meaning and Severity Cut-points on Patient Reported Outcome Measures in Juvenile Idiopathic Arthritis

Esi Morgan DeWitt, MD,MSCE [1,2]; Constance A. Mara, PhD [3], Bin Huang, PhD [1,4]; Kimberly Barnett, BS [3]; Adam C. Carle, PhD [1,3]; Jennifer E. Farrell, MA[2]; Karon F. Cook, PhD[6]

Presenting Author: **Constance A. Mara, PhD** Constance.Mara@cchmc.org

Institutional Affiliations: [1] Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH. [2] Division of Rheumatology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH [3], James M. Anderson Center for Health Systems Excellence, Cincinnati Children's Hospital Medical Center, Cincinnati, OH [4]Division of Behavioral Medicine and Child Psychology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH [5]Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH [6]Center for Behavior and Health, Institute for Public Health and Medicine, Northwestern University, Chicago, IL

Objective: Patient reported outcome measures (PROs) are used increasingly in clinical care. A framework to interpret scores according to degree of clinical severity would enhance their practical use. Furthermore, use of PROs in evaluation of treatment effectiveness over time requires establishment of minimal important differences (MID) in change scores.

Methods: We identified clinical severity thresholds and MID for measures of mobility, upper extremity function, fatigue, and pain interference working with patients with juvenile idiopathic arthritis (JIA), parents of JIA patients, and clinical providers who treat JIA using standard setting methodology modifiedfrom educational testing. Data from Patient-Reported Outcomes Measurement Information System (PROMIS) item bank longitudinal validation collected on 121 JIA patients was used to develop clinical vignettes across a range of symptom severity. Vignettes were created based on most likely item responses at different levels on the T-score metric [mean = 50; SD = 10]. Vignettes were anchored at 5-point intervals (0.5 SDs). Parents, patients, and clinical providers participated as expert panelists in separate one-day meetings. Vignettes were ordered and placed on cards. Panelists identified adjacent vignettes considered to represent upper and lower boundaries separating category cut points (i.e., none /mild problems, mild/moderate, moderate/severe). Cut scores were defined as mean score for boundary vignettes. To define MIDs, panelists responded to items to represent "just enough improvement to make a difference". Average change scores served as estimates of MID.

Results: For fatigue, pain interference, and upper extremity function, parents and clinical providers set identical cut points for severity. Patients set higher cut points for severity than parents and clinical providers, typically by 0.5 SD. For mobility, parents tended to set lower cut points for severity than clinical providers and patients. Size of MID varied according to severity classification of the symptom. Clinicians consistently estimated MID scores smaller than parents. Patient MID score estimates varied in relation but tended to be between clinicians and parents estimates. MIDs estimated by the panelists were typically larger than the MIDs determined using statistical methods.

Conclusions: We used a modified educational standard setting method to estimate clinically relevant cut points to classify severity for PROMIS measures of mobility, upper extremity, fatigue, and pain interference. Parallel exercises identified these cut points from the perspectives of patients with JIA, parents of a child with JIA, and clinical providers who treat JIA. We explored a novel means of determining MID

from the patient/parent/provider perspective. This allows for meaningful interpretation of PROMIS measures in a clinical setting.

Effectiveness of acupuncture and oriental medicine in the management of pain--a prospective cohort study

Ben Marx, MAcOM, LAc, Deb Espesete, MAcOM, LAc, MPH, Beth Howlett, MAcOM, LAc , Beth Burch, ND, Tim Chapman, PhD

Presenting Author: **Ben Marx, MAcOM**, LAc; bmarx@ocom.edu

Institutional Affiliations: Oregon College of Oriental Medicine (OCOM)

Objective: Compare first and fifth visit outcomes for patients seeking treatment for pain in a high-volume acupuncture and oriental medicine (AOM) teaching clinic.
Methods: Between Nov 2009 and Dec 2013 three PROMIS instruments--the Global Short Form, Adult Physical Function, and Adult Pain Interference--were administered to a cohort of consenting patients treated for pain at the OCOM intern clinic. Instruments were administered at visit 1 and again at visit 5. Patient demographics and changes in global physical and mental health, physical functioning, and pain interference were assessed.
Results: A cohort comprising 374 patients was analysed. Demographics were similar to those reported in other AOM clinics. The majority of patients were white females over the age of 50. Changes in mean scores from visit 1 to visit 5 for the four outcomes measures were as follows: Global Physical Health 12.68 to 13.68 ($p<.05$), Global Mental Health 13.74 to 13.77 ($p=.79$), Physical Functioning 37.94 to 40.11 ($p<.05$), and Pain Interference 18.40 to 14.55 ($p<.05$).
Conclusions: Statistically significant improvements in Global Physical Health, Physical Functioning and Pain Interference from visit 1 to visit 5 suggest that intern delivered AOM is an effective intervention for the management of pain. Future large-scale cohort studies are warranted, and should address pain outcomes related to particular AOM treatments and Chinese pattern diagnoses

Comparison of PROMIS® survey between scleroderma patients in an academic center and patient-based scleroderma foundations

Vivek Nagaraja MD [1], Veronica Berrocal PhD [2], Kerri Connolly [3], Ann Kennedy [4], Daniela Seelmann MD [5] and Dinesh Khanna MD MS [2]

Presenting author: **Vivek Nagaraja MD**; Vivek.Nagaraja@utoledo.edu

Institutional Affiliations: [1] University of Toledo, Toledo, OH [2] University of Michigan, Ann Arbor, MI [3] Scleroderma Foundation, Boston, MA [4] Federation of European Scleroderma Associations, Tournai,Belgium [5] Universidad de Los Andes, Santiago, Chile

Objective: The National Institutes of Health Patient- Reported Outcomes Measurement Information System (PROMIS®) roadmap initiative is a cooperative research program designed to develop, evaluate, and standardize item banks to measure patient-reported outcomes (PROs) across different medical conditions as well as the US population (www.nihpromis. org). It has comprehensive items banks that assess physical, mental, and social well-being. The aim of this study was to compare the PROMIS survey between the scleroderma patients at an academic center and patient-based foundations as this has implications for large epidemiological studies (such as Scleroderma Patient Intervention Network).
Methods: A study titled 'PROMIS in rheumatology' was created in the

Assessment center website. This study contained 13-PROMIS instruments. Patients seeking care in the academic Scleroderma clinic (UM – University of Michigan) were approached to participate in the PROMIS survey. Patients were also recruited from the Scleroderma patient-based foundations (SF) namely – the Scleroderma Foundation and the Federation of European Scleroderma Associations, through the respective social media pages and e-newsletters. Average T-scores of the patient-based foundation scleroderma cohort were compared with those of the UM scleroderma patient cohort.
Results: Thirty-seven patients at UM and 243 patients from SF have so far completed the survey. In both groups, the T-scores in the following domains were approximately 1 standard deviation worse than the United States (US) general population (GP) – fatigue, physical function, pain interference, satisfaction in roles and activities. Anger and social isolation banks were comparable to US GP. In comparison to the UM scleroderma cohort, the T-scores of the SF patient cohort was significantly worse for pain behavior and social isolation; however, the differences were not clinically meaningful.
Conclusions: The patients with SSc have decrements in the health- related quality of life (HRQOL) on PROMIS measures when compared to the US general population. The scores are similar in patients recruited from the scleroderma clinic and patient foundations and both groups can be recruited in future studies assessing HRQOL.

Tracking adolescent outcomes in a child welfare setting using emotional distress level 2 PROMIS measures

Bernie S. Newman, Ph.D., MSW, Melinda Moskal, MSW

Presenting Author: **Bernie S. Newman, Ph.D., MSW** bernienewman@comcast.net

Institutional Affiliations: People R Us, Philadelphia, Pa. 19012

Objective: To pilot test the clinical utility of using a composite of PROMIS Level-2 Emotional Distress Measures and the adaptation of the Affective Reactivity Index (ARI) to guide treatment and assess progress in an adolescent community residential setting.
Methods: We developed a composite index consisting of three PROMIS Emotional Distress pediatric subscales in the areas of anger, depression and anxiety and the Affective Reactivity Index (ARI) measuring irritability to use as an initial domain specific severity score. These domains were selected as they represent common mental health issues for adolescents in the child welfare system.. It was our intention to create a short and simple inventory that youth could complete in less than 10 minutes on their first day at PRU Community Services and intermittently during and immediately following treatment in this group home setting. An adapted measure was created for the primary clinician to complete during the first week of the youth's stay, during and following treatment. Although preliminary data analysis relies on a small sample of the self-reported pretests only from this newly opened facility, data are collected as new adolescents enter the program and ongoing assessment will provide a larger sample of interim and post-test data from both youth and clinicians. Alpha coefficients and inter-rater differences will be computed as our data set grows. A time series design will reflect degree of change following initiation of treatment from both the youth and the clinician's perspectives.
Results: The composite measure was easily integrated into the initial intake process for both staff and youth. Adolescents were willing to complete the measure. Initial scores on a small sample reflect a range from slight to moderate on anger, depression and anxiety and slight to moderate-severe on irritability.

Conclusions: The Emotional Distress Pediatric Bank items lend themselves to practical application in terms of feasibility and ease of use. The degree to which they are useful as a treatment planning and evaluative tool over time will be presented

PROMIS in Sweden?

Bo Palaszewski [1] PhD, Evalill Nilsson [2,3] PhD, John Chaplin [4] PhD

Presenting Author: Evalill **Nilsson;**
Evalill.Nilsson@regionostergotland.se

Institutional Affiliations: [1] Centre of Registers, Västra Götaland, Sweden [2] The Research & Development Department of Local Health Care, Region Östergötland, Sweden, [3] Quality Register Center QRC Stockholm, Stockholm, Sweden [4] Inst. Clinical Sciences, Sahlgrenska Academy at the University Gothenburg, Sweden

Background: The health services of Sweden's 21 Region and County Councils are encouraged by central government to contribute to the evaluation of healthcare across the regions and to include patient-reported outcomes measures (PROMs). The most systematic use of patient-reported data in the Swedish healthcare system today is seen in the National Quality Registers (NQRs). The registers are designed to be used in an integrated and interactive way for quality control, research, and clinical management. There are over 100 different NQRs, each focused on a specific diagnosis, condition or intervention. The health services collect individual patient data and these are input into the national registries. In 2015 almost all NQRs report including patient- reported measures, either to measure outcomes (PRO) or to measure patient satisfaction with healthcare, or both.
Objective: to describe and evaluate the possible use of PROMIS in Swedish national quality registries (NQRs).
Methods: In the NQR implementation of PROs the need for both generic and disease-specific measures, and both electronic and paper-and pencil administrations are often included. It is considered desirable that comparison can be made across conditions and between national healthcare systems. Therefore, a generic PRO measure (or set of measures) is required. Barriers for the inclusion of PROs in the NQRs include the availability of validated and reliable relevant measures as well as the feasibility of using such measures consistently and repeatedly across the country.
Results: PROMIS satisfies a major part of the NRQs need of PROs. The PROMIS domain specific approach has advantages in that it allows comparisons within and across different diagnoses. This facilitates, for example, comparisons, meta-analyses and health policy. Further, the PROMIS item bank approach allows for flexible administration and different modes of administration, and as the aim of the international PROMIS initiative is to implement common metrics across countries PROMIS will facilitate international comparisons.
Conclusions: NQRs in Sweden promote the use of PROs in clinical practice as part of the process of setting goals and objectives for treatment and monitoring response to treatment as well as shared decision-making in clinical encounters, as a basis for individual care plans. If the NQRs adopt PROMIS, this would mean that the major part of the Swedish healthcare system would be involved, forming a solid base for quality improvement and research. In the evaluation, PROMIS will be compared with other alternative instruments, and the pros and cons of all options are assessed.

Assessment of Health Related Quality of Life in Psoriatic Arthritis using NIH-PROMIS

Ana-Maria Orbai MD MHS [1], Susan J. Bartlett PhD [1,2], Trisha Duncan MD1, Michelle Jones [1], Clifton O. Bingham III MD [1]

Presenting Author: **Ana-Maria Orbai, MD MHS**, aorbai1@jhmi.edu

Institutional Affiliations: [1] Johns Hopkins University Division of Rheumatology, Baltimore, USA; [2] McGill University Montreal, Canada

Objective: Psoriatic arthritis (PsA) impacts health related quality of life (HRQL) manifesting with periods of exacerbation or flares. There is limited multidimensional characterization of PsA impact on HRQL. PROMIS is an IRT calibrated item bank covering physical, mental, and social health with population-normalized T scores (Mean 50, SD 10). PROMIS evaluation in specific diseases is limited and not studied in PsA. Our objective was to gather preliminarydata on HRQL in PsA patient-reported flare using multiple PROMIS measures.
Methods: PsA patients completed 10 PROMIS computer adaptive tests (CAT) on a tablet: pain interference, fatigue, physical function, sleep disturbance/impairment, depression/anxiety/anger, ability to participate social, satisfaction social roles. Routine PROs (patient global/pain/fatigue VAS, mHAQ, flare self-report) and physician assessments (tender/swollen joints, physician- global) were collected.
Results: We report on the first 42 patients: White 90%, Female 63%, Mean (SD) age 49(11) years, PsA duration 14(9) years. In the overall PsA population [disease activity: mean CDAI 7.6(9)] PROMIS pain interference, fatigue, function and sleep were worse than population norms. Patients reporting flare (n=15) had worse scores across all PROMIS domains versus no flare (n=27), statistically significant for most domains (Table available upon request).
Conclusions: These preliminary data on PROMIS CAT measures of physical, mental, and social health in PsA, which provide population normalized estimates, suggest patients with PsA suffer impairment in HRQL areas encompassing pain, fatigue, function, sleep, mood and participation. PROMIS CATs may help differentiate flare vs no flare states in PsA. Our results highlight impact of PsA on HRQL areas not traditionally assessed in clinical care and research.

A Pilot Study of the Effects of a Resiliency Program on Promis-29 Measures

Elyse R. Park PhD[1,2], Darshan Mehta MD[1,2], Roberta Goldman BA[1], Emma Chad-Friedman BA[1]

Presenting Author: **Elyse R. Park**; epark@mgh.harvard.edu

Institutional Affiliations: [1]Massachusetts General Hospital [2]Harvard Medical School

Objective: The Relaxation Response Resiliency Program (3RP) is an 8-week, evidence-based, mind-body group intervention, developed at the MGH Benson-Henry Institute (BHI) for Mind Body Medicine, which integrates aspects of cognitive behavioral therapy, positive psychology, and relaxation strategies to promote adaptation to stress and enhance resiliency. The core of the program is the elicitation of the Relaxation Response, a physiological state of rest counter to the stress response that is modulated by decreased activation of the sympathetic nervous system and manifests as a decrease in breathing rate and heart rate. At the BHI we have integrated an electronic data collection system that assesses pre and post treatment health status for all clinic patients. The objective of this study was to determine whether the 3RP improves patient health status as measured by PROMIS-29 instruments.
Methods: At the BHI clinic between May 15 and November 25, 2014, the PROMIS-29 was

administered to 103 patients who were participating in the 3RP, before and after the intervention.
Results: Of 103 patients, 25 completed the pre and post questionnaires. In order to determine representativeness of this subgroup, we compared the baseline scores of the subgroup to the full 103 participants and found that the mean scores for the groups did not differ except with respect to pain disturbance (mean difference= 1.9, p=.02) and pain intensity (mean difference=1.4, p=.011), both of which were higher in the subgroup. Using t-tests, we compared pre-post intervention scores on PROMIS-29 measures including Anxiety, Depression, Fatigue, Physical Function, Sleep Disturbance, Satisfaction with Participation in Social Roles, Pain Intensity, and Pain Interference. Symptoms of Anxiety (p=0.05, Cohen's d=0.45), Depression (p=0.031, Cohen's d=0.39), Sleep Disturbance (p=0.01, Cohen's d=0.72), and Fatigue (p<0.001, Cohen's d=0.74) improved significantly from pre to post intervention. No significant difference was found from pre to post on Physical Function, Pain Interference, or Pain Intensity, although Satisfaction with Participation in Social Roles showed a trend toward improvement (mean increased from 14.2 to 15.9, p=0.09).
Conclusion: In accordance with the PHO Clinical Practice theme, we integrated a data collection system utilizing the PROMIS-29 to assess health status changes among patients attending a group resiliency treatment program. After the 3RP program, patients reported statistically significant improvement in their experiences of psychological distress. Patients also showed significant improvement in stress-mediated physical symptoms and behaviors, in particular sleep disturbance and fatigue. In keeping with the PROMIS' goal to assess the patient reported efficacy of real world treatments, these preliminary data support that efficacy of the 3RP in improving psychological, behavioral, and physical symptoms associated with stress.

Changes in clinical and patient- reported outcome measures during a post-acute coronary event rehabilitation program in Uruguay

Laura Ramos, Roberto Ricca , Gerardo Amilivia , Alejandra Maccio, Juan J. Dapueto

Presenting Author: **Laura Ramos, MD**
lramos@hc.edu.uy

Institutional Affiliations [all]
Hospital de Clínicas, Facultad de Medicina, Universidad de la República, Uruguay

Objective: To assess the impact of a multidisciplinary cardiovascular rehabilitation program (MCVRP) on patients early after acute myocardial infarction (AMI) using clinical and patient reported outcomes.
Methods: The sample consisted of patients enrolled in the MCVRP 10 days after *ST-elevation Acute Coronary Syndrome (STE-ACS), who had undergone anglioplasty, had a negative ergometry for ischemia, and normal left ventricular ejection fraction; between* March 2013 – September 2014. The MCVRP consisted of 30 sessions of physical exercise thrice a week, cardiovascular evaluation, and at least three psychosocial interviews. The assessment consisted of clinical data, EKG, ergometry (METS), and echocardiogram. Quality of life was assessed with a set of PROMIS instruments consisting of PROMIS Global Health Measures (PGHM), Emotional Distress – Anxiety (ED-A) and Depression (ED-D), Emotional Distress – Anger, Emotional Support, Emotional Social Isolation, Physical Functioning. Family APGAR was included. Only preliminary data on PGHM, ED-A, ED-D will be presented here. Wilcoxon signed ranks test for paired sample was used for analyses
Results: Fifty eight percent of eligible patients participated in the study (21/36), 83% males, 33% married, mean age 59.6 (range= 40-77, SD: 9.2 years). Functional capacity improved (∆ 1.23 METS, p=.003, effect size .76) as well as PGHM Physical

(Δ 5.76 points, p=.000, effect size .90), PGHM Emotional (Δ 2.81points, p=.000, effect size .56), ED-A (Δ −5.07points, p=.000, effect size .55), ED-D (Δ −2.77points, p=.000, effect size .37)
Conclusions: This is a pilot study on the implementation and impact of a cutting-edge multimodal rehabilitation program for patients early after (STE-ACS) carried on by a multidisciplinary team. Significant changes were observed based on clinical and ergometric data pre and post intervention. PGHM, ED-A and ED-D proved to be useful instruments to assess changes in physical and mental well-being in association with clinical measures.

Translation of eight Pediatric PROMIS® item banks into Spanish and German

Prof. Dr. Ulrike Ravens-Sieberer, MPH [1], Dr. Veronika Ottová-Jordan, MPH [1], Janine Devine, PhD [1], JeanHee Moon, PhD [2], Dr. med. Lara Philippa Gerigk [1], Franka Metzner [1], Prof. Michael Herdman, PhD [3], Margarita P. Hurtado, PhD, MHS [4], Graciella Castillo [5], Anne-Catherine Haller [1], Helena Correia [6], Christopher B. Forrest, MD, PhD [2]

Presenting author: **Ulrike Ravens-Sieberer,** MPH; ravens-sieberer@uke.de

Institutional Affiliations
[1] University Medical Center Hamburg-Eppendorf, Hamburg, Germany [2] Children's Hospital of Philadelphia, Philadelphia, PA, USA [3] Insight Consulting & Research, Barcelona, Spain [4] Independent Consultant, Columbia, MD, USA [5] American Institutes for Research (AIR),Washington, DC, USA [6] Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Objective: To describe the German and Spanish translation of eight PROMIS pediatric item banks: Physical Activity (PAC), Subjective Well-Being (SWB), Experiences of Stress (EOS), and Family Relations (FAM) covering Strength Impact, Physical Activity, Positive Affect, Life Satisfaction, Meaning and Purpose, Family Involvement, Family Belonging, Psychological Stress Experiences, and Physical Stress Experiences.
Methods: We followed the methods outlined in the current ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Task Force recommendations. Ten professional translators (two translators each for German, Swiss, Austrian, US and European based Spanish) generated forward translations (FTs) of each item bank. FTs were compared within a country and cross-culturally to identify problems and to produce a consensus translated version, which was then back translated (BT), evaluated and revised again, if necessary. The items underwent cognitive interviews(CI) in 126 children (55 German, 3 Austrian,2 Swiss, 38 US-Spanish, and 28 Spanish) before finalization.
Results: Eight pediatric PROMIS® item banks were successfully translated covering: PAC (22 items), SWB (125 items), EoS (45 items), and FAM (66 items). The average translation difficulty rating was 1.08 to 1.54 on a 1 (easy to translate) to 3 (difficult to translate) scale. The percentage of item revisions to ensure conceptual equivalence and comprehensibility ranged from none to 18.4%. The cognitive interviews showed that 91% of the final items were appropriate for children aged 8 to 17 years.
Conclusions: German and Spanish translations of the eight PROMIS Pediatric item banks were created for use in clinical trials and routine pediatric health care. The translation methodology helped ensure conceptual equivalence and comprehensibility. Next steps will include linking the item banks to the English versions to create internationally normed and comparable PRO metrics.

Integrating health-related quality of life assessment in oncology practice: sample plan for practitioners

Renee M. Retton, B.S. [Anticipated M.S., May, 2015]

Presenting Author: **Renee M. Retton, B.S**; rettonr@my.canisius.edu
Institutional Affiliations Canisius College, Buffalo NY

Objective: To demonstrate steps practitioners can take to assess: current culture of patient-based care, facility demographics, available internal and external resources accessible to patients, and initiation of HRQOL incorporation, monitoring, and evaluation of surveying as a basic component of care. Thus, opening lines of communication, identifying patient needs, facilitation of individualized care plans, and fostering transparency across the practice [1]. To gauge patient satisfaction of care provided [2].
Methods: Using valid and reliable HRQOL questionnaires in standard oncology practice that can be implemented for any stage or type of cancer within the practice; regardless of facility type. This sample work plan speaks to oncology practitioners in a way that makes assessing, implementing, and evaluating progress feasible while filling a predetermined gap exhibited in peer-reviewed literature in modern oncology care.
Results: Depiction of planning stages including logical steps engrained in evidence-based medicine assist in selecting an appropriate tool, use of the tool, evaluating patient progress or impediments necessitating interventions, and performing aggregate snapshots within the practice at pre-determined points in time. The example HRQOL tool selected for the sample work plan illustration was the FACT-G (Version 4) for primary aim/objective post capturing individual and facility demographic baseline data. IBM SPSS and other avenues of data analysis were provided [1]. The secondary objective of data driven patient satisfaction surveying was accomplished through the display/use of FACIT-TS-G (Version 1), [2]. This sample work plan includes extensive background to support incorporating instruments currently used in clinical trials on an individual basis in standard care practices (oncology), where general HRQOL seems impossible to assess the subscales of: physical well-being, social/family well-being, emotional well-being, and functional well-being; which pertain directly to the domain of health-related-quality of life concepts of: perceived self-efficacy, social supports, and coping abilities in cancer survivors/patients.

Developing the fibrodysplasia ossificans progressiva-physical function questionnaire for adults: intelligent test design using a PROMIS item bank

Dennis Revicki, PhD [1], Maria S. Mattera, MS [1], Frederick S. Kaplan, MD [2], Robert J. Pignolo, MD, PhD [2], Donna Grogan, MD [3]

Presenting Author: **Dennis A. Revicki, PhD**; dennis.revicki@evidera.com

Institutional [1] Evidera, Bethesda, MD Affliations [2] University of Pennsylvania, Philadelphia, PA
[3] Clementia Pharmaceuticals, Newton, MA

Objective: Fibrodysplasia Ossificans Progressiva (FOP) is a rare and disabling genetic condition of progressive extraskeletal bone formation. Physical functioning declines significantly as FOP progresses. The objective was to develop a measure of physical function (PF) in adults with FOP. The PROMIS item banks provide a rich source of item content for developing disease targeted outcome measures.
Methods: We reviewed the Patient Reported Outcomes Measurement Information System (PROMIS) PF item bank

for relevant items for FOP, and 44 PF items were identified. We then conducted concept elicitation (CE) interviews in 21 patients diagnosed with FOP (with varying levels of disease severity) who attended the International FOP Association meeting. The selected PF items were administered after the CE interviews. Interview data were analyzed to identify categories of physical functioning that were impacted by FOP. Based on the CE findings and PF item data, 26 items were initially selected for the new measure. Clinical experts in FOP reviewed the proposed set of items. Five additional items were incorporated into the draft measure, and cognitive interviews (CIs) were conducted in 10 patients, and revisions were made to the final FOP-Physical Function-Questionnaire (FOP-PFQ; 28 items).
Results: For the CE interviews, mean age was 30 years (range 16–54) and 58% were female. For the CIs, mean age was 31 years (range 16–57) and 50% were female. CE interviews demonstrated substantial impacts of FOP on mobility, upper extremity function, and related activities. The CE findings, PROMIS PF item descriptive data, and discussion with clinical experts resulted in 31 relevant items which were included in the draft FOP-PFQ. Based on the CIs, the majority of patients understood the instructions, questions, and response scales; three items were deleted due to redundancy or item removal from the original PROMIS item bank. The final FOP-PFQ contains 28 items covering mobility, upper extremity function, and transferring between various positions (e.g., lying in bed to standing).
Conclusions: This qualitative research supports the content validity of the FOP-PFQ and illustrates the application of PROMIS item banks for efficient new instrument development in an ultra-rare and disabling genetic disease.

Measuring Physical Functioning in Children with Respiratory Insufficiency

Angie Mae Rodday, MS [1,2,3], Robert J Graham, MD [4,5], Ruth Ann Weidner, MBA, MRP [1], Michael L McManus, MD, MPH [4,5], Nan E Rothrock, PhD [6], Darren A Dewalt, MD, MPH [7], Susan K Parsons, MD, MRP [1,2,3]

Presenting Author: **Angie Mae Rodday, MS**; arodday@tuftsmedicalcenter.org

Institutional Affiliations: [1] Institute for Clinical Research and Health Policy Studies, Tufts Medical Center [2] Sackler School of Graduate Biomedical Sciences, Tufts Universit [3] Department of Medicine, Tufts University School of Medicine [4] Department of Anesthesiology, Perioperative and Pain Medicine, Division of Critical Care Medicine, Boston Children's Hospital [5] Department of Anesthesia, Harvard Medical School [6] Department of Medical Social Sciences, Northwestern University Feinberg School of MedicinE [7] Division of General Medicine and Clinical Epidemiology, University of North Carolina School of Medicine

Objective: Pediatric muscular and neuromuscular illnesses often result in decreased health-related quality of life, notably in the area of physical functioning. In this population, non-disease- specific physical functioning scales demonstrate substantial floor effects and lack discrimination at lower levels of functioning. Although pediatric neuromuscular instruments exist, they also do not target assessment of lower levels of physical functioning. Thus, we describe the preliminary psychometric performance of two custom parent-proxy physical functioning short forms created using pediatric Patient Reported Outcomes Measurement Information System (PROMIS) item banks among children at risk for low levels of physical functioning.
Methods: Participants were parents of children aged 5-21 years with respiratory

insufficiency due to muscular and neuromuscular conditions. Parents completed single summary items about the child's mental and physical health using the Child Health Rating Inventories. Two custom parent-proxy short forms from pediatric PROMIS Upper Extremity (13 items) and Mobility (13 items) item banks were created and administered. We selected short form items based on item difficulties and relevance to activities of daily living, including use of assistive devices. We excluded potentially insensitive or redundant items. The custom short forms were scored according to algorithms provided by PROMIS. Standardized T-scores have a mean of 50 (SD=10) with high scores indicating better functioning. Psychometric properties, including construct validity, were explored for the custom SFs.
Results: Fifty-seven parents completed the parent-proxy custom short forms. The median Upper Extremity score was 17 (25th-75th percentile: 12, 23); the median Mobility score was 17 (25th-75th percentile: 15, 30). Known groups comparisons showed that those with lower clinician-rated disease severity had better median Upper Extremity (22 vs. 14, $p<0.001$) and Mobility (28 vs. 16, $p=0.004$) scores than those with worse clinical severity. Both Upper Extremity and Mobility scores were higher in the subgroups defined by better physical health item scores ($r=0.28$, $r=0.15$, respectively). Upper Extremity and Mobility scores also were higher among those with better mental health scores ($r=0.35$, $r=0.21$, respectively).
Conclusions: Upper Extremity and Mobility scores were nearly three standard deviations below the PROMIS pediatric calibration population norms. Preliminary psychometrics demonstrated the richness of PROMIS item banks and the potential to more accurately measure this population's physical functioning. Further field testing is planned to assess scale performances in large patient cohorts.

The Future of PRO Development: Using Technology to Advance Patient- Centered Measure Development

Shimon Rura, BA [1], David Cella, PhD [2], Ari Gnanasakthy, PhD [3], Tamara Kear, PhD, RN [4]

Presenting Author: **Ana Maria Rodriguez, PhD MSc PT;**
amrodriguez@patientslikeme.com

Institutional Affiliations:[1] PatientsLikeMe, Cambridge, MA [2] Northwestern University Feinberg School of Medicine, Chicago, IL [3] RTI Health Solutions, Research Triangle Park, NC [4] Villanova University, Villanova, PA

Patient Reported Outcome (PRO) measure development relies heavily on patient input and as a result can be a burdensome process for recruitment and data collection. Therefore, to produce better measures, patients need a seamless vehicle through which they can engage and contribute to the research process. New technologies such as electronic PROs and computer adaptive testing (CAT) have enabled more time efficient designs, but may also present challenges in balancing new, patient-centered methodologies with goals of scientific rigor. As researchers and technologists, how do we maintain this balance to help promote the potential of PROs when methods and technologies continue to advance? How do we continue innovating so development becomes more patient-centered? David Cella will give an overview of PROs and technology, specifically electronic PROs, patient-powered research networks, CAT, IVR, and "Bring Your Own Device," with a view to where technology can be an asset and where it risks becoming a distraction. How can technology help create measures that assess important patient needs in a truly patient-centered way? How do various technology-enabled methods affect measure validity and scientific rigor? Ari

Gnanasakthy will discuss specific scenarios where research and patients converge and how technology can offer the best solution to successfully capture the patient perspective. Shimon Rura, a technologist, will describe the underpinnings of new online approaches to PROs and how they can help elicit better feedback and rapidly reach patients at scale. Tamara Kear will offer her firsthand account of using new, technology-enabled methods to develop the High Blood Pressure Management, Adherence, Attitudes, and Health Behaviors Scale for patients. She will explain how these technologies can empower researchers and practitioners, who might not start off as measurement experts, to contribute to the process of PRO development. Finally, Dr. Cella will summarize these arguments and invite audience discussion.

PRO-triggered Palliative Care Reduces End-of-Life ICU Utilization in Patients with Advanced Cancer

Andrew M. Romano MD, Kristine E. Gade BS, Paul W. Read MD, PhD, George J. Stukenborg PhD, James H. Harrison MD PhD, Josh Barclay MD, Leslie J. Blackhall MD, Patrick M. Dillon MD,

Presenting Author: **Andrew M. Romano, MD;** amr2h@hscmail.mcc.virginia.edu

Institutional Affiliations: University of Virginia Health System, Charlottesville, VA

Objective: End-of-life care for cancer patients varies widely and increased intensity does not result in improvements in mortality or quality of life. Up to 20% of these patients die in an intensive care setting. Prospective studies show that concurrent palliative care programs for patients with advanced cancer improve quality of life and possibly survival, lead to fewer hospital admission and lower medical costs, however palliative care resources are scarce. This study utilizes a phased intervention of a NIH PROMIS based Patient Reported Outcomes tool (UVA MyCourse) measuring cancer-related symptoms and functional status and integrated in the EMR. The objective is to identify at-risk patients in need of increased palliative care services and measure impact on health resource utilization including ICU admission at end of life.

Methods: As part of a 2012 CMS Innovation Award, CARE Track is a 3-year phased concurrent outpatient palliative care program for patients with advanced cancer, which utilizes the MyCourse PRO tool completed at each visit to identify at risk individuals who would benefit from increased supportive care services. In this preliminary analysis, the CARE Track database was used to identify 191 deceased patients who had received care under this program. These cases were compared to 143 patients with advanced cancer who were not enrolled in this program. Individual charts were reviewed for demographic data, cancer characteristics, number of ICU admissions during the last 6 months of life, whether patients were seen in the ICU during their terminal admission, and place of death. These data were analyzed using a chi-square test for categorical variables.

Results: 29 of 191 (15.2%) cases enrolled in CARE Track were admitted to the ICU in the last 6 months of life, as compared to 44 of 143 (30.7%) control patients ($p=0.0008$). 15 of 191 (7.9%) of CARE Track patients were admitted to the ICU in the final month of life, as compared to 36 of 143 (25.2%) controls ($p<0.0001$). 7 of 191 CARE Track patients (3.7%) were admitted to an ICU and expired during the same hospitalization, as compared to 28 of 143 (19.6%) controls ($p<0.0001$).

Conclusions: A concurrent palliative care program utilizing PRO data for cancer patients at the end of life significantly decreases the utilization of intensive care in the final 6 months of life, particularly in the final month. Palliative care also

substantially reduces ICU utilization during terminal admission for patients who expire in the hospital, (which based on prior studies is consistent with patient preferences). Further analyses will look at the relationship between the increased utilization of the PRO component of this intervention and outcomes as data becomes available.

Simple web-based IRT score estimation for common metrics: http://www.common-metrics.org

Felix Fischer[1,2], Sandra Nolte [1,3], Matthias Rose1 [4]

Presenting Author: **Matthias Rose**; matthias.rose@charite.de

Institutional Affiliations: [1] Department of Psychosomatic Medicine, Clinic for Internal Medicine, Charité Universitätsmedizin Berlin, Germany [2] Institute for Social Medicine, Epidemiology and Health Economics, Charité Universitätsmedizin Berlin, Germany [3] Population Health Strategic Research Centre, School of Health and Social Development, Deakin University, Burwood, Australia [4] Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester

Objective: Item Response Theory allows calibrating instruments on the same scale. Such models have been recently used to calibrate the PROMIS measures for Anxiety and Depression on a common scale with other well established instruments. We developed a web-application (http://www.common-metrics.org) to allow IRT-score estimation for such models.
Methods: The web-application was programmed using open source software. It allows IRT-score estimation under the ML, MAP, WLE and EAP approach using item parameters from various published IRT models. Several different priors can be used.
Results: Researchers can easily upload raw data and receive latent trait score estimates, which can later be analyzed in standard statistical software packages. Compared to cross-walk tables, this approach tolerates missing data and takes the full response pattern into account.
Conclusions: http://www.common-metrics.org facilitates the estimation of theta for researchers less familiar with advanced psychometric techniques. Additional IRT models, e.g. correcting for DIF in different age-groups or languages, can be easily added. Simple scoring on common metrics is an important step towards the establishment of instrument independent scales.

Validation of PROMIS Mobility Computer Adaptive Test and Short Form in Patients with Orthopedic Trauma

Nan Rothrock, PhD [1], Mark Vrahas, MD [2], Robert V. O'Toole, MD [3], Sarah Buono [1], Suzanne Morrison [2], Aaron Kaat, PhD [1], and Richard Gershon, PhD [1]

Presenting Author: **Nan Rothrock**, PhD; n-rothrock@northwestern.edu

Institutional Affiliations:[1] Northwestern University, Chicago, IL [2] Harvard Medical School Orthopedic Trauma Initiative, Boston, MA [3] University of Maryland School of Medicine, Baltimore, MD

Objective: This study evaluated the validity of the PROMIS Mobility computer adaptive test (CAT) and PROMIS 8-item Physical Function short form (PROMIS PF 8a) in a sample of 402 patients with an orthopedic trauma isolated to the lower body (pelvis to heel).
Methods: At the time of a scheduled post-treatment clinic visit, participants completed the PROMIS Mobility CAT, PROMIS PF 8a, and legacy measures including the 10-item physical function scale from the RAND-36 (PF-10), Foot and Ankle Ability Measure ADL Index (FAAM) and Sports Subscale (FAAM Sport), Short

Musculoskeletal Function Assessment (SMFA), and UCLA Activity Scale on a tablet computer. Six months later, 120 participants completed a second identical assessment. Surgeons classified the fracture as mild, moderate, or severe.
Results: Participants were middle-aged (mean=45.1, SD=16.9), predominantly White (77.9%), and male (mean=56%). At time 1, injuries were 4.9 (SD=3.8) months old with treatment occurring 4.3 (SD=3.0) months ago on average. Content validity for PROMIS measures was supported through very little missing data (<1.7%) whereas the FAAM and FAAM Sport had high levels of "not applicable" responses (7.9% and 14.7% respectively). At time 1, floor effects were demonstrated for the PROMIS PF 8a, PF-10, and FAAM Sport (6.6%, 5.9%, and 21.3% with worst possible score). No ceiling effects were found. At time 2, no floor effects were observed. However, ceiling effects were found for the Mobility CAT (10%), PROMIS PF 8a (18.5%), PF-10 (9%), SMFA (6.1%), FAAM (14.8%), and FAAM Sport (13.8%). Convergent validity was demonstrated for PROMIS measures with strong correlations with other measures apart from the UCLA Activity Scale (mean r=0.80 for Mobility CAT, mean r=0.84 for PROMIS PF 8a, all p's<0.001). All measures showed significant improvement between time 1 and 2 (Mobility CAT effect size d=0.81, time 1 mean=35.5, SD=8.5, time 2 mean=43.7, SD=7.9; PROMIS PF 8a d=0.88, time 1 mean=34.2, SD=9.1, time 2 mean=44.7, SD=9.1). Known groups validity based on fracture severity was demonstrated at time 2 with significant differences between mild and severe groups for the Mobility CAT ($F(2, 117)=5.03$, $p<.01$) and PROMIS PF 8a ($F(2, 116)=4.59$, $p=0.01$). Additionally, PROMIS scores varied by the degree of restriction on weight bearing activities (all p's<.01). Strong correlations were found between change scores for PROMIS and other multi-item measures (mean r=0.53 for Mobility CAT, mean r=0.68 for PROMIS PF 8a). Completion times for the Mobility CAT were fast (median 0.9 min, 90% of patients answered <5 items), particularly compared to longer legacy instruments (median=4.3 minutes for SMFA, 1.9 minutes for FAAM).
Conclusion: The PROMIS Mobility CAT and PROMIS SF 8a performed as expected demonstrating physical functioning about 1 • standard deviations worse that the population mean post treatment. Both measures demonstrated validity in orthopedic trauma patients. All measures demonstrated floor or ceiling effects. As the PROMIS item banks can be extended with the addition of items at the extremes of physical function, it may be most appropriate for clinical samples showing large changes in function.

Validation of Patient-Reported Outcomes Measurement Information System (PROMIS) Computer Adaptive Tests in Lumbar Stenosis Patients

Patel AA[1], Dodwad SM[1], Bhatt SA[1], Savage J[1], Hsu WK [1], Rothrock N[2]

Presenting Author: **Nan Rothrock**, PhD; n-rothrock@northwestern.edu

Institutional Affiliations: [1]Department of Orthopaedic Surgery [2] Department of Medical Social Sciences Northwestern University Feinberg School of Medicine

Objectives: Lumbar spinal stenosis is a common, neurological condition in patients over the age of 50 that is often treated surgically. PROMIS computer adaptive tests (CATs) offer brief yet precise assessment in a range of domains of self-reported health. The objective of this project is to validate PROMIS CATs in Pain Behavior, Pain Interference, and Physical Function in surgically treated patients with lumbar spinal stenosis against commonly used static outcome instruments including Zurich Claudication Questionnaire (ZCQ; Pain and Physical Function scales), Oswestry

Disability Index (ODI), and the Short-Form 12 (SF-12; Physical Component Score [PCS]).
Methods: PROMIS Pain Behavior, Pain Interference, and Physical Function CATs, ODI, ZCQ, and SF-12 were administered to 92 consecutive tertiary hospital patients treated surgically for lumbar spinal stenosis. Assessments were completed prior to surgery (T1) and postoperatively at 6 weeks (T2) and 3 months (T3).
Results: Of the 92 patients enrolled (mean age = 62.2, SD = 13.6; 63% male), 65% completed the T3 assessment. At T1, PROMIS means were 1 to 1.5 SDs worse that the general population at baseline (Pain Behavior = 60.3, SD=4.8, Pain Interference = 64.5, SD=7.2, Physical Function = 34.9, SD=6.1). At T1, PROMIS CATs had moderate to strong correlations with the ZCQ Pain ($r=0.60, 0.65, p<0.01$ for Pain Behavior and Pain Interference respectively), ZCQ Physical Function ($r= -0.60, p<0.01$ for Physical Function), ODI ($r=0.73, p<0.01$ with Pain Interference), and SF-12 PCS ($r= -0.34, -0.42, 0.50, p<0.01$ for Pain Behavior, Pain Interference and Physical Function respectively). Grouping patients by ODI severity, PROMIS CAT scores varied in the expected direction (effect size 0.53 – 1.20, $p<0.001$). All instruments showed improvement between T1 and T2 as well as between T1 and T3 (all p's <0.001). Patients reported 7.7 to 10.9 point improvement in PROMIS scores between T1 and T3 ($p's<0.001$). Patients who reported doing "much better" at T2 had 6.7 to 10.4 point improvements in PROMIS scores whereas those reporting doing "slightly better" to "much worse" had minimal score changes (0.4 to 2.7; all p's <0.05). Patients categorized as improved with the ODI and ZCQ at T2 had better PROMIS and PCS scores (SRMs = 0.61 to 1.35) compared with patients categorized as unchanged (all p's <0.05). At T1, none of the measures demonstrated floor or ceiling effects. PROMIS CATs were completed quickly (mean=1.0, 0.8, 0.8 minutes for Pain Behavior, Pain Interference, and Physical Function respectively). Other measures required more time (ZCQ=3.6 minutes, ODI=3.1 minutes, SF-12=3.1 minutes).
Conclusions: PROMIS CATs are valid and responsive measures of surgically treated patients with lumbar spinal stenosis. They perform at least as well as legacy instruments in detecting change over time in addition to being more efficient.

Choosing between the PROMIS Global and EQ-5D for Comparative Effectiveness Research: Are they really different?

Shannon K. Runge, MA [1], Benjamin M. Craig, PhD [2]

Presenting Author: **Shannon K. Runge, MA**; skrunge@mail.usf.edu

Institutional Affiliations: [1] University of South Florida, Tampa, FL [2] Moffitt Cancer Center, Tampa, FL

Objective: To inform the design of comparative effectiveness studies (CER), a head-to-head psychometric comparison of the Patient-Reported Outcomes Measurement Information System (PROMIS) Global and EQ-5D instruments is needed.
Methods: In 2013, 2289 US adults completed an online survey including 10 items from the PROMIS Global and 5 items from the 3 EQ-5D versions (EQ-5D-3L, EQ-5D-Y, andEQ-5D-5L) in random order. After testing for unidimensionality between each pair of these 25 items, we conducted 3 separate exploratory factor analyses (EFA) between the PROMIS Global and each EQ-5D version. Next, we performed an item-response-theory (IRT) analysis for factors shared between the 2 instruments. Item levels with insufficient responses (n<40) were collapsed. The relative degree of information obtained by each instrument was assessed using an unconstrained graded response model.
Results: All items were positively correlated (0.32 to 0.98). Regardless of EQ-

5D version, EFA analyses identified 3 factors (eigenvalue > 1): Physical Health (PH), Mental Health (MH), and Quality of Life (QL). Each item uniquely loaded to a single factor after rotation. Unlike PH and MH, QL included only PROMIS items. At a threshold of 0.5 in standard error, the IRT analyses showed similar PH information function ranges by instrument (-0.2 to 2.8 for EQ-5D-5L vs. -0.2 to 2.4 for PROMIS Global). However, the MH information range for EQ-5D-5L was substantially narrower than the range for the PROMIS Global (0.3 to 2.4 vs. - 0.8 to 2.8).

Conclusions: The PROMIS Global includes 5 items that extend the measurement of general health beyond the 2 factors shared with the EQ-5D. When comparing the remaining 5 items of the PROMIS Global to the EQ-5D, the instruments appear to share information ranges in PH, but the PROMIS Global has a broader MH range than the EQ-5D. These similarities and differences are important considerations when choosing between the PROMIS Global and EQ-5D for CER.

Creating Condition-Specific Patient Reported Outcome Measures

Karen Schifferdecker PhD MPH [1], Karen Kaiser PhD2, Anna Adachi- Mejia1, Susan Yount PhD [2], Dave Cella PhD [2], Kathleen Carluzzo, MS [1], Elliott Fisher MD MPH [2]

Presenting Author: **Karen Schifferdecker, PhD, MPH** karen.e.schifferdecker@dartmouth.edu

Institutional Affiliations: [1] Geisel School of Medicine at Dartmouth, [2] Northwestern University

Objective: To develop a generalizable methodology to create tailored PROMIS Condition-Specific Impact Assessments (PROMIS-CSIA) that integrate generic and condition-specific patient reported outcome (PRO) measures. We are starting with two high-impact conditions: osteoarthritis of the knee (OA-K) and heart failure (HF).

Methods: We are using a mixed-methods approach to develop the instruments: (1) qualitative methods (focus groups, individual interviews, content analysis) to identify condition-specific health impacts important to patients and clinicians, select relevant existing PROMIS items, develop new domains and items for those impacts not covered by PROMIS, and construct expanded PROMIS-CSIA for OA-K and HF, and perform cognitive testing on new items; and (2) quantitative methods to evaluate the measurement properties (validity, reliability, precision, sensitivity, relevance) of the expanded PROMIS-CSIA.

Results: Our analysis of qualitative data from 16 patient focus groups and 17 physician interviews resulted in identification of 32 potential new items for HF (development of potential new items for OA-K is ongoing but will be available by the time of the conference). Potential new HF items undergoing cognitive testing cover symptoms (e.g., swelling, chest pains), the quality of life impact of diet restrictions, anxiety related to one's health condition, health behavior outcomes, and fatigue. Deliberations regarding quantitative data collection, resulted in a two-prong approach to increase efficiency of data collection while evaluating measurement properties: a cross-sectional panel-company sample for psychometric assessment, scale stability, and item parameter stability; and a longitudinal clinical sample to conduct preliminary validity and responsiveness assessments. Preliminary cost estimates for development of condition-specific instruments will be shared.

Conclusions: The methodology will produce a process for creating condition specific instruments that build on and add to PROMIS item banks for both clinical and research applications. It will fill a critical gap; specifically, the current lack of a generalizable model to create generic and

condition-specific measures that allow the full impact of specific conditions to be assessed from the patient's perspective. The expansion of the scope and utility of existing generic measures for key conditions will lead to greater uptake and support for routine and systematic collection of critical patient-centered outcomes in clinical practice, by disease registries, and in comparative effectiveness research.

PROMIS®29: A tool for improving mental healthcare and wellbeing of US Peace Corps Volunteers

Priscilla M. Schulz, MSW, LCSW-C [1], Susan Henderson, MD, MPH [2], Megan Brown [1]

Presenting Author: **Priscilla M. Schulz**, MSW, LCSW-C; pschulz@peacecorps.gov

Institutional Affiliations: 1] United States Peace Corps Office of Health Services Counseling and Outreach Unit, Washington, DC [2] United States Peace Corps Office of Health Services Epidemiology Unit, Washington, DC

Background: In 2014 approximately 6900 Peace Corps Volunteers were engaged in international development and citizen diplomacy work in 63 countries. To support their health and wellbeing, the US Peace Corps employs approximately 132 US- and foreign-trained Medical Officers (PCMOs). As Volunteers' primary healthcare providers, PCMOs play a critical role in assessing emotional problems and determining need for mental health intervention. Medical training in psychiatry differs around the world. This difference results in a range of knowledge, experience and skills among the US Peace Corps' PCMOs. Language and culture underlie expression and identification of illness and well-being. For most PCMOs English is a second language.

Objectives: Because variations in professional training and language skill level may affect detection and prioritization of mental health issues, the US Peace Corps adopted PROMIS®29 in 2014 as part of ongoing efforts to improve the quality of Volunteer medical care. As a validated, brief, objective measure of health-related quality of life (QOL), PROMIS®29 augured increased standardization of mental health assessments and minimization of the impact of psychiatry expertise or cross-cultural differences. Three improvements were anticipated: 1. Earlier detection of latent, clinically relevant problems; 2. Improved prioritization, management and outcomes evaluation of mental health cases; 3. Identification of common QOL vulnerabilities in Peace Corps Volunteers from aggregated PROMIS®29 data gathered as a service entry baseline measure (future) to be used to strengthen Volunteer resilience training.

Methods: In 2014 the US Peace Corps implemented a policy change requiring that PCMOs to administer PROMIS®29 during mental health assessments. Trainings taught PCMOs how to use PROMIS®29: • to identify focus of care; • to normalize distress, or consider psycho- education to address specific issues (e.g., sleep hygiene information with the discovery of sleep disturbances), • to assess impact of supportive counseling; • to identify individuals needing referral for professional psychological intervention.

Results: Peace Corps took several steps to incorporate PROMIS®29 into standard mental healthcare. These included: • piloting PROMIS®29 in the field at 3 Peace Corps medical units; • based on field recommendations, revising the paper/pencil PROMIS®29 to improve ease of scoring, and understanding of measure results; • large group trainings at Peace Corps' 2014 Continuing Medical Education conferences; individual, on-the-job tutorials. Trainings included didactics, experiential exercises in which PROMIS®29 use is rehearsed, and distribution of key PROMIS®29 academic papers. Preliminary results suggest PROMIS®29 implementation

is achieving anticipated objectives. Evaluation is ongoing.
Conclusion: Continued medical education will be needed to maintain meaningful ongoing use of PROMIS®29 in the US Peace Corps' medical program.

Pain severity and pain interference associated with psychosocial factors

Annalisa F. Sciullo, B.A.[1], Fred Foley, Ph.D.[1,2] & Elizabeth K. Seng, Ph.D.[1,3]

Presenting Author: **Annalisa F. Sciullo**; annalisa.sciullo@gmail.com

Institutional Affiliations: [1]Ferkauf Graduate School of Psychology, Yeshiva University, Bronx, NY [2 3]Saul R. Korey Department of Neurology, Albert Einstein College of Medicine of Yeshiva University

Objective: Multiple Sclerosis (MS) is a chronic inflammatory, demyelinating neurological disease of the central nervous system, which can result in many debilitating symptoms including pain. Chronic pain is comorbid with MS, and can interfere with quality of life. According to the biopsychosocial model of pain, modifiable psychosocial factors can aide in the maintenance and exacerbation of the pain. Few studies have examined relationships between pain interference in people with MS and pain-relevant psychosocial factors, such as kinesiophobia (fear of movement due to pain), catastrophizing, anxiety and depression. This study attempts to add to the limited data available examining psychosocial factors associated with pain interference in people with MS.
Methods: Participants were 26 adults with MS or clinically-isolated syndrome (people who have experienced a single episode of neuronal inflammation or degeneration) recruited from a larger ongoing observational study conducted at the outpatient Multiple Sclerosis Clinic within Holy Name Medical Center in New Jersey. Participants completed questionnaires assessing pain symptoms (pain severity rated on a 0-10 numeric rating scale, and the PROMIS pain interference scale), pain-related cognitions [Tampa Kinesiophobia Scale (TSK) and Pain Catastrophizing Scale (PCS)], and psychiatric symptoms (Hospital Anxiety and Depression Scales; HADS). Due to the non-normal distribution of the symptom variables, Spearman's rho was used to examine correlations between pain severity and interference and measures of pain-related cognitions and psychiatric symptoms.
Results: Higher pain severity was strongly related to higher pain interference, $r = .70$, $p < .001$. Higher kinesiophobia was strongly associated with higher pain interference, $r = .59$, $p = .002$, but was not significantly associated with pain severity, $r = .30$, $p = .133$. Higher catastrophizing was strongly associated with higher pain interference, $r = .50$, $p = .009$, and higher pain severity $r = .54$, $p = .004$. Higher anxiety was associated with higher pain interference, $r = .44$, $p = .024$, and higher pain severity $r = .57$, $p = .002$. Higher depression was strongly associated with higher pain interference, $r = .64$, $p < .001$, and higher pain severity $r = .52$, $p = .007$.
Conclusions: Pain severity and pain interference are highly correlated with psychosocial factors. Thus, treatment for pain in MS may benefit from targeted treatment of psychosocial factors.

A comparison of calibrated projection with conventional IRT based linking

Eisuke Segawa, PhD [1], Li Cai, PhD [2], Benjamin D. Shalet, PhD [1], David Cella, PhD [1]
Presenting Author: **Eisuke Segawa**; eisuke.segawa@northwestern.edu

Institutional Affiliations [1] Northwestern

University, Chicago, IL[2] University of California Los Angeles, Los Angeles, CA

Objective: Compare a recently developed IRT based linking method, calibrated projection (Thissen et al., 2011), and the conventional IRT based approach through linking of GAD-7 with PROMIS anxiety.
Methods: A two-dimensional item factor model, in which PROMIS anxiety items load on one factor and GAD-7 items load on another, was used in calibrated projection. A one-dimensional item factor model was used in the conventional IRT approach. All analyses were conducted using IRTPRO (Cai and DuToit, 2012).
Results: The two-dimensional IRT model fitted the data much better than one dimensional IRT. The difference of PROMIS T-score between the calibrated projection and the conventional IRT was largest (2.3) at the smallest GAD-7 summed score of 0. The difference gradually decreased as the summed score increased, to 0 at the summed score of 12 (which is linked to PROMIS T-score of 62.3) and -2.0 at the largest GAD-7 summed score of 21. Standard Errors of linking in calibrated projection were larger than those using conventional IRT.
Conclusion: Because the differences were not trivial, further investigation of calibrated projection, e.g., applying it to more instruments and examining it through simulation, is needed.

The burden of itch in the primary care setting.

Jonathan I. Silverberg, MD, PhD, MPH [1] , Keiki Hinami, PhD and David Cella, PhD

Presenting Author: **Jonathan Silverberg, MD, PHD;** jonathanisilverberg@gmail.com

Institutional Affiliations: [1] Northwestern University Feinberg School of Medicine, Chicago, IL USA [2] Cook County Health and Hospitals System, Chicago, IL, USA [3] Northwestern University Feinberg School of Medicine, Chicago, IL USA

Importance: Itch is a common symptoms encountered in dermatology and other specialties. However, little is known about the epidemiology and patient-burden of itch in primary care.
Objectives: We sought to determine the prevalence and impact of itch on quality of life (QOL) in primary care.
Design: Cross-sectional study using audio computer-assisted self-interview.
Setting: Outpatient general internal medicine clinic.
Participants: 2,076 adults (age >18 years).
Exposure: History of itch in the past week.
Main Outcome Measure: Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health 10-item, 17 physical symptoms from the Memorial Symptom Assessment Scale and Patient Health Questionnaire (PHQ-2).
Results: The prevalence (95% CI) of itch was 39.9% (37.7–42.0%) and increased with age from 33.1% (25.4–40.8%) at age 19-39 years to 45.9% (35.3–56.5%) at age ≥80 years. In multivariable models controlling for age, gender, race/ethnicity and langue of interview, even being a little or somewhat distressed by itch was significantly associated with lower PROMIS Global total raw scores, physical and mental health T-scores and EQ5D scores overall, particularly poorer overall health, QOL, physical health, mental health and satisfaction with social activities and relationships, inability to carry out every day physical activities and usual social activities and roles, more severe fatigue, and being bothered by emotional problems (P≤0.01 for all). Further, quite a lot or very much distress from itch was associated with higher odds of depressed mood (4.91 [3.36–7.18]) and anhedonia (4.46 [3.07–6.47]). The overall patient-burden of itch was similar to that of pain, constipation, sexual dysfunction, cough and weight loss.
Conclusions and Relevance: Itch occurs

commonly in the primary care setting and is associated with poor quality of life. Physicians should inquire about itch during review of systems and treat accordingly. Given its importance, development of a PROMIS itch item bank is underway.

Patient Reported Outcomes Measurement Information System (PROMIS) measures in pediatric adenotonsillectomy patients at ambulatory surgicenters

Allan F. Simpao, MDa, Elicia C. Wartmana, Jorge A. Galvez, MD,a Arul Lingappan, MD,a Luis M. Ahumada, MSCS,b Abbas F. Jawad, PhDc Mohamed A. Rehman, MDa,

Presenting Author: **Allan F. Simpao, MD**; simpaoa@email.chop.edu

Institutional Affiliations: a Dept. of Anesthesiology and Critical Care, The Children's Hospital of Philadelphia, PA, USAb Enterprise Analytics and Reporting, The Children's Hospital of Philadelphia, PA, USAc Department of Biostatistics in Pediatrics, The Children's Hospital of Philadelphia, PA, USA

Objective: The Patient Reported Outcomes Measurement Information System (PROMIS) is an NIH-funded system of highly reliable, precise question-and-answer measures of patient- reported physical, social, and mental well being that can be administered via paper-based or electronic means. Adenotonsillectomy (T&A) procedures are one of the most common procedures performed at our institution, yet systematic measurement of patient-reported health status was not routinely performed. We initiated the pre- and post-surgery administration of PROMIS tools using tablet computers and e-mail to T&A patients at our ambulatory surgery centers.

Methods: A research assistant enrolled patients of ages 5-17 years old on the day of surgery who were scheduled for an T&A or adenoidectomy at an ambulatory surgery center. The PROMIS tools (measuring anxiety, depression, fatigue, physical function and peer relations) were administered to the patients in the preoperative waiting area prior to the procedure using a PROMIS-enabled tablet computer. Patients of ages 5-8 years were encouraged to complete the electronic questionnaires with a parent or guardian proxy. At one week post-procedure, a link to the PROMIS tools was sent via e-mail to the patient's caregiver; if necessary, a research assistant reminded the caregiver to complete the PROMIS assessment via a telephone call. The PROMIS tools results generated both raw and scaled PROMIS scores. The raw PROMIS scores (z-scores) were converted to the scaled PROMIS scores. The scaled scores are based on standard curves in healthy patients, with each curve having a mean of 50 (no units) and standard deviation of 10. We analyzed the PROMIS results were analyzed and generated descriptive statistics following ten months of data collection.

Results: Fifty-six patients completed the pre-procedure and one-week follow-up PROMIS assessments during February 1, 2014 to November 7, 2014. Eighteen patients were in the 9-17 years old age group and completed the PROMIS tools themselves; 38 patients were in the 5-8 year old age group that consisted of caregiver proxy scores.

Conclusions: The PROMIS tools allowed for convenient measurement of patient-reported outcomes in pediatric T&A patients. Upon initial review, no dramatic changes in the six measured outcomes were noted in the 9-17 years age group. There was a 10-point (1 SD) increase in fatigue scores (and a similar decrease in mobility scores) at the 1 week follow up point in patients of ages 5-8 years, and an increase in pain interference scores, suggesting that children in this age group continue to be affected by pain and fatigue one week after surgery. Future plans include continued patient recruitment and administration of the PROMIS tools, more rigorous statistical

analysis, and investigation of the causes of the score changes to determine possible options for amelioration of pain and fatigue symptoms post-procedure.

Does health-related quality of life predict physical activity among African American breast cancer survivors?

Selina A. Smith, PhD, MDiv [1,2], Mechelle D. Claridy, MPH [3], Mary S. Whitehead, MPH [2], Joyce Q. Sheats [1], Benjamin E. Ansa, MD, MSCR [1], Ernest Alema-Mensah, PhD, DMin [2]

Presenting Author: **Selina A. Smith** SESMITH@gru.edu

Institutional Affiliations: [1] Georgia Regents University, Institute of Public & Preventive Health, Augusta, GA [2] SISTAAH Talk Breast Cancer Support Group, Miami, FL [3] Morehouse School of Medicine Cancer Research Program, Atlanta, GA

Objectives: To determine correlations between physical activity (150min/week recommendation) and post-treatment outcomes (physical functioning, fatigue, pain) and investigate the impact of treatment exposure (chemotherapy, radiation, surgery) on health-related quality of life (HR-QoL) and physical activity.
Methods: Two-hundred and forty (240) African American breast cancer survivors completed a 45-minute lifestyle assessment tool capturing physical activity (150min/week recommendation) and HR-QoL using the PROMIS Global Health Scale.
Results: Female breast cancer survivors with ages 18 years and older were identified. A majority (58%) were 55 years and older, married (42%), with college or graduate school education (46%), and presented with stage I cancer at diagnosis (42%); more than half (51%) earned an annual income between $25,000-$49,999, and 57% reported meeting the physical activity recommendation of 150 minutes/week [p=0.0042]. Survivors with college or graduate school education (65%) and those who presented with stage I cancer (62%) met recommended weekly physical activity requirement [p=0.0003; p=0.0288]. Fewer survivors who had chemotherapy (49%) engaged in physical activity compared with those who had none (60%) [p=0.0868], while more of those that had radiotherapy (52%) and surgery (55%) met the required physical activity level [p=0.5622 and p=0.3942]. Most survivors (56%) who participated in physical activity did not report breast cancer recurrence [p=0.4383]. Pain and fatigue did not have a significant impact on physical activity. Eighty eight percent of survivors who had radiotherapy reported having excellent to good physical health [p=0.0080].
Conclusions: Results reveal that education, income and stage at diagnosis were significant determinants of physical activity and more survivors reporting radiotherapy had excellent-good physical health, compared to those reporting chemotherapy and surgery. More effort is needed to develop strategies that will enable and enhance physical activity among breast cancer survivors.

Health Related Quality of Life Among African American Female Breast Cancer Survivors

Selina A. Smith, PhD, MDiv [1], Mechelle D. Claridy, MPH [2], Benjamin E. Ansa, MD, MSCR [1], Francesca Damus, MPH [2], Ernest Alema- Mensah, PhD, DMin [2]

Presenting Author: **Selina A. Smith**; SESMITH@gru.edu

Institutional Affiliations: [1] Georgia Regents University, Institute of Public & Preventive Health, Augusta, GA[2] Morehouse School of Medicine Cancer Research Program, Atlanta, GA

Objective: The purpose of this study was to compare the health-related quality of life (HR-QoL) of African American female breast cancer survivors and survivors of other cancers to African American females with no history of cancer.
Methods: African American female breast cancer survivors (n=62), survivors of other cancers (n=74), and those with no history of cancer (n=1,566) were identified from the 2010 National Health Interview
Survey (NHIS). The PROMIS Global Health Scale was used to assess HR-QoL.
Results: In multivariate logistic regression models adjusting for age, marital status, and education, the overall effect of cancer status was statistically significant for all HR-QoL outcome variables except fatigue;
breast cancer survivors and those with no history of cancer had better physical health t-scores compared to those with other cancers and breast cancer survivors and those with no history of cancer had better mental health t-scores compared to those with other cancers. All results
were weighted.
Conclusions: This study shows that African American female breast cancer survivors reported a fairly good HR-QoL. However, there is an HR-QoL disparity for African American female survivors of other cancers. This information regarding cancer survivorship can be used to assist with the national efforts aimed at increasing HR-QoL for this population

Collection and Use of PROMIS measures for Palliative Care Cancer Patients

George Stukenborg PhD, Leslie Blackhall MD, James Harrison MD, Paul Read MD.

Presenting Author: **George J. Stukenborg.**
Email: gstukenborg@virginia.edu

Institutional Affiliations: University of Virginia School of Medicine. Departments of Public Health Sciences, Medicine, and Radiation Oncology

Objective: Patient Reported Outcomes (PROs) are a key information resource for cancer patient care, beginning with the initial presentation after diagnosis and continuing for some patients through end of life care. This abstract describes ongoing efforts underway at the University of Virginia Health System to collect and utilize PROMIS domain measures and additional PROs for targeted clinical care. The project is funded through a CMS Innovation Challenge Award, which has the objective of improving the quality of care provided to patients with incurable malignancies.
Methods: A tablet computer based PRO collection tool was implemented in 2013 for use with advanced cancer patients. The assessment is implemented using the EpicSystems Corporation Electronic Health Record (EHR) online browser application tool (MyChart). A core set of five PROMIS domains are captured and stored (anxiety, depression, fatigue, pain interference, and physical function) using the version 1.0 short form item series. Additional PROs are collected addressing cancer specific symptoms and measures (bowel function, nausea, vomiting, anorexia, dyspnea, and peripheral neuropathy). Assessment measures are scored and displayed in the EHR with results tracked over time to guide therapy and assess response to treatment. Current assessment results are collected from patients in clinic at each routine follow-up appointment, before their clinical encounter. Current and longitudinal results are available to physicians while patients are in clinic, so that they can be used in real time to guide therapy and assess response to treatment. The project is organized in three phases. Phase 1 is a developmental phase in which the PRO collection capacity was developed. Phase 2 is a first use phase where the PRO data displays were made available for use in the clinic. Phase 3 incorporates additional EHR displays of alerts and related triggers based on

threshold values of the PROMIS and other PRO reported data. Research databases have also been developed to link PRO data with data regarding hospital visits, hospice utilization, survival, utilization of chemotherapy and radiation therapy. The research databases will be used as a resource for the development of statistical models for predicting patient reported outcomes and program evaluation.
Results: As of January 2015 a total of 1,868 assessments were completed by 489 patients, with repeated assessments available for most patients: >1 (69%), >2 (51%), >3 (36%). PROMIS domain median scores (and standard deviations) are as follows: Anxiety (50, 19.6), Depression (30, 19.1), Fatigue (55, 17.6), Pain Interference (53, 24.5), Physical Function (85, 19.2). Analysis of longitudinal trends in PROMIS measures is currently in progress using random effects models.
Conclusions: With PROs available in routine EHR reports, patients with severe symptoms or trend patterns of concern can be referred to appropriate services prior to symptom escalation to the point of requiring inpatient management. The PRO data is currently being used to identify patients for referral to palliative treatments such as radiation therapy, social work, or chaplaincy and also to identify patients with severe or complex problems requiring enhanced care coordination, addressed by a weekly multi-disciplinary team-based care planning meeting.

PROMIS assessment for patients undergoing colorectal surgery in enhanced recovery protocol

Tracy Hedrick MD, Amy Harrigan RN, George J. Stukenborg PhD

Presenting Author: **George J. Stukenborg PhD**

Institutional Affiliations: University of Virginia School of Medicine, Departments of Surgery and Public Health Sciences

Objective: Patient Reported Outcomes measurement provides a unique patient-centered perspective on the impact of colorectal surgery on patient quality of life. Few studies of colorectal surgery outcomes have assessed these types of measures. The absence of this information makes it difficult to inform patients about the near term effects of surgery, beyond the outcomes assessed by traditional clinical measures. This study was designed to provide information about the effects of colorectal surgery among patients in an enhanced recovery program on physical, mental, and social wellbeing outcomes.
Methods: Four domains were selected based on patient consultation and clinical experience: depression, pain interference, social participation, and interest in sex. The NIH Patient Reported Outcomes Measurement Information System (PROMIS®) Assessment Center was used to collect patient responses prior to surgery and at a follow-up postoperative visit occurring within 30 days. Patients undergoing elective colorectal surgery were consented and administered the survey using a tablet computer, with available support provided by a clinical research coordinator. Each domain was measured with computerized adaptive testing based scoring, following a series of questions with hierarchically structured responses on the underlying domain. Key clinical measures (including gender, colon cancer, diverticultis, inflammatory bowel disease, radiation therapy, and erythema) were separately collected for each patient from the Society for Thoracic Surgery patient registry database. Multilevel random coefficient models were estimated to assess the overall direction of scores during the follow-up period, and to assess the statistical significance of differences in trends between patients grouped by key clinical measures.

Results: In total, 90 patients were consented and 169 assessments were completed, with 79 patients completing assessments before and after surgery (87.7%). The mean number of days between assessment dates was 56.3 days. Technical or scheduling difficulties accounted for the majority of follow-up omissions. There was one immediate postoperative death. Overall mean scores and mean differences from baseline to follow-up were as follows: depression (51.20: -3.09), pain interference (55.48: -2.32), social participation (48.83: 0.40), and interest in sex (47.24: 1.19). No statistically significant changes in scores were demonstrated over the follow-up period for any of the four domains, either overall or for differences between patients grouped by key clinical measures.
Conclusions: Colorectal patients reported median pain, depression, social isolation, and sexual dysfunction scores that were nearly equal to those reported by the general population, both at baseline and following surgery. There were no statistically significant changes in scores over the follow-up period for any domain. These results suggest that the majority of patients quickly return to baseline physical, mental and social function following surgery within the follow-up period considered in an enhanced recovery program.

Validation of the Dutch-Flemish PROMIS pain item banks in patients with rheumatoid arthritis

Martine H.P. Crins, Msc [1], Caroline Terwee, PhD [2,3], Rene Westhovens, MD, PhD [4,5], Dirkjan van Schaardenburg, MD, PhD [6], Niels Smits, PhD [3,7], Johan Joly, Msc [4,5], Patrick Verschueren, MD, PhD [4,5], Kristien van der Elstc,d, Msc [4,5] David Cella, PhD [8], Karon F. Cook, PhD [8], Joost Dekker, PhD [9], Maarten Boers, MD, PhD [2], Leo D. Roorda, MD, PhD [1]

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

Institutional Affiliations
[1] Amsterdam Rehabilitation Research Center | Reade, Amsterdam, The Netherlands [2] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands [3] The EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands [4] Department of Development and Regeneration, Skeletal Biology and Engineering Research Center, KU Leuven, Louvain, Belgium [5] Department of Rheumatology, University Hospitals, KU Leuven, Louvain, Belgium [6] Jan van Breemen Research Institute | Reade, Amsterdam, The Netherlands [7] Department of Clinical Psychology, VU University, Amsterdam, the Netherlands [8] Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA [9] Department of Rehabilitation Medicine and Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands.

Objective: To validate the Dutch-Flemish translation of the PROMIS Pain Behavior (DF-PROMIS-PB) and Pain Interference (DF-PROMIS-PI) item banks in Dutch and Flemish RA patients.
Methods: A paper-and-pencil or web-based survey, including the full DF- PROMIS-PB (39 items, 6-point Likert scale) and DF-PROMIS-PI (40 items, 5-point Likert scale), was completed by 826 Dutch and 618 Flemish RA patients. Unidimensionality was evaluated by one-factor confirmatory factor analysis. With the future strategy to develop computer adaptive tests (CAT), item response theory (IRT) models were used to evaluate the item characteristics of the two item banks. A graded item response model (GRM) was fitted and construct validity was studied. To analyse cross- cultural validity, Differential Item Functioning (DIF) was evaluated for e.g. language (Dutch-Flemish vs. English and Dutch vs. Flemish) by ordinal

regression models.
Results: The interim analysis showed that unidimensionality of the DF- PROMIS-PB and DF-PROMIS-PI was supported (CFI=0.975; 0.997 resp. and TLI=0.974;0.997 resp.). The first factor accounted for 49% (DF-PROMIS-PB) and 81% (DF-PROMIS-PI) of the questionnaire variance. Thirteen out of 741 (1.8%) DF-PROMIS-PB item pairs and 20 out of 780 (2.6%) DF- PROMIS-PI item pairs were marked as possibly locally dependent. The data of the two item banks fit the GRM, and showed good coverage across the pain behavior and pain interference continuum. Additional analyses of the DF-PROMIS-PB and the DF-PROMIS-PI are in progress and will be presented at the conference.
Conclusion: The first results indicate that the items of the DF-PROMIS-PB and the DF-PROMIS-PI fit a GRM and demonstrate good coverage across the range of the pain behavior and pain interference domain. The interim conclusion is that the DF-PROMIS-PB and DF-PROMIS-PI can be used to develop a CAT for measuring pain behavior and pain interference in Dutch and Flemish RA patients.

Patient-friendly, web-based, longitudinal administration of PROMIS ools

Vincent Tsugranes (BS, MIS, MS), Pamela Gavin (MBA), Deanna Portero, Casey Morris, Jacqui Kraska

Presenting Author: **Vincent Tsugranes;** vince@rarediseases.org

Institutional Affiliations: National Organization for Rare Disorders (NORD)

Objective: To develop a platform for web-based, patient-centered outcomes driven, longitudinal patient registries, that is ideal for delivering instruments like PROMIS tools.
Methods: Ideas for improvements and enhancements to NORD's platform are regularly elicited from a diverse group of stakeholders, including patients and caregivers, non-profit patient advocacy organizations, clinical researchers, regulators, drug and biologic developers, research coordinators and thought leaders in digital engagement. On a quarterly basis, NORD coordinates a review of potential improvements and enhancements with non-profit patient advocacy organizations that currently use NORD's platform to host patient registries, or plan on using the platform in the near future. The patient organizations are asked to rank the features by desirability. Their rankings are synthesized. NORD's technical team reviews the top ranked features, and prepares an estimate of the expected time to develop, test and deploy each feature. To complete the selection of the next round of enhancements and improvements, the technical team reconciles the feature rankings and the estimated effort against the fixed resource of three months of developer bandwidth.
Results: Response to the NORD platform design and functionality has been strongly positive, which supports the case of centering patient user perspectives when designing PROMIS administration and scoring software. The platform is less than one year old, yet over 100 organizations have reached out to NORD to learn about the NORD Registry Platform and how to begin using it. Despite a competitive market with viable alternatives, approximately 50% of organizations that have seen demonstrations of the platform, and learned about the democratic roadmap have chosen to use the NORD Registry Platform. A list of notable platform features is available upon request.
Conclusion: There is a business case for prioritizing the patient user experience in software design for administering and scoring PROMIS. The NORD Patient Registry Platform development model—an iterative development approach, which leverages a diverse network of patient data stakeholders, and centralizes the priorities and perspectives of patient advocacy organizations in setting the development

roadmap—is an effective manner of improving and enhancing software for administering and scoring PROMIS.

Integrating patient reported outcomes assessment in advanced liver disease

Manisha Verma, MD, MPH; Victor Navarro, MD

Presenting Authors **Manisha Verma** VermaM@einstein.edu; **Victor Navarro** NavarroV@einstein.edu

Institutional Affiliations Einstein Healthcare Network, Philadelphia

Objectives: Patient centered care is a goal of the health care reform, as it improves patient-clinician communication, quality of life, and disease outcomes. The overall goal of this project is to test the feasibility and acceptability of integration of Patient Reported Outcomes (PRO) assessment in patients with advanced liver disease (Cirrhosis and Hepatocellular Cancer/ HCC) in an ambulatory setting. The purpose of this project is to identify if PRO assessment has any influence on patient centered care, through enhanced patient physician communication, increased patient well-being/ health related quality of life, detection of physical or psychological issues, and supporting patient directed clinical decision making.

Methods: This is a prospective non randomized longitudinal study. The PRO assessment as an experimental intervention includes 3 elements: a) Training of physicians (hepatologists and oncologists) to understand the content and interpretation of questionnaire and results, b) Administering the PRO assessment at 3 consecutive time points (1, 3 and 6 months visit) using PROMIS assessment center, and c) PRO results (T-scores) reported to patients and physicians at each point of care for use during the clinical interaction. T-score metric is used to measure the symptom assessed, with a mean of 50 for the US general population. Higher score reflects more of the symptom evaluated. The study instruments include a patient and a provider feedback survey to assess the feasibility and acceptability of PRO assessment in routine clinical practice, and Observing Patient Involvement in Decision Making (OPTION) to quantify and assess patient physician communication. The OPTION score is based on the audio recording of the office visit, and ranges from 0 to 100. Higher score reflects more of patient involvement in decision making. The study assesses PROMIS-29, Alcohol use questionnaire, Chronic Liver Disease Questionnaire and Sexual Function questions from the Short Form-Liver Disease Quality of Life. All these are assessed at baseline, 3 and 6 month visits.

Results: This study is currently active and is in data collection phase; 20 patients have been enrolled since Jan'15. The mean age is 59 years, 14 females and 6 males. 16 have cirrhosis and 4 also have liver cancer. The mean T- scores for enrolled patients at baseline are as follows:- physical function, 43.98; anxiety, 55.3; depression, 51.41; fatigue, 43.98; sleep disturbance, 53.41; ability to participate in social roles, 50.89; and pain interference, 56.75. 95% patients found it useful to fill out a questionnaire to inform their providers about how they feel physically and emotionally, and 90% understood the T-scores. The mean baseline OPTION score is 57.3 (range 45 to 88). Barriers to implementation include: disruption of office flow, and need to address symptoms in addition to disease specific problems. Standard clinical algorithms to follow-up specific cut-off T-scores will be implemented for patients with no improvements over time.

Conclusion: There have been no studies to date assessing the feasibility and acceptability of PRO assessment for patients with advanced liver disease. This is a pilot demonstration project to show its value for advanced liver disease.

The Spanish PROMIS® Depression measures are valid and responsive in a clinical sample of mental health patients

Gemma Vilagut, MSc.[1], Carlos G. Forero, PhD[2], Elena Olariu, MSc.[1], Núria D. Adroher, MSc.[1], Jordi Alonso, MD, PhD[1]

Presenting Author: **Gemma Vilagut**; 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.

Objective: We aimed to assess measurement properties of the Spanish version of PROMIS Depression item bank (IB) and 8-item static short form (SF) in a clinical sample of mental health patients from Spain.
Methods: A 3-month follow-up study was carried out on a sample of patients seeking assistance for active mood/anxiety symptoms (n=233) from primary care or specialized mental health care. Clinician-administered Mini- International Neuropsychiatric Interview (MINI) was used to establish the DSM-IV diagnoses (gold standard): major depression episode (MDE), panic disorder (PD) and generalized anxiety disorder(GAD). In addition to the PROMIS depression Item bank, patients were administered: the Patient health Questionnaire (PHQ-9) self-reported measure for depression and the Beck Anxiety Inventory for anxiety. The 8-item PROMIS Depression short form was scored based on the answers to the full item bank. Hamilton Rating Scales for Depression (HAM-D) and Anxiety (HAM-A) were administered to asses clinical severity. Unidimensionality was evaluated with Confirmatory Item Factor Analysis, using Comparative Fit Index (CFI)>0.95 and Root Mean Square Error of Approximation<0.08 for goodness of fit. We estimated internal consistency reliability (Cronbach's alpha and Lambda-2) and used Test Information Curves for the assessment of reliability throughout the scale continuum. Construct validity was evaluated through correlations with other self-reported and clinical measures, and known-groups validity comparing PROMIS means across MINI diagnostics (No disorder, pure MDE, pure anxiety and comorbid MDE and anxiety). We also assessed 3-month responsiveness to change through Cohen's effect sizes (d) in stable and recovered patients.
Results: The unidimensional model showed adequate fit (CFI=0.97, RMSEA=0.08). Internal consistency for both the item bank and the short form were high (Cronbach'salpha≥0.95). Information Curves had reliabilities over 0.90 throughout most of the score continuum (from almost 1 SD below the reference mean of 0 to around 1.7SD above the mean). As expected, we observed high correlations with external self-reported depression, and moderate with self-reported anxiety and clinical measures. The item bank showed an increasing severity gradient from no disorder (mean=48, SE=0.6) to pure anxiety (mean=50, SE=1.0), pure MDE (mean=54, SE=0.5) and depression with comorbid anxiety (mean=55.8, SE=0.4). PROMIS detected depression disorder with great accuracy (AUC=0.89). Both PROMIS Depression measures (item bank and the 8-item fixed form) showed high responsiveness to change in recovered patients (d>0.8) and small in stable patients (d=0.1).
Conclusions: The good metric properties of the Spanish PROMIS Depression measures provide further evidence of their adequacy for monitoring depression levels of patients in clinical settings. This double check of quality (within countries and populations) supports the ability of PROMIS measures for guaranteeing fair comparisons across languages and countries in specific clinical populations.

Comparing PROMIS® Physical Function Short-Form10a (PF-10a) to a Clinical Disease Activity Measure in Rheumatoid Arthritis Patients

Elizabeth Wahl, MD [1], Andrew Gross, MD [2], Vladimir Chernitskiy [2], Laura Trupin, MPH [2], Jinoos Yazdany, MD MPH [2], Patricia Katz, PhD [2]

Presenting Author: **Elizabeth Wahl, MD**; elizabeth.wahl@ucsf.edu

Institutional Affiliations: [1] San Francisco Veterans Affairs Medical Center, San Francisco, CA [2] University of California San Francisco, San Francisco, CA

Objective: Physical functioning is considered a key outcome in rheumatoid arthritis (RA), and measurement of physical function is now mandated by some health insurance payers. Our objective was to evaluate the relationship between changes in physical function as measured by PF-10a, and changes in a composite clinical measure of disease activity, among patients with RA.
Methods: Data were collected during routine clinical care at a university- based rheumatology clinic. All patients with RA were administered the PF- 10a by clinical staff at each encounter, and had Clinical Disease Activity Index (CDAI) scores recorded by a clinician. The CDAI is a widely used measure of disease activity in RA, and includes physician assessments of tender and swollen joints, overall disease activity, and a patient assessment of disease activity. CDAI scores can be categorized as indicating 'Remission', 'Low', 'Medium', or 'High' disease activity using established cutpoints; lower scores reflect less active disease. Clinical data were abstracted from the electronic health record from February 2013 – March 2014 for all patients with ≥2 encounters for RA, who had PF-10a and CDAI scores recorded for ≥2 encounters at least 1 month apart. We calculated the change in PF-10a score, the change in CDAI score, and change in CDAI category between each encounter for each individual. We then classified change in disease activity between visits as improved, stable, or worse in two ways—by looking at change in CDAI score (based on a Minimally Important Difference (MID) of 12 points), and by looking at change in CDAI category. Using ANOVA, we compared the mean change in PF-10a across change groups. We then used multivariate linear regression to determine the relationship between change in PF-10a and change in CDAI, adjusting for baseline CDAI score.
Results: Data from 205 patients were included in the analysis. Mean age was 59, 83% were women, and 53% were Caucasian. A total of 735 encounters had PF-10a and CDAI scores recorded (mean visit number 2.4), and changes in CDAI and PF-10a between these encounters was calculated. Of 12% of encounters where CDAI improved (MID≥12), the mean±SD change in PF-10a was 4.5±6.5; among those without a significant change in CDAI (84%), the mean±SD change in PF-10a was 0.39±6.0, and among those with worsening CDAI (5%), the mean±SD change in PF-10a was -3.9 ±8.0. These differences were significant by ANOVA ($p<0.001$), and were consistent whether change in CDAI score was classified by MID or by disease activity category. Multivariate analysis controlling for baseline CDAI showed that increasing change in CDAI score was associated with a decreasing mean PF-10a score (0.29 points lower PF-10a score per unit increase in CDAI; 95% CI -0.38 to -0.21, $p < 0.001$). For each 12 point increase in CDAI, there is a mean decrease in PF-10a of -3.5 points; 95% CI -4.5 to -2.5, $p<0.001$). Adjusting for baseline CDAI did not affect the model.
Conclusions: PF-10a is sensitive to changes in RA disease activity as measured by CDAI, suggesting it could play an important role in the clinical setting.

"Basic PROMIS Panel" a proposal for a unified visual interface to summarize multiple PROMIS domains

James H. Willig, MD, MSPH [1], Jeffrey Curtis, MD [1], Alfredo Guzman, MEng, [1], Bernadette Johnson, BA [1], Ben Nowell PhD, MSW [2], Monika Safford, MD [1], Cooper Filby, MS [1], David M. Condon, PhD [3], David Cella, PhD [3]

Presenting Author: **James H. Willig**, MD, MSPH; jwillig@uab.edu

Institutional Affiliations:[1] University of Alabama at Birmingham 2] Global Healthy Living Foundation 3] Northwestern University

Objective: The PCORnet Patient Reported Outcomes (PRO) common measures working group recommended a focus on domains, rather than specific conditions or diseases, as they are applicable to multiple conditions. To move these domains into standard clinical practice across various conditions, standardized ways to visually summarize data are needed. In the current abstract, we propose a way to summarize and visualize data across multiple domains.
Methods: In order to create a visual way to summarize and visualize data across multiple domains, we sought visual cues from the clinical domain with which to anchor our proposed diagram. Our starting point was to use wireframe on which clinicians typically document the "basic metabolic panel" (BMP) a summary of seven electrolyte (sodium, potassium, chloride, bicarbonate), renal function (urea, creatinine) and glucose measures that is inculcated early in medical school and becomes an inextricable part of the visual vocabulary of physicians by their third year of training. We selected the PROMIS instruments for the domains of anxiety, depression, sleep, fatigue, pain interference, social, and mobility as our measure analogous to the BMP creating a "basic PROMIS panel" (BPP) that would be applicable across multiple disease states. We rotated the wireframe 90 degrees to the right, giving it the look of a stick figure standing. Two variations for presenting BPP data will be presented to the conference audience and feedback will be requested from attendees This feedback, in conjunction with feedback from members of the ARPOWER PPRN will be used to determine the preferred visual representation to add to our at home data collection app.
Results: Each section of the wireframe partitioning the BPP visual displays sections of an area that is represented by a semitransparent tile that will change color from green (above average), to yellow (average) to red (below average) depending on the scores. This way, at a glance, through a color cue, the user can rapidly process overall performance across the BPP domains. In addition, we propose that while the overall figure represents the latest BPP score, selecting a specific tile will bring up a graph of the five previous results for review of longitudinal measures. Screenshots of the two visual display options are presented below, the audience will receive a questionnaire to collect their feedback.
Conclusions: The "Basic PROMIS Panel" is a proposed way to visually summarize an aggregate of seven PROMIS instruments that would provide valuable point of care data for multiple chronic disease states. By modifying visual cues already ingrained in medical training (basic metabolic panel display), we present two visual options to summarize BPP data. We will integrate PHO audience feedback and subsequently patient feedback to optimize our visual display – in effect "researching" the optimal visual display in the PHO conference. Standardized visual approaches to rapidly summarize, facilitate display and expedite interpretation of PROMIS instrument results will facilitate their adoption in the clinical arena.

PRO implementation in clinical practice: Failure as a steppingstone to success

James H. Willig, MD, MSPH [1], Heidi Crane, MD, MPH [2] Michael Mugavero MD, MHSc [1], Paul Crane MD, MPH [2], Michael S. Saag, MD [1], James L. Raper PhD, CRNP, JD [1]

Presenting Author: **James H. Willig**, MD, MSPH; jwillig@uab.edu

Institutional Affiliations: [1] University of Alabama at Birmingham [2] University of Washington

Objective: The implementation of patient reported outcomes (PROs) in routine care settings can benefit patients and providers by providing data to augment point of care decision-making; clinical researchers by providing information on key behaviors, symptoms, and outcomes; and psychometricians by augmenting data capture for refinement of existing, and development of, new instruments. We share the lessons learned from two implementations.

Methods: The UAB 1917 Clinic serves >3,000 individuals with HIV/AIDS. In 2004, we deployed a kiosk in the waiting room that used software designed to capture multiple PRO instruments during clinic visits. The rollout halted clinic workflow; the first patient took >90 minutes to complete their PRO assessment. After 3 hours, implementation was abandoned. In 2008, we again implemented PROs, in partnership with CNICS (CFAR Network for Integrated Clinical Systems), across multiple touchscreens in our waiting area (n=3) and one per clinic room (n=8). We targeted PRO completion within 12 minutes. PRO utilization has continued uninterrupted in our high volume clinical setting since that time. Lessons derived from these experiences for successful point-of-care PRO deployment are shared.

Results: Balance research and clinical imperatives: The desire for PRO capture for the expansion of available research data drove our short-lived 2004 implementation attempt. We asked individual researchers to recommend instruments in their interest areas. The sum of instruments led to a classic game theory "tragedy of the commons" scenario where the total time required by all instruments was overwhelming, leading to patient fatigue, delay in clinic operations and ultimately derailed implementation. Instrument selection must take into account both the clinical and research viewpoints. Brevity matters: One must weigh the overall "time cost" of a panel not individual instruments. A "panel" of clinically relevant compact instruments that prioritizes brevity has a higher chance of successful implementation than one with greater psychometric precision but concomitantly greater length. Brief instruments or those that utilize computer adaptive testing can alleviate "time cost." Comprehensive stakeholder engagement: Involvement of all stakeholders in planning will increase buy-in which in turn provides benefits including; access to in-depth workflow insights that will facilitate implementation, better understanding of patient perspectives which will inform acceptable length and frequency of PRO administration, enhancement of PRO utilization in point-of-care clinical decision making. We recommend prioritizing instruments that impact point-of-care medical decision making, and some research (we suggest a 3:1 "clinical/research" instrument ratio) instruments. Metrics throughout: We recommend electronic monitoring of PRO completion in clinic in order to provide real-time assistance as needed. Implementation should be continually assessed via metrics such as completion rates and the rates of physician utilization of data during encounters to inform changes needed to achieve proposed outcomes.

Conclusions: Our failed 2004 PRO implementation provided key lessons that guided subsequent successful implementation in 2008 of ongoing data

capture across multiple instruments in a high volume clinic setting. Best practices for point of care implementation and at home (between visit) PRO capture need to be shared to fully realize the potential of these tools to aid patients.

Satisfaction with social roles and adults aging with a disability

Rozanne Wilson, PhD, Dagmar Amtmann, PhD, Alyssa Bamer, MPH, Rana Salem, MA

Presenting Author: **Rozanne Wilson, PhD**; rozanne9@uw.edu

Institutional Affiliations: University of Washington

Objective: To compare self-reported satisfaction with participation in social roles (e.g., work, home, and family responsibilities) in four chronic disability groups to an age group-matched Patient Reported Outcomes Measurement Information System (PROMIS) U.S. general population normative sample.
Methods: A total of 1822 participants completed the PROMIS Satisfaction with Participation in Social Roles Short Form 7a version 1.0 including individuals with multiple sclerosis (MS: n = 572), spinal cord injury (SCI: n = 483), post- polio syndrome (PPS: n =432), and muscular dystrophy (MD: n = 335). Sample means were compared to PROMIS age adjusted means to examine differences in satisfaction with participation in social roles between individuals aging with disabilities (i.e., MS, SCI, PPS, MD) and the PROMIS normative sample. Differences were examined for the overall sample, by disability subgroups, and by age group cohorts.
Results: Overall, individuals living with disabilities reported lower levels of satisfaction with participation in social roles ($M = 45.01$, $SD = 8.31$) as compared to the PROMIS normative sample ($M = 50$, $SD = 10$) and this difference was statistically significant ($t(1821) = -25.65$; $p < .001$). All clinical samples reported lower levels of satisfaction with participation in social roles as compared to the PROMIS normative sample, with mean differences ranging from -4.45 (SCI) to -6.15 (PPS) points. In general, satisfaction decreased with age across all disability groups, although the patterns were slightly different for different diagnostic groups.
Conclusions: Results suggest that individuals living with chronic disabilities experience clinically significantly (i.e., about a half SD) lower levels of satisfaction with participation in social roles as they age compared to the age matched general population sample. Two main patterns of satisfaction with social role participation were observed. Older adults living with SCI and MS showed a declining pattern of satisfaction with age, while the MD and PPS samples reported lower levels of satisfaction that were similar across age cohorts. More research is needed to understand the reasons that contribute to decreased satisfaction with participation in social roles in older adults aging with disabilities, and whether or how these reasons differ between people with disabilities and the general aging population. Some caution is advised in interpreting the results for the youngest and oldest groups because the sample sizes for these groups were small.

Measurement of symptoms: The PROMIS dyspnea item banks

Susan Yount, PhD, David Victorson, PhD, Seung Choi, PhD, David Cella, PhD

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

Institutional Affiliations: Feinberg School of Medicine, Northwestern University, Chicago IL

Objective: The Patient Reported Outcomes Measurement Information System (PROMIS) represents a collection of psychometrically sound and efficient measures of patient-reported outcomes (PROs) for people with a wide range of chronic diseases and demographic characteristics. Among other domains of health status, the PROMIS library includes approaches to a wide range of symptoms, from general symptoms, such as pain, fatigue, depression and anxiety, to more focused symptoms, such as gastrointestinal, pulmonary and cognitive symptoms. We will describe how the PROMIS Dyspnea item banks were developed and provide initial evidence of their validity in people with chronic obstructive pulmonary disease (COPD).

Methods: Using methods consistent with both the Food and Drug Administration (FDA) Guidance for PROs and the National Institutes of Health Roadmap PROMIS initiative, we developed two item- response theory (IRT)-based item banks to measure dyspnea severity (DS) and functional limitations (FL) caused by dyspnea for use in clinical research of people with COPD. The process of developing the patient-centered conceptual model and framework involved qualitative input from patients and pulmonary medicine clinicians as well as health outcomes and measurement experts. To maximize geographic diversity and efficiency of data collection, items were tested in a calibration sample of 608 individuals with self- reported COPD from an internet panel company; participants completed new bank items and legacy measures (36-item Short Form Health Survey [SF-36], Chronic Respiratory Questionnaire-Self Administered Standard, Medical Research Council Dyspnea Scale, Hospital Anxiety and Depression Scale) at baseline (n=608) and 7-10 days post-baseline for test-retest reliability (n=236). After calibrated items were selected for the final two 33-item DS and FL banks and two 10-item short forms, they were tested in a different clinical sample of COPD patients recruited from outpatient pulmonary medicine and internal medicine clinics, who completed the assessment at baseline (n=102) and 7-10 days later to evaluate test-retest reliability (stability; n=82).

Results: In the calibration sample, exploratory factor analysis suggested a dominant first factor, and confirmatory factor analysis confirmed a unidimensional model, with measurement equivalence across sex. Based on calibration results, two 10-item short forms were selected, which produced comparable scores (r=0.98) and a computerized adaptive-testing (CAT) simulation indicated efficient measurement with fewer items. Both the calibration and clinical samples produced consistent evidence of reliability and validity. Across both samples, internal consistency (alpha 0.92 – >0.97) and test-retest reliability (r>0.89) were high. There was also strong evidence of convergent (r=0.71-0.88) and divergent (r=0.35-0.65) validity and known groups validity (MRC level; p<.05 to p<.0001). In both testings, CAT required an average of 4-5 items. DS and FL scores are highly correlated (r=0.95), so researchers can reduce respondent burden by administering only one of the measures.

Conclusions: The PROMIS Dyspnea banks and short forms provide options for brief, psychometrically sound measures of patient-reported dyspnea and/or FL in COPD.

Author Index