ABSTRACT BOOK

2017 HealthMeasures User Conference

September 27-28, 2017
Prentice Women's Hospital Conference Center
250 E. Superior Street, Chicago, Illinois

Contents

ii

Abstracts featured in the Second Annual HealthMeasures User Conference

Capturing the Effect of COMT Inhibition on Sequelae of Brain Injuries with NIHTB and TBI-QoL

Emily Adams Berich, Alexander Maclay, PT, PhD, Margo Lauterbach, MD, MPH, Robert Schloesser

Sheppard Pratt-Lieber Research Institute

Objective: The Sheppard Pratt Lieber Research Institute (SPL) is conducting translational research studies in psychiatry informed by neuroscience and genetics. SPLs projects are conducted at the point-of-care and are facilitated by multidisciplinary teams of scientists and clinicians. Here, we are presenting data from an ongoing project examining whether pharmacological inhibition of Catechol-O-Methyltransferase (COMT) can improve cognition, behavioral function and quality of life (QOL) in subjects with a history of brain injuries (BI). This study takes place at a large tertiary referral outpatient clinic for patients with neuropsychiatric conditions. This intervention trial makes use of the NIH Toolbox (NIHTB) and the Traumatic Brain Injury Quality of Life Measurement System (TBI-QoL). Data capture, management and reporting processes are designed to integrate the research process into the clinical environment. Method: This is an open label, proof-of-concept study to evaluate the effects of pharmacological COMT inhibition in patients with a history of BI. Subjects had a remote (>1 year) history of mild to severe BI. Outcome measures include cognitive assessments as well as patient and proxy reported outcome measures. Cognitive tests include the NIHTB cognitive measures as well as a battery of other legacy cognitive tests including Nelsons modified Wisconsin Card Sorting Task, Brief Visuospatial Memory Test Revised, Calibrated Ideational Fluency Assessment and others. Patient and proxy reported outcomes include the TBI-QoL as well as the Frontal Systems Behavior Scale (FrsBe), Geriatric Depression Scale and others. Subjects receive the study medication, Tolcapone, 200mg TID for 14 days and are tested prior to and on the last day of drug administration. In addition to psychometric and cognitive testing, biosamples are collected for analysis of DNA and plasma levels of the study medication. The study is registered with clinicaltrials.gov (NCT02652598). Results: We were able to successfully utilize NIHTB and TBI-QoL in a point-of-care, open label intervention trial studying a population of subjects with BI. Over 20 patients have completed the study so far, including 8 patients with mild and 12 patients with severe BI. Compared to other instruments used in this study, NIHTB and TBI-QoL offered improved efficacy in data collection due to the CAT nature of the instruments. We also developed solutions to integrate the measures with data capture (REDCap) and reporting systems (Drupal, R Shiny) to allow improved engagement of multidisciplinary teams, including clinicians at the point-of-care. Results from this study show significantly improved cognition and quality of live ratings while patients were taking the study medication. Cluster analysis algorithms allowed for stratification of participant groups from NIHTB and TBI-QoL data, differing in cognitive profiles and responses to study medication. Conclusions: NIHTB and TBI-QoL are useful in evaluation of cognition and QOL at the point-of-care as well as in our intervention trial. Utilizing these instruments, we were able to show significant effects of open-label Tolcapone treatment on cognition and QOL in a population of patients with BI. Data from this study has been critical to inform cohort selection and study design for a follow-up placebo controlled trial currently under development.

The Assessment of Quality of Life of Children using Child and Parent-proxy PROMIS Measures.

Joel Alcantara, DC[1], Jeanne Ohm, DC[1], Junjoe Alcantara, DC[2]

The International Chiropractic Pediatric Association[1], Alcantara Chiropractic - Philippines[2]

Objectives: To gain insights on the utilization of chiropractic services by children (8-17 years), investigate their quality-of-life (QoL) prior to and following a trial of care and determine the level of agreement between child and parent-proxy measure.

Methods: In addition to obtaining socio-demographic information (i.e., age, gender and level of education) and clinical correlates of the history and physical examination (i.e., motivation for care), children and their parents independently completed an online survey using the PROMIS-25 and PROMIS proxy measures, respectively. Data was entered into an online data processing center created specifically for the purpose of this study and exported to an Excel spreadsheet for analysis and analyzed using descriptive statistics (i.e., frequencies and percentages, means and standard deviations). The PROMIS data was analyzed using the appropriate PROMIS manual converting raw data to a T score metric (mean=50; SD=10). The greater the T score, the greater the measured QoL domain. Agreement of child and parent scoring was assessed using Intraclass Correlation Coefficient (ICC) via 2-way ANOVA.

Results: Two hundred and eighty nine parent/child pairings completed the questionnaires. The mean age of the children (164 females; 125 males) and parents (247 females; 42 males) were 12.52 years (SD=2.77) and 41.26 years (SD=6.88), respectively. The parents were highly educated with 87% attaining some college or higher education. The majority of children (N=203) were presenting for wellness care and to relieve a symptom(s) followed by wellness care alone (N=47) and symptom care alone (N=39). Median number of visits at baseline and comparative was 1 and 5, respectively. The pre-treatment versus comparative child mean T scores are: mobility (50.82 vs. 52.58), anxiety (46.73 vs. 44.21), depression (45.18 vs. 43.60), fatigue (45.59 vs. 43.92), peer-relationship (52.15 vs. 52.88) and pain interference (47.47 vs. 44.8). The pre-treatment versus comparative parent-proxy mean T scores are: mobility (49.91 vs. 51.83), anxiety (48.25 vs. 45.17), depression (55.63 vs. 55.04), fatigue (43.27 vs.41.29), peer-relationship (51.20 vs.52.06) and pain interference (46.85 vs.44.50). The ICC for the pre-treatment vs. comparative measures are: mobility (0.4629 vs. 0.2581), anxiety (0.4872 vs. 0.4836), depression (0.1208 vs. 0.0849), fatigue (0.4288 vs. 0.3633), peer-relationship (0.3245 vs. 0.3974) and pain interference (0.4779 vs. -0.0607). Overall, the QoL of children improved following a trial of chiropractic care based on self-reported and parent PROMIS measures. Parents had a tendency to over-estimate the decrease in their childs anxiety and depression in terms of pre-treatment and comparative mean T score measures. At baseline, there was fair agreement in mean T scores in mobility, anxiety, fatigue and pain interference and poor agreement with depression and peer-relationships. Following a trial of chiropractic care, the agreement changed for mobility and pain interference from fair to poor while the poor agreement for depression became worse.

Conclusion: There is fair agreement between parent/guardian and childrens rating of their QoL except for depression. Given that parents' perception of their child's state of health influences healthcare utilization, the observed differences and changes in levels of agreement following a trial of care requires continuation in this line of research.

Generalizability and Validation of PROMIS Scores to Predict Surgical Success in Foot and Ankle Patients

Judith Baumhauer MD MPH[1], Michael Anderson MD[1], Charles Saltzman MD[2], Florian Nickisch MD[2], Alexej Barg MD[2], Timothy Beals[2], Jeffrey Houck[1]

University of Rochester Medical Center[1], University of Utah[2]

Introduction: Patient-reported outcomes are advancing clinical care by improving patient satisfaction and engagement. A recent publication reported preoperative PROMIS scores to be highly predictive in selecting patients who would and would not benefit from foot and ankle (F/A) surgery. Although this publication used the data from 5 fellowship trained foot and ankle surgeons at one institution, the generalizability to other patient populations and geographic areas is unknown. The objective of this study is to assess the pre-operative PROMIS physical function (PF) and pain interference (PI) t-scores as a predictor of post-operative success from a separate geographic area.

Methods: Prospective consecutive patient visits to a multi-surgeon tertiary F/A clinic were obtained between 1/2014-11/2016 resulting in 18,565 unique visits and 1,408 new patients. Patients undergoing elective operative intervention for F/A were identified by ICD-9/10; CPT code. PROMIS PF and PI were

assessed at initial and follow-up visits (minimum 6 months, mean 7.8 months). Two-way ANOVA was used to determine differences in PROMIS PF and PI from pre to post surgery with age and gender as co-variates. The distributive method of estimating a minimal clinical important difference (MCID) was used. Receiver operator curve (ROC) analysis was used to determine cut offs for achieving and failing to achieve MCID. To determine the validity of previously published cut offs, 1) they were compared to cut offs for this data set and 2) the percentage of patients achieving and failing to achieve MCID based on previous cut offs were evaluated using a chi-square analysis.

Results: There were significant improvements in PROMIS PF scores (mean=6.0;sd=11.6;p<0.01) and PI scores (mean=-7.0;sd=8.4;p<0.01). The AUC for PROMIS PF (0.77) was significant (p <0.01) and the cut offs for achieving MCID (current data = <23.8 versus previous study= <29.7) and failing to achieve MCID (current data=>41.1 versus previous study=>42) were comparable (Figure 1). Of the patients identified as unlikely to achieve MCID, a significant proportion (88.9%) failed to achieve an MCID ((Chi square=4.7;p=0.03). Of the patients identified as likely to achieve MCID, a significant proportion (84.2%) achieved MCID ((Chi square=17.8;p<0.01). This validates the prior preoperative PROMIS PF thresholds for patients undergoing F/A surgery who will and will not demonstrate MCID improvement in PROMIS PF. The AUC for PROMIS PI was not significant.

Conclusions: PROMIS PF cut offs from published data were successful in classifying patients who would improve in PF with surgery from a different geographic area and academic institution with a broad unique array of surgical procedures, diagnoses, and a diverse patient population. This study provides validation evidence to support using the PROMIS PF as a potential tool for surgical selection to help identify patients who would benefit from surgery as well as those who would not. This can allow for appropriate utilization of healthcare dollars and manpower resources to benefit our patients.

How Much Will I Improve After My Surgery and Will I Be Normal? The Critical Importance of Collecting and Discussing Patient Reported Outcomes Measures With Adult Spinal Deformity Patients

Shay Bess, MD[1], Breton Line, BS[1], Christopher Ames, MD[2], Douglas Burton, MD[3], Renaud Lafage, MS[4], Virginie Lafage, PhD[4], Robert Hart, MD[5], Michael Kelly MD[6], Han Jo Kim, MD[4], Eric Klineberg, MD[7], Richard Hostin, MD[8], Gregory Mundis, MD[9], Munish Gupta, MD[6], Khaled Kebaish, MD[10], Frank Schwab, MD[4], Christopher Shaffrey, MD[11], Justin Smith, MD, PhD[11]

Presbyterian/St. Luke's Medical Center & Rocky Mountain Hospital for Children[1], University of California San Francisco[2], University of Kansas[3], Hospital for Special Surgery[4], Swedish[5], Washington University[6], University of California Davis[7], Baylor University[8], Scripps[9], Johns Hopkins University[10], University of Virginia[11]

Introduction: The surgical decision-making process hinges upon the potential for improvement in pain and function. Surgeons are often asked to quantify the amount of anticipated improvement and the potential to return to normal. Purpose: To quantify the amount of anticipated improvement in pain and function and the ability to return to generational normative patient reported outcomes measure scores (PROMS) following adult spinal deformity (ASD) surgery according to spine deformity type, and with 2 year follow up.

Methods: Surgically treated ASD patients (4 levels fused) were identified from a multi-center prospective, observational ASD database. Patients were organized into spine deformity type according to SRS-Schwab scheme. Demographic, radiographic, and operative data evaluated, and change in PROM (SF-36, SRS-22r, ODI, NRS-back pain, NRS-leg pain) scores from baseline to last follow up within each deformity type quantified and final SF-36 and SRS-22r scores compared to normative values.

Results: 377/582 patients eligible for study (mean age 57.8 years, mean scoliosis 42.1, mean SVA 65.4 mm, mean 11.8 levels fused) had 2yr follow up (mean 3.2 yrs). Improvement for all patients was 44.4% for back pain, 34.8% for leg pain, 20.7% for function, 50.0% for appearance, 3.8% for general health, and 20.1% for social functioning (p<0.05; Table 1). Improvements differed according to deformity type; back pain improvement was consistent across groups whereas activity was most improved in SAGITTAL and MIXED (Table 1). The majority of deformity groups demonstrated >50% return to population normative values, however largest return to normative values included appearance scores in THORACIC (11.1 to 73.7%) and activity scores in DOUBLE (24.2% to 63.5%). Conversely, SAGITTAL had worst return to normative values for all measured domains.

Conclusion: Surgical counseling and establishment of realistic patient expectations requires accurate discussions regarding anticipated surgical outcomes. Quantification of postoperative PROM improvement

and comparison to normative PROM values may provide a consistent framework for surgical counseling. Evaluation of postoperative improvement and comparison to normative values for different ASD types demonstrated ASD patients will demonstrate substantial and quantifiable gains in pain, function, social and psychological measures at 2 year, however, most ASD patients will not return to age based normative values. Additionally, different deformity types had different improvements. We anticipate these data will assist the patient counseling process and encourage surgeons to collect PROMS and to accurately measure spinal alignment to aid these discussions.

Incorporating PROs into the EMR: The Northwestern Medicine Orthopaedic Surgery Joint Replacement Pilot

Surabhi Bhatt, BS, Begum S. Kutay, MS, Nicole M. McGuire, BA, Kathryn Thomas, MSN, RN, SCRN, Cynthia Barnard, PhD, MBA, MSJS, Michael D. Stover, MD, David W. Manning, MD

Northwestern Medicine

Background: Patient reported outcome (PRO) measurement facilitates the documentation of healthcare outcomes, quantifies patient perception of outcome, and can improve provider-patient communication regarding care. Historically, PRO survey results have never been fully integrated into the electronic medical record (EMR) at Northwestern Medicine.

Objective: Implement PROs as standard of care within the confines of normative clinical workflow, embed PROs scores within EMR, and display scores real time to enhance the patient-provider interaction. Additionally we seek to analyze patterns of patient responses to treatment over time and identify areas for quality improvement and research.

Methods: NMPRO (Northwestern Medicine Patient Reported Outcomes) is a tool implemented within the EMR to collect PROMIS PROs through patient remote use or at point of care in clinic. Implementation required custom programming and scoring integrated into Epic using Assessment Center API. The Orthopaedic Surgery pilot population included total hip and knee replacements. PROs were collected through the use of three order sets (pre-surgery, post-surgery, and ad hoc/on-demand) built to include assessments of Physical Function, Pain Interference, Pain Intensity, and Ability to Participate in Social Roles and Activities.

Workflows were developed to place the pre-surgical and post-surgical orders at appropriate points in care. Pre-surgical assessments were completed during the preoperative visit in clinic. Post-surgery order set was placed in Epic the week prior to surgery and triggered an automatic cascade of five survey windows: 3 weeks (2 weeks), 5 weeks (2 weeks), 10 weeks (4 weeks), 23 weeks (4 weeks), 1 year (2 months), and 2 years (2 months). These could be completed at home through the EMR portal or in clinic during routine visits.

The ad hoc/on-demand order set supplemented the pre and post survey order sets by providing a freestanding survey option should additional time points be needed during the course of care.

Once completed, all survey results and trended graphs were available for immediate review in the Epic electronic medical record.

Results: Over the 4 month period since NMPROs launch, pre- and post-op surveys ordered for eligible patients reached 82.74%. Of the patients with ordered surveys, the completion rate was 59.11%. 51.49% of these patients completed the survey on their Epic MyChart account at home, while 48.51% completed the survey in clinic. Average in-clinic survey completion time was 4.31 minutes, which has allowed survey administration to fit into existing workflow with minimal disruption.

Conclusions: PROs can be administered, scored, and made immediately available within the EMR to patients and their providers without disruption of efficient clinical workflows. Measuring quality of care from the patients perspective is likely an indicator of disease impact and trended data over time ought to serve a valuable role in patient care as the program continues to grow. As a result of the high staff and clinician engagement and success of NMPRO implementation in the Orthopaedics Joint Program, Northwestern Medicine plans to incrementally roll out NMPRO throughout our health system.

Using the PROMIS Pediatric Peer Relationships Scale for Children with Cleft Lip and/or Palate

Tara Braun, BS, Amy Xue, MD, Christy Hernandez, RN, Laura Monson, MD

Baylor College of Medicine

Objective: Camp Keep Smiling is a weekend-long recreational retreat that is part of the standard treatment plan for children with cleft lip and/or palate who are managed in the Plastic Surgery Department at Texas Childrens Hospital. After our first camp session in 2014, the anecdotal feedback we received from campers, parents, and counselors indicated that there was a dramatic improvement in campers attitude, mood, and confidence upon their return. The current analysis aimed to detect any changes in peer relationships after camp using an objective outcomes measure.

Methods: At the most recent camp session in March of 2017, we administered the PROMIS Pediatric Peer Relationships questionnaire to all campers whose parents consented to their participation in the study. The questionnaire was administered on the first morning during camp registration, and again two days later on the last day of camp.

Results: A total of 36 out of 44 campers participated in the study and completed pre- and post-camp surveys. The mean (SD) pre-camp raw score was 23.6 (7.0) and the mean (SD) post-camp raw score was 25.0 (6.8). On average, the score increased by 1.4 points after camp, but this increase was not statistically significant. We are still gathering long-term follow-up data to analyze scores a month after attending camp.

Conclusions: The average score on the PROMIS Peer Relationships questionnaire increased after attending camp, though these results were not statistically significant. In the future, we plan to administer the pre-camp survey a month prior to the first day of camp, since the majority of participants were return campers who reported exchanging contact information, staying in touch over social media between sessions, and re-establishing communication in the weeks prior to the most recent camp session.

A Pilot Test of ASCQ-Me Measures in Adults with Sickle Cell Disease

Dominique Bulgin, BSN, RN, Paula Tanabe, PhD, Susan Silva, PhD

Duke University

Objective: Sickle cell disease (SCD) is a genetic blood disorder that has serious health complications including stroke, renal failure, vaso-occlusive crisis, leg ulcers, acute and chronic pain. These complications can negatively influence quality of life (QoL). This pilot study tests measures for stigma [Sickle Cell Disease Health Related Stigma Scale (SCD-HRSS)] and QoL [Adult Sickle Cell Quality of Life Measurement System (ASCQ-Me)] developed specifically for individuals with SCD. The SCD-HRSS measures perceived stigma by health care providers (nurses and doctors), the general public, and family. The aim of this pilot study is to test the ASCQ-Me and SCD-HRSS instruments to determine: 1) recruitment rate, 2) percent completion of the instrument battery, and 3) patient perceptions of QoL outcomes and perceived stigma in a convenience sample of adults with SCD.

Methods: A survey design was used. Participants completed self-administered assessments, including demographic and clinical characteristics, ASCQ-Me, and the SCD-HRSS. SCD-HRSS scores range 10-50 for subscales and 50-200 total scale, with higher scores indicate greater perceived stigma. Total raw scores for emotional distress, pain impact, sleep impact, social functioning impact, and stiffness impact range 5-25. Total raw scores for pain severity range 0-22 and 0-11 for pain frequency. Higher scores indicate worse health status on the subscales. Recruitment began at a SCD conference and was completed at the Duke Adult SCD Clinic. Inclusion criteria were >18 years of age, ability to read English, and self-reported diagnosis of one of the following SCD genotypes: HbSS , HbSC, Hb SB0or SB+. A total of 16/20 individuals have been recruited to date. Descriptive statistics were used to analyze data.

Results: After 18 hours of data collection over three weeks, 16 individuals participated in the study, with 1 refusal. Participants took an average of 15 minutes to complete the survey; 62.5% reported being female; and all except one reported being Black or African American. Ages ranged from 22 to 67. The median, 25th and 75th percentiles on the ASCQ-Me domains were: emotional distress (10; 6, 12), pain impact (12; 8, 12), sleep impact (13.5; 12, 15.5), social functioning impact (12.5; 8, 15), stiffness impact (9.5; 6, 11.5), pain severity (17.5; 13.5, 19), and pain frequency (10; 7.5, 11). The median, 25th and 75th percentiles on the SCD-HRSS domains were: nurses (30.5; 24.5, 33.5), doctors (34; 28, 37.5), general public (32.5; 27, 36.5), family (17; 13.5, 24.5); and total (115; 100.5, 124.5).

Conclusions: Individuals were able to complete the survey without assistance and generally did not skip items on the surveys. Participants reported the moderately high scores for pain, sleep, and social functioning impact, as well as overall pain severity. Stigma scores were moderate for all domains except family where scores were low. Future research is required to explore the relationships between perceived stigma and QoL in SCD.

Patient Centered Outcomes for Clinical Use in Pediatric Manual Therapy

Beth Carleo, DC[1], Katie Pohlman, DC, MS, PhD(c)[2], Kris Anderson, DC, MS, Carol Parnell Prevost, DC, DICCP[1], Brian Gleberzon, DC, MSc[3],

Palmer College of Chiropractic[1], Parker University[2], Canadian Memorial Chiropractic College[3],

Objective: As one of the most commonly utilized complementary and alternative medicine modalities in North America, manual therapy represents an area for further research efforts to elucidate and substantiate preliminary findings. With these research efforts using valid, reliable, and relevant outcome measures, which is the central effort in the development of the PROMIS measures and NIH Toolbox. The objective for this project is to evaluate patient-centered outcomes currently utilized for the pediatric population undergoing manual therapy and obtain expert and consumer consensus on measures that should be used in the clinical setting, including the use of PROMIS measures and NIH toolbox. The main goal of this project will be to determine clinically appropriate, condition specific outcomes to make research driven suggestions that can be easily used in pediatric manual therapy healthcare settings.

Methods: This multi-step evaluation starts with a systematic review of manual therapy research studies in the pediatric population. Inclusion criterion are: patients under the age of 18, published between 2001 and January 2016, and manual therapy (chiropractic, physiotherapy, or osteopathic). Observational and experimental designs were included, while case reports were not. Previous comparable reviews were cross referenced. The second part of this project is to evaluate the outcome measures, including patient-centered and objective outcomes, used in each study. For the identified patient-centered outcomes, all psychometric properties referenced will be appraised using the COSMIN (Consensus-based Standards for the selection of health Measurement Instruments) checklist. These outcome measures will also be compared to what is currently being used in research with other modalities and what is available using PROMIS and the NIH Toolbox. A Delphi panel of manual therapist scientist and practitioners will be convened to assess the clinical approach going forward for pediatric manual therapy. Final clinical recommendations will be published and posted on relevant websites for integration into relevant research studies and clinical practice.

Results: The first part of this project is complete. Forty-four articles met the inclusion criteria of the systematic review. Of which, the following categories were studied: Structural / Musculoskeletal (n=16); Respiratory (n=8); Gastrointestinal (n=11); Special Needs (n=9).The second part of this study has just started with the outcome measures utilized within the structural / musculoskeletal articles.

Conclusions: Currently, there is a lack of reviewed, easily accessible, evidence-based patient-centered outcomes for manual therapy practitioners to use in regard to specific presenting complaints in the pediatric population. The end-product of this project will be easily-accessible, valid, reliable, and relevant patient-centered outcome measures that have been reviewed and critiqued by manual therapy scientists and practitioners. By having easy to access instruments, like PROMIS and the NIH Toolbox, both researchers and practitioners can more efficiently assess the effectiveness of manual therapy for the pediatric population.

Understanding patient-reported outcome measures in individuals with neurodegenerative diseases that include cognitive decline

Noelle E. Carlozzi[1], Lai, Jin-Shei[2], Hahn, Elizabeth[2], Stout, J[3], Paulsen, JS[4], Kratz, AL[1], Gershon, Richard[2], Cella, Dave[2], Schilling, SG[1]

University of Michigan[1], Northwestern University[2], Monash University[3], University of Iowa[4]

Objective: Patient-reported outcomes measures can be difficult to use with individuals with neurodegenerative diseases given associated cognitive declines. We examined the reliability of several new PROs from the HDQLIFE measurement system (an affiliate of the Neuro-QoL system) in individuals with Huntington disease (HD) in order to establish clinical cutoffs for cognition for when to use these PROs.

Methods: 501 individuals with prodromal, early-stage or late-stage HD completed three different PROs from HDQLIFE: Speech Difficulties, Swallowing Difficulties, and Chorea. In addition, participants completed clinician-administered assessments from the Unified Huntington Disease Rating Scales (UHDRS) including two measures of cognition (Stroop and Symbol Digit Modalities; total sum score was calculated by summing the total scores across these assessments) and items from the Total Motor Scale (which assesses chorea, dysarthria and dysphagia items from the UHDRS). Item bias using differential item functioning (DIF) both across HD stage (prodromal, early-, late-) and relative to cognitive performance. We also examined correlations between self-report and associated clinician ratings. A simple linear regression model, a heterogeneous variance model, and a heterogeneous variance model that examined total cognition were used to examine the psychometric reliability of the PROs.

Results: The majority of PRO items did not exhibit DIF (either for staging or cognition). In addition, there were modest correlations between PROs and associated clinician report (ranged from -0.40 to -0.60); correlations by staging group were lower (especially for those with later-stage HD). Modeling analyses indicated that measurement reliability decreases (i.e. split-half reliability <0.80) as cognition declines and staging increases. Critical cutoffs for reliability (i.e. split-half reliability <0.80) when total cognition scores are <136 for Chorea, <109 for Speech Difficulties, and <179 for Swallowing Difficulties.

Conclusions: As HD progresses and cognition declines, high error variance and low reliability can negatively affect the psychometric properties of a PRO. In cases where cognitive scores do not meet critical cutoffs, PRO measures should only be considered in conjunction with other assessments. Furthermore, findings suggest that cognitive complexity may vary across PRO measures given that recommended clinical cutoffs differ for the different PROs.

The Reliability and Validity of Neuro-QoL in Individuals with Huntington Disease

Noelle Carlozzi[1], Jin-Shei Lai[2], S Goodnight[1], Elizabeth Hahn[2], RE Ready[3], P Dayalu[1], J Stout[4], JS Paulsen[5], S Frank[6], M McCormack[7], NR Downing[8], C Ross[9], Richard Gershon[2], JA Miner[1], AL Kratz[1], N Boileau[1], J Russell[1], Dave Cella[2]

University of Michigan[1], Northwestern University[2], UMASS Amherst[3], Monash University[4], University of Iowa[5], Beth Isreal Deaconess[6], Rowan University[7], Texas A&M University[8], Johns Hopkins University[9]

Objective: There is a need for meaningful and sensitive patient reported outcome (PRO) measures that examine health-related quality of life (HRQOL) in individuals with Huntington disease (HD), an autosomal dominant neurodegenerative disease. While the Neuro-QoL offers a potential solution to this shortcoming, data regarding the reliability and validity of Neuro-QoL computer adaptive tests (CAT) and short form (SFs) in individuals with HD are needed before they can confidently be used in research and clinical care.

Participants/Methods: 512 individuals with HD completed 12 Neuro-QoL measures (Lower Extremity Function, Upper Extremity Function, Satisfaction with Social Roles and Activities, Ability to Participate in Social Roles and Activities, Applied Cognition-Executive Function, Applied Cognition-General Concerns, Emotional and Behavioral Dyscontrol, Positive Affect and Well-Being, Stigma, Anxiety, Depression, and PROMIS Anger), 2 other generic HRQOL measures (EQ-5D and WHODAS 2.0) and several clinician-rated measures of motor, emotion, cognitive, and functional status. We examined internal consistency (Cronbachs alpha), convergent and discriminant validity (Pearson correlations with other self-report and clinician-rated measures), and known groups validity (univariate analyses to examine the ability of the measures to differentiate between prodromal, early, and late HD).

Results: Cronbachs alphas were all 0.88. The Neuro-QOL measures generally had moderate to high correlations with the generic self-report measures (all rs .64) and less robust correlations with clinician-rated composite measures, supporting discriminant validity. Univariate analyses indicated that all Neuro-QOL measures (except Anger, Anxiety and Depression) were able to differentiate between prodromal, early, and late HD (p <.03). While data supported the reliability and validity of both CATs and SFs, CATs sometimes offered slight advantages.

Conclusions: Results support the reliability and validity of these Neuro-QoL measures in evaluating HRQOL and suggest that they are appropriate for use across the full HD spectrum. Work is underway to collect data that examines change over time to determine their clinical utility in repeat trials.

The Reliability and Validity of PROMIS in Caregivers of Individuals with Civilian- or Military-related Traumatic Brain Injury

Noelle E. Carlozzi[1], Ianni, P[1], Kallen, Michael[2], Hahn, Elizabeth[2], Brickell, T[3], Lange, RT[3], French, L[3], Kratz, AL[1], Gershon, Richard[2], Cella, Dave[2], Miner, JA[1], Tulsky, DS[4]

University of Michigan[1], Northwestern University[2], Walter Reed National Military Medical Center[3], University of Delaware[4]

Objective: There is a need for meaningful and sensitive patient reported outcome (PRO) measures that examine health-related quality of life (HRQOL) in caregivers of individuals with military- or civilian-related traumatic brain injury (TBI). While the Patient-Reported Outcomes Measurement Information System (PROMIS) offers a potential solution to this shortcoming, data is needed to establish the reliability and validity of these measures in these caregivers before they can be used with confidence in research and clinical care.

Participants/Methods: 558 caregivers of individuals with TBI (344 civilian and 214 military) completed 10 PROMIS measures (Anger, Anxiety, Depression, Sleep Disturbance, Fatigue, Emotional Support, Informational Support, Social Isolation, Ability to Participate in Social Roles and Activities, and Satisfaction with Social Roles and Activities), a generic measure of HRQOL (the RAND-12), 2 measures of caregiver burden (the Zarit Burden Interview and the Caregiver Appraisal Scale), and a measure to estimate the overall functioning of the individual with TBI for whom they cared (the Mayo-Portland Adaptability Inventory-4). We examined internal consistency (Cronbachs alpha), 2-week test-retest reliability, convergent and discriminant validity (Pearson correlations with other HRQOL and caregiver burden measures), and known groups validity (univariate analyses to examine the ability to differentiate between caregivers of individuals with low versus medium versus high functioning TBI), of the PROMIS computer adaptive tests (CATs) and short forms (SFs).

Results: Cronbachs alphas were all .88 and 2-week test retest reliability was .63 (with the majority of the measures being .71). The PROMIS measures generally had moderate to high correlations with other measures of similar constructs (all rs .66) and less robust correlations with different domains of HRQOL. Univariate analyses indicated that all PROMIS measures were able to differentiate between caregivers of individuals that were low, medium and high functioning (p <.05). While data supported the reliability and validity of both the CATs and SFs, CATs sometimes offered slight advantages.

Conclusions: Results provide support for the reliability and validity of these PROMIS measures in evaluating HRQOL among caregivers with military and civilian-related TBI. Future work is needed to examines the sensitivity of these measures to change over time in this caregivers to establish their clinical utility in repeat trials.

Making a Swedish PROMIS: A Needs Assessment

J.E. Chaplin PhD AFBPsS[1], E Nilsson[2]

Institute of Clinical Sciences, Sahlgrenska Academy at University of Gothenburg, Gothenburg[1], Quality Register Center QRC Stockholm, Stockholm, Sweden[2]

Objective: To identify the processes needed for the successful integration of an item banking system into an established framework of quality control and PRO measurement.

Methods: Sweden has over 100 quality registries with the aim of creating and improving standards of healthcare. Each NQR focuses on a specific diagnosis collecting data on patients and the procedures. Recently, greater emphasis has been placed on the incorporation of patient-reported outcome data. A thorough investigation conducted in 2015 concluded that Sweden should implement the PROMIS item banking system at a national level in order to take advantage of future computer-assisted technology options and utilises the possibilities of large-scale data analysis. A project team, financed by the NQR organisation, has carried out a needs assessment and initiated a series of joint projects which are aimed at leading to translation and validation of all PROMIS item banks in Sweden.

Results: The majority of adult PROMIS item banks are in the process of being translated and an agreement has been reached to translate all the paediatric item banks. A series of translation quality control groups has been established to evaluate the translations which were carried out according to the PROMIS protocol. These groups consist of clinicians with different specialisms, linguists and patient representatives. Cooperation with a number of NQRs has been established to start the validation procedure which has been integrated into ongoing or newly established research projects. The implementation of a nationwide change in the way PROM data is collected is a major challenge due to differences in data collection procedures. Some NQRs have established PRO instruments, some have electronic data collections. Those looking to implement PRO measurement are the most receptive to the implementation of the new procedure. The possibility of using a Computer Adaptive Test (CAT) is very appealing to all NQRs. Integration at a national level will require a secure patient portal through which identified patients can gain access to the short forms and the CAT item banks. This future e-health system will deliver the CAT software and scoring manuals that can be incorporated into the existing NQR infrastructure for data collection and data storage. Different administrative options are currently being assessed.

Conclusion: A sustainable implementation strategy will have to address the resistance to replacing existing PRO data collection procedures. It is considered vital for smooth integration to involve patient groups and future users as co-designers and advisors. Co-design will help to ensure that item banking will be rapidly integrated into the Swedish NQRs. Further research on the Swedish translations will increase confidence in their use. This includes calibration of the Swedish translations, examination of their content validity in specific populations, testretest reliability, construct validity and responsiveness in relevant patient populations and in the general population. Swedish-specific reference scores and short-forms will facilitate the interpretation and usability of PROMIS and further improve confidence.

A Validation Protocol for NeuroQoL Item-banks and Short Forms in Sweden

J.E. Chaplin PhD AFBPsS[1], K. Lycke[2], L. Egertz[3], E. Helmersson[4], L. Stawiarz[2], J. Hillert[2]

University of Gothenburg[1], Karolinska Institute[2], Stockholm County Council[3], Swedish MS Registry[4]

Objective: The objective of this research paper is to describe the protocol for an assessment of the reliability and validation of the recently translated Neuro-QOL item banks and short forms (SFs) with a Swedish population of adults with MS.

Methods: A prioritised subset of Neuro-QoL item banks and short forms have been translated into Swedish. These will complement patient reported outcome measures currently being used in the Swedish Neuro-registries/Multiple Sclerosis registry (NEUROreg/MSreg). NEUROreg aims to ensure that neurological care is equitable and of high quality and that treatment guidelines are followed. Patients contributing to the quality registry at their normal clinic appointment will be asked to complete the standard PRO instruments and the Neuro-QoL item-banks and SFs. Inclusion criteria an are age older than 18 years, confirmed diagnosis of MS and completion of informed consent. Current PROs include: MS Impact Scale (MSIS-29), Fatigue for motor and cognitive functioning (FSMC); Treatment Satisfaction Questionnaire (TSQ) and the VAS scale from the RAND-36. Neuro-QoL Fatigue and Cognitive functioning item-banks,

SF for upper and lower functioning and Ability to participate in social roles and activities will be added to the registry. The PROMIS 10-Item Global Health Scale (GHS) will provide an overall assessment of health. A Global Rating of Change (GRC) will be constructed to assess patients subjective evaluation of the amount of change experienced in Physical, Emotional, Cognitive, Social/Family, Well-being and overall QOL over a six-month period.

Results: The degree to which the translations are free from measurement error will be measured via both internal and external reliability. Internal reliability will be calculated using Cronbachs alpha with coefficient scores of 0.70 or above. External reliability will be assessed via a test-retest procedure comparing baseline results with a follow-up between 7 and 21 days after the first. Intraclass correlation coefficients (ICC) and corresponding 95% confidence intervals (CIs), with coefficient scores of 0.70 or higher will be considered acceptable. To measure the degree to which a translations are able to measure the relevant concepts, construct validity will be assessed by comparison at baseline of the MSIS-29 and FSMC; known groups analysis using analysis of variance and ROC curves will compare baseline mean Neuro-QoL scores between MS severity categories defined by clinical assessment and self-reported health. Bland and Altman plots will also be used to identify proportional bias. In order to detect changes over time the responsiveness to change will be evaluated using general linear models of each patients GRC score between baseline and a 6-month follow-up.

Conclusions: The Swedish Neuro-registries/Multiple Sclerosis registry provides an ideal platform for the assessment of validity and reliability of Neuro-QoL, however, the additional burden on the patient due to the completion of extra questions requires careful integration into the data collection routine which will do done electronically. Participation and coverage are good within the registry and therefore it is not anticipated that there will be substantial non-participation or problems with missing data.

The Distribution of PROMIS Domains among Patients Undergoing Spine Surgery

Amy M. Cizik, PhD, MPH[1], Katie Odem-Davis, PhD[1], Jesse Fan, MD, MPH[1], Jeffrey G. Jarvik, MD, MPH[1], Richard A. Deyo, MD, MPH[2], David R. Flum, MD, MPH[1], Danielle Lavallee, PhD, PharmD[1]

University of Washington[1], Oregon Health and Science University[2]

Objective: In 2014, the NIH Task Force on Research Standards for Chronic Low Back Pain (cLBP) recommended a broad set of standardized questions to be used when studying patients with back pain. These included Patient-Reported Outcomes Measurement Information System (PROMIS) 4-item profile domains. Pain Interference, Pain Intensity, Physical Function, Depression, and Sleep Disturbance were included. It is unknown if these domains are independent or if some have greater weight and the number of domains evaluated can be reduced. It may also be that other domains not previously recommended should also be included, such as anxiety. Our goal was to examine the recommended measures and determine the extent to which the domains correlated, in a cohort of patients prior to undergoing lumbar fusion.

Methods: A prospective observational study of patients undergoing lumbar fusion from seven participating clinics, which included assessment of PROMIS domains as described previously. Patients completed a baseline questionnaire prior to surgery and PROMIS items were scored using the Health Measures scoring service. The T-score metric was used to standardize the raw score for each 4-item response to the reference population, where 50 is the mean and 10 is the standard deviation (sd). One-half a standard deviation (5) was considered to be a meaningful difference in scores. Correlation statistics and a principal-component analysis (PCA) were conducted to describe relationships and identify dimensions across the PROMIS domains. The number of components was selected by the minimum number required to explain 80% of the total variance among the patients in the cohort.

Results: We report on 87 patients cared for by 11 spine surgeons who completed the baseline survey. Sixty-three percent were female, 43% were retired and 15% were disabled due to back pain or other reasons. The majority of patients (77%) experienced radiculopathy within 2 weeks prior to the baseline. Low back-pain duration was >1 year for 38% of patients and 41% experienced back pain almost every day for the past 6 months. Anxiety and depression were highly correlated (r=0.73, p<0.001). Pain Interference correlated moderately with other scores, except social support (r=0.03, p>0.05). Based on the T-score metric, baseline Pain Interference (mean=67, sd=6), Anxiety (mean=57, sd=9) and Social Support (mean=55, sd=9) were elevated at baseline relative to the PROMIS reference population. The first component had three domains with correlations 0.50: anxiety, depression, and social support, explaining 43% of the variance of all the domains. The second component also had three domains with correlations 0.50: pain, sleep, and social

support, explaining 29% variance. The third component had two domains with correlations 0.50: global health (2-item) and pain, explaining 27% variance.

Conclusions: Anxiety and depression are highly correlated and with social support explain almost half of the spread of PROMIS data across this cohort. The PCA analysis of baseline data suggests that we may be able to reduce the number of PROMIS domains assessed. This finding also helps to design biopsychosocial programs targeting these specific combinations of responses to PROMIS profile domains for patients undergoing lumbar fusion.

Factors associated with sleep disturbance as individuals diagnosed with colorectal cancer transition to early survivorship

Theresa Coles, MS[1,2], Xianming Tan, PhD[1], Antonia V. Bennett, PhD[1], Hanna K. Sanoff, MD[1], Ethan Basch, MD[1], Roxanne E. Jensen, PhD[3], Bryce B. Reeve, PhD[1]

University of North Carolina at Chapel Hill[1], RTI-Health Solutions[2], Georgetown University[3]

Objective: Sleep disturbance is associated with decreased cognitive functioning, fatigue, and increased healthcare utilization. Sleep disturbance is also a risk factor for cardiovascular and infectious diseases, as well as depression. Individuals diagnosed with cancer manage unique disease and treatment impacts throughout the cancer continuum depending on the cancer site and stage of the disease. Sleep disturbance is understudied in individuals diagnosed with colorectal cancer (CRC). The purpose of this study was to identify patient, disease, and treatment characteristics associated with sleep disturbance (and change in sleep disturbance) in a sample of individuals diagnosed with localized CRC. We also investigated variation in patient, disease, and treatment characteristics within levels of sleep disturbance severity and magnitude of change in sleep disturbance in early survivorship.

Methods: Data were obtained from the MY-Health study, a community-based observational study of adults diagnosed with cancer collected through four Surveillance, Epidemiology and End Results (SEER) cancer registries. Patient-Reported Outcomes Measurement Information System (PROMIS) Sleep Disturbance (6-item custom form), Anxiety, Depression, Fatigue, and Pain Interference measures were administered to individuals with cancer. Participants self-reported demographic information, comorbidities, treatment type and dates of treatment. Data were collected at two timepoints after diagnosis: 6-21 months (mean =10 months) and 13-26 months (mean = 17 months). Regression mixture and multiple regression models were used to evaluate the relationship between sleep disturbance diagnosis and other patient, disease and treatment characteristics cross-sectionally at an average of 10 months after diagnosis and the magnitude of change in sleep disturbance over a six month period, from approximately 10 and 17 months post-diagnosis.

Results: Patient, disease and, treatment characteristics associated with sleep disturbance did not differ by severity of sleep disturbance or change in the severity of sleep disturbance between the 10-month and 17-month assessment points. Pain (B = 0.09), anxiety (B = 0.22), fatigue (B = 0.29), retirement (B = -2.49, reference category: full-time or part-time work, student), and being diagnosed with 2 or more self-reported comorbid conditions (B = 1.53) had statistically-significant relationships with sleep disturbance (P <0.05) in the cross-sectional model. The change model revealed that worsening anxiety (B = 0.14) and fatigue (B = 0.20), and lower PROMIS Sleep Disturbance scores (better sleep quality) 10 months after diagnosis (B = -0.21) were associated with worsening sleep disturbance from 10 to 17 months after CRC diagnosis (P <0.05).

Conclusions: Although statistically-significant, correlation coefficients associated with pain, anxiety and fatigue were less than 1, meaning that over approximately 6-months, a 40-unit improvement in PROMIS anxiety scores would be associated with only a 5.4-unit improvement in PROMIS Sleep Disturbance scores. Other statistically-significant patient characteristics also yielded small coefficients (<|2.5|) in the models. Together, these results suggest that no patient, disease and treatment factors included in the models were obvious indicators of sleep disturbance. Given the negative consequences of sleep disturbance on health outcomes, sleep disturbance screening may be warranted.

COMPASS-CP: Precision Health at Its Best

Pamela Duncan, PhD, Cheryl Bushnell, MD, Scott Rushing, Mysha Sissine, MSPH, Rica Abbott, MPH

Wake Forest Baptist Health

The COMPASS Study is a Patient Centered Research Outcomes Institute pragmatic, cluster-randomized trial of 41 hospitals in North Carolina designed to determine the effectiveness of comprehensive models of coordinated post-acute stroke care. The COMPASS care model is consistent with Center for Medicare and Medicaid Services (CMS) value-based care models for management of patients with complex needs, such as stroke. CMS care models require care coordination, development of a care plan for complex patients, beneficiary engagement for self-management, and shared decision making. Individualized care plans must be available to all patients and health providers to manage care coordination and optimize the patients ability to manage their own health and independence. As part of the COMPASS trial, a Wake Forest Baptist Medical Center interdisciplinary team is solving the barriers to chronic care management with an interoperable electronic application, COMPASS-CP for Health. COMPASS-CP is easily scalable and will provide opportunities for health systems to integrate care across the delivery system. eCOMPASS is a patient-centered electronic application that captures the social and functional determinants of an individuals health and their goals of care at the point of clinical care.

Objective: 1) Demonstrate how patient-reported outcomes using a web-based application can support the generation of individualized electronic care plans in real-time (COMPASS-CP); 2) demonstrate the utility of COMPASS-CP in identifying social and functional determinants of health; and 3) describe how COMPASS-CP measures quality and meets CMS and MACRA performance metrics for value-based care.

Methods: The COMPASS study includes enrollment at discharge, a follow-up call within 2 days after hospital discharge, and clinic visit within 14 days. An advanced practice provider and nurse perform standardized assessments that captures patient-reported outcomes measures in real-time at the point of clinical care (goals for care, physical, social and medical function, health literacy, medication management). COMPASS-CP uses proprietary algorithms to immediately generate provider reports and individualized care plans with recommendations for self-management, prevention, referrals to rehab/community resources, and support.

Results: Preliminary analysis of the first 219 patients (3,000 anticipated) enrolled in the intervention arm of the COMPASS trial revealed 36.5% had systolic BP >140 mmHg, 32.4% required additional rebab services, 39.2% needed social (e.g transportation, meals on wheels, personal assistants) and/or caregiver support, and 79.4% needed financial support or assistance medication management. There were significant differences in patients who were flagged for needed services by age, gender, insurance type, stroke severity, and primary stroke center status, but not race/ethnicity, or geography.

Conclusion: Patient-reported comprehensive assessments at the point of care identified a substantial number of unmet patient needs post-acute stroke discharge. We have demonstrated the feasibility and value of eCOMPASS within the context of post-acute clinical practice. In the future, we will expand the COMPASS-CP application to acquire PROMIS10 outcomes at 90 days.

Assessment of Health-Related Quality of Life in Adults Hospitalized with Sickle Cell Disease Vaso-Occlusive Crisis

Kimberly S. Esham, MD [1], Angie Mae Rodday, PhD, MS[1], Ruth Ann Weidner, MBA, MRP[1], Rachel J. Buchsbaum, MD[1], Hedy P Smith, MD, PhD[2], Susan K Parsons, MD, MRP[1]

Tufts Medical Center[1], Boston Medical Center[2]

Objective: To describe the vaso-occlusive (VOC) pain experience and its effect on health-related quality of life in hospitalized adults with sickle cell disease (SCD), using two patient reported outcomes (PRO) measures, the Patient-Reported Outcomes Measurement Information system (PROMIS) Global Health and the Adult Sickle Cell Quality of Life Measurement System (ASCQ-Me).

Methods: Adults with SCD hospitalized at a US-based academic medical center with VOC were eligible. The local Institutional Review Board approved the study protocol and all participants provided written informed consent. Following study enrollment at hospital admission (T1), participants completed the ASCQ-Me Pain Episode assessment, which yields a Frequency score (range 0-11) and Severity score (range 0-22), where higher scores reflect worse disease over the prior year. Seven days after hospital discharge (T2), participants also completed PROMIS Global and ASCQ-Me questionnaires. These measures yield scores in domains of global physical and mental health, and SCD-related domains of emotional, social

12

functioning, pain, stiffness, and sleep impact, respectfully. Standardized domain scores have a mean of 50 (SD=10) where higher scores indicate better functioning. Changes of 3 points (0.3 SD) are considered clinically meaningful. We calculated means and standard deviations for domain scores at T1 and T2.

Results: We report preliminary results from our larger inception cohort. To date, 28 of 29 eligible patients consented and completed T1 assessments (97%). 18 also completed T2 assessments. Of the 10 participants who did not complete T2 assessments, 50% were readmitted within 1 week and 20% within one month; the remainder (n=3) were non-respondents at T2.

The mean age was 30.6 years (SD 9.9) with 57% females and 71% self-identified as black, non-Hispanic. The remainder was Hispanic. The most common genotype was HbSS (57%). T1 mean ASCQ-Me Pain Episode Frequency score was 9 (SD 2.2, range 4-11) and Severity score was 18 (SD 3.1, range 8-22), reflecting recurrent and prolonged painful VOC episodes over the prior year.

The mean domain scores for all T1 PROMIS and ASCQ-Me assessments were well below population norms: PROMIS physical health score 37(SD=6.7), mental health 44.5(SD=8); ASCQ-Me emotional impact 44.9(SD=7.2), social functioning 46.2(SD=6.8), pain impact 42.2(SD=6.4), stiffness 48.9(SD=7.5), sleep 45.1(SD=9).

The mean domain scores for T2 PROMIS and ASCQ-Me assessments also remained well below population norms: PROMIS physical health score 40.3(SD=9.2), mental health 44.2(SD=7.5); ASCQ-Me emotional impact 46.4(SD=7.9), social functioning 43.9(SD=7), pain impact 44.9(SD=8.4), stiffness 46.2(SD=8.5), sleep 44.9(SD=5.8).

Conclusions: Administering both general and disease-specific PRO measures at the point-of-care is feasible and informative. Among this cohort of hospitalized patients with VOC, baseline ASCQ-Me Pain Episode scores reflect recurrent and prolonged painful VOC episodes over the prior year. All ASCQ-Me domain and PROMIS Global scores were well below population norms, indicating substantial suffering from the multidimensional impact of the VOC pain experience both upon admission and following discharge. Despite the intensity of the patients condition, we demonstrate successful enrollment (97%) and data collection on admission. 64% data completeness at T2 indicates some degree of non-ignorable missing data, which will be addressed in final data analysis.

PROMIS and the American Joint Replacement Registry: National Implementation of PROMIS in a Clinical Data Registry

Caryn D. Etkin, PhD, MPH[1], September R. Cahue, MPH[1], Bryan D. Springer, MD[2], Terence J. Gioe, MD[1,3], David G. Lewallen, MD[1,4]

American Joint Replacement Registry[1], OrthoCarolina Hip and Knee Center[2], UCSF Medical Center[3], Mayo Clinic[4]

Objective: The American Joint Replacement Registry (AJRR) was established in 2010 as an independent 501(c)3 organization. AJRRs goal is to track 90% of all total joint arthroplasty (TJA) procedures conducted across the U.S. Data collection reflects different levels, specifically Level I (procedural), Level II (patient comorbidities and operative/post-operative complications), and Level III data (Patient-Reported Outcomes [PROs]). This presentation will describe AJRRs history relating to PRO collection, how a national legislative initiative focused the PRO direction for AJRR and TJA in general, and if the implementation of the initiative increased the submission of PRO data to AJRR.

Methods: AJRR contracts with hospitals, private practice groups and ambulatory surgery centers to receive data on all TJAs performed either at the institution, or by surgeons in the practice. At the onset of data collection efforts in 2012, AJRR intended to provide a web-based service for participants to capture PROs. AJRR sought to provide a warehouse for PRO data of any sort, related to quality of life or specifically related to TJA. Seven PRO instruments were available, and participants were neither advised nor required to conduct follow-up assessments at a specified time point. However, in 2015, the Centers for Medicare & Medicaid Services (CMS) announced the Comprehensive Care for Joint Replacement (CJR) bundled payment initiative. CJR was mandated for hospitals in 67 Metropolitan Statistical Areas (MSAs). To receive additional points toward the institutions total quality score, hospitals were asked to submit PRO data to CMS for baseline (pre-operative) and 12 months post-operative. Institutions were required to submit one joint-specific measure and one quality of life measure (either VR-12 or PROMIS-10 Global) for each patient.

Results: Early uptake of AJRRs Level III system was sporadic, and included mainly institutions that were already capturing PRO data for their own internal research purposes. Indeed, the utilization of the Level III systems, or PROs in general, was not a priority related to TJA. With the inclusion of PROMIS in

CJR, a surge in PROMIS utilization was expected, as AJRR has enrolled over 900 participating institutions including 135 CJR hospitals. At this time, AJRR has received n=1,369 pre-operative assessments and n=738 12-month post-operative assessments from only 66 participating institutions.

Conclusions: One year after implementation of CJR, uptake of PROs by hospitals remains slow. Developing a process for data capture can be a daunting process, involving numerous staff and new patient protocols. As bundled payment programs proliferate and the status of PROs in the realm of health care quality increases, AJRR expects increases in the use of PROMIS.

The Passport to Wellness: Mobilizing the Learner Workforce to Describe Health Determinants of Defined Populations

Leah FitzGerald, PhD, FNP[1], Brian Prestwich, MD[1,2]

Mount Saint Mary's University[1], University of Southern California[2]

Abstract: To achieve the Triple Aim of better patient experience/care processes, better outcomes, and lower health costs primary care clinics will require informatics platforms capable of screening and risk stratifying defined populations. Drivers of wellness may then be identified and prioritized. The same is true for any organization committed to improving the health of their own defined population. The authors designed a web-based service-learning platform to support nursing students evaluating healthcare needs of individuals in community settings. A two-page patient report survey collects metrics across 4 domains of wellness (oral health, behavioral health, physical fitness/ nutrition/rehabilitation, and medical) and social determinants of health. The software generates a Wellness Plan which can be shared with participants in hard copy and/or emailed electronic format. The platform generates dashboards to minimize cognitive workload for learners and faculty as they review the survey results, care recommendations, and learner competency.

Methods: Study participants included 30 community members (age 31 52 years) who were caretakers of children attending a free dental clinic in South Los Angeles and 524 college students (age 19 31 years) from a private non-profit urban University. Participants completed the two-page survey and received software generated Wellness Plans. Participants gave written informed consent prior to the investigation. The research protocol was approved by Mount Saint Marys University, Institutional Review Board for use of Human subjects according to the Helsinki declaration. Data were analyzed to describe the wellness profile of both defined communities and are expressed as means SE for continuous variables and as percentages for categorical/binary variables. Results: Dental clinic participants: 33% F (45.2 years) 87% non-Caucasian. University students participants: 87% F (21.7 years) 42.4% non-Caucasian. Prevalent factors identified in dental clinic participants: no health insurance 33.3%; lack PCP 26.7%; 43.3% overweight/13% obese; 43.3% report eating junk food 2/wk; 23% lack confidence to maintain lifestyle during stress; 50% without DDS or dental cleaning within past year; 23% witness or personal violence; 36.7% not enough money to meet basic needs; 50% English language difficulty. Prevalent factors identified in the University participants: No health insurance 3.1%; lack PCP 13.7%; 25.4% overweight/9.7% obese; 65.1% report eating junk food 2/wk; 26.3% lack confidence to maintain lifestyle during stress; 18% without DDS; 25% without dental cleaning within past year; 30.5% witness or personal violence; 35.7% not enough money to meet basic needs;11.2% English language difficulty.

Conclusions: A web-based service-learning platform may be useful for screening and risk stratifying defined populations across 4 domains of wellness while generating individualized Wellness Plans to engage and activate participants in advocating for their own wellness. This data may identify different care needs within defined communities, facilitating strategic allocation of resources more likely to improve the wellbeing of respective communities. Future research may establish that this approach is scaleable and that both individual and collective community wellbeing are improved when this approach is utilized across communities.

Implementing PROMIS for Routine Screening in Ambulatory Cancer Care

Sofia Garcia, PhD[1], Frank Penedo, PhD[1], Lynne Wagner, PhD[2], Michael Bass, MS[1], Oana Popescu, MS[1], Timothy Pearman, PhD[1], Yanina Guevara, BA[1], David Cella, PhD[1]

Northwestern University[1], Wake Forest University [2]

Objectives: Oncology patients commonly experience subjective symptoms that are best evaluated through patient-report and can otherwise go unrecognized by clinicians. Leading organizations are increasingly promoting routine patient-reported symptom assessment as a critical component of quality cancer care. To help meet Commission on Cancer distress screening standards, we implemented an EHR-integrated ePRO assessment that includes PROMIS CATs to assess symptoms, and checklists to assess psychosocial and nutritional concerns, as part of standard oncology care within medical oncology clinics at the Robert H. Lurie Comprehensive Cancer Center of Northwestern University.

Methods: We used Assessment Center for the administration of PROMIS computer adaptive tests (CATs) and supportive care needs checklists. Through custom programming, we integrated assessment administration, scoring, and clinician notification within the local EHR (Epic). Prior to scheduled medical oncology visits, patients received instructions to complete online assessments at home through MyChart, Epics patient communication portal. PROMIS CATs were administered to assess anxiety, depression, fatigue, pain interference, and physical function. Checklists identified psychosocial concerns, informational and nutritional needs, and risk factors for inadequate nutrition. Assessment results (including PROMIS T scores) were immediately populated in Epic, and designated clinicians were notified of clinically elevated symptoms via EHR messages. We summarized assessment metrics as part of a clinical quality improvement initiative.

Results: In 2016, we sent MyChart messages to 4,536 outpatients (all cancer types) to complete clinical assessments; 1,050 patients (23.15%) completed them. A minority of patients who completed the assessments had PROMIS CAT scoresAnxiety (7.17%), Depression (7.79%), Fatigue (2.58%), Pain Interference (1.80%), and Physical Function (4.91%)above established institutional thresholds for clinician alerts (T>65, 60, 70, 70, & <30, respectively). Greater numbers of patients who completed assessments endorsed nutritional (42%), psychosocial (32%), and health learning (20%) needs that prompted clinician alerts. Queries of Epic encounters indicated that, as prescribed by the initiatives distress screening guidelines, psychosocial and dietician follow-up contacts after assessment alerts occurred within three days; health learning center follow-up for informational needs occurred within four days.

Conclusions: We used PROMIS CATs, along with psychosocial and nutritional needs checklists, to measure common cancer symptoms and concerns in routine oncology outpatient care. EHR integration facilitated the use of symptom reporting as the basis for referral to supportive and follow-up care. This demonstrated that a brief and validated ePRO assessment could successfully be integrated within an EHR and clinical workflows on a large scale. However, less than a quarter of patients who were invited to complete the assessments did so, perhaps because the onus fell on them to initiate it. We are now seeking to extend the initiatives reach by expanding the ePRO assessment platform to allow for administration in clinics, with providers encouraging patients to participate.

Technical Evaluation of Implementing Patient-reported Outcomes Measures into an Electronic Health Record

Heather T. Gold, PhD, Alissa Link, Raj Karia, Rachel Lebwohl, Devin Mann, Joseph Zuckerman, Thomas Errico, James Slover, Aaron Buckland, Michael Cantor

NYU School of Medicine

Objective: There is much interest in learning about how healthcare systems implement electronic platforms for collection of PROMs as they become increasingly important in the care and evaluation of patients, as well as research efforts. The project bridged formal human-computer interaction (HCI) techniques into a large academic medical centers clinical and Information Technology (IT) operations enterprise, with the goal of ensuring that the PROMs assessments are usable, efficient, and acceptable to patients and clinicians. This report details challenges and how we overcame them.

Methods: A group of committed physician champions worked together with a multidisciplinary team from clinical fields, information technology, user experience design, medical informatics, and population health research. Our approach leveraged REDCap, Epic Systems, MyChart, Oracle, and Symfony, and

integrated the API for PROMIS-CAT from Northwestern University. Software developers and study investigators used rapid-cycle design, meeting with the full team bi-weekly to review design components for visual acceptance and usability. The team conducted initial user interviews with Epic super-users in the institution and a workshop with orthopaedic surgery faculty to ascertain their preferences surrounding PROMs data presentation and placement.

We expected the following areas to be challenges in developing this application: Developing a mobile-responsive, user friendly web interface Integrating clinical data capture with a research-focused backend Integrating the PRO assessment into the paperless registration process Displaying PRO-related information for use by both clinicians and patients Working within the technical constraints of our EHR system

Results: The team developed a mobile-responsive web application using the Symfony framework. The platform uses REDCap for configuring disease-specific assessments and for de-identified data storage. An Oracle database maintains much of the back-end logic for configuring and assigning PROM assessments and the identification linkages to medical record numbers (MRN). The system integrates with Epic using Interconnect web services and Epics API. The clinic staff scan a QR code during check-in/registration that passes encrypted MRN and CSN (encounter ID) to the PROMs application. The application then uses a web service to call appointment and demographic details from Epic, assigning the correct PRO assessments according to the patients provider, visit type, age and language. Instruments are scored in real-time and data are saved to REDCap and Epic flowsheets. Providers can view, graph and trend the PROMs data for each patient in an Epic Synopsis Report. The patient can also view his/her scores in MyChart via a Health Trends report. A portal for querying PROMs and clinical data will be created for researchers. Finally, MyChart will be integrated so patients can complete PROMs outside the office, and custom analytics dashboards will be built for providers to enhance score interpretation.

Conclusion: This custom-built PROMs platform leverages software commonly used by academic medical centers and integrates with one of the most widely used EHRs, Epic. The efforts of a multidisciplinary team helped overcome operational and technical challenges to build a scalable platform for the routine collection of PROMs allowing incorporation into clinical practice.

Disease and Treatment Factors Associated with Lower Quality of Life Scores in Adults with MEN-1

Sneha Goswami, BA, Benjamin J. Peipert, BA, Susan Yount, PhD, Cord Sturgeon, MD, MS

Northwestern University

Objective: The physical and psychosocial morbidity of multiple endocrine neoplasia type 1 (MEN-1) is ill-defined. Due to lack of genotype-phenotype relationships, heterogeneity in clinical manifestations, and infrequent surgical cures, there is controversy regarding optimal management. Thus it is important to understand how disease and treatment-related factors relate to patient-reported outcomes (PROs) such as health-related quality of life (HRQOL). We hypothesized that disease and treatment burden negatively impacts HRQOL in adults with MEN-1.

Methods: Adults >18 years with MEN-1 (n=174) were recruited through an MEN-1 support group. An online survey on demographics, disease features, and treatment history was administered, and respondents were grouped and compared based on their answers. The PROMIS-29 instrument was used to evaluate HRQOL in 7 domains. T scores generated from MEN-1 survey responses were compared to United States (US) normative data using a 1-sample t-test. The Mann-Whitney U test was used for subgroup analysis of categorical variables and linear regression was performed on continuous variables. Holms-Bonferroni was used to correct for multiple comparisons. Data are reported as mean T scores standard deviation.

Results: Adults with MEN-1 reported scores reflecting poorer physical, mental, and social well-being than US normative data in all 7 domains ($p<0.001$): anxiety (61.110.2), depression (57.110.5), fatigue (60.711.7), pain interference (55.611.1), physical function (44.79.5), sleep disturbance (57.29.0), and social function (44.910.6). Respondents with recurrent hyperparathyroidism (42%) reported worse scores than those without recurrence in 5 of 7 domains: anxiety (63.28.6 vs 58.110.0, $p<0.01$), depression (58.58.9 vs 54.311.3, $p<0.05$), fatigue (63.210.8 vs 57.911.9, $p<0.05$), pain interference (58.89.8 vs 53.012.0, $p<0.01$) and social functioning (43.39.3 vs 48.211.4, $p<0.01$). Requirement of prescription medication for MEN-1 (73%) was associated with worse PRO scores across all 7 domains ($p<0.05$). Increased frequency of specialty doctor appointments was associated with higher levels of anxiety, depression and pain interference and lower levels of physical and social functioning ($r = 0.19$-0.28, $p<0.05$). Traveling >50 miles for specialty care was associated with greater anxiety ($p<0.01$). Increased frequency of hospitalization was associated with higher

anxiety (r=0.26, p=0.001). No significant association was found between presence of pancreatic tumors or extent of pancreatic surgery and HRQOL. PRO scores were independent of age, sex, and ethnicity.

Conclusion: This study is the first to explore PROs in MEN-1. Adults with MEN-1 report lower HRQOL than US normative data in all 7 PROMIS-29 domains. Factors associated with lower HRQOL include recurrent hyperparathyroidism, increased travel distance and frequency of specialty appointments, hospitalization frequency, and requirement of prescription medications.

Validity of the Healing Encounters and Attitudes Lists (HEAL) in Persons with Ongoing Pain

Carol M Greco, PhD[1], Lan Yu, PhD[1], Nathan Dodds[1], Kelly Johnston, MS[1], Jeremiah Slutsky[1], Ronald Glick, MD[1], Natalia Morone, MD, MS[1], Nicole Kelly, Jelena Janjic, PhD[2], Ajay Wasan, MD[1], Michael Schneider, DC, PhD[1], Christine McFarland[1], Clem Adams, Suzanne Lawrence, MS[1], Paul Pilkonis, PhD[1]

University of Pittsburgh[1], Duquesne University[2]

Objective: Treatment outcomes are influenced not only by the specific actions of a modality, but also by nonspecific factors such as the patients views of the therapeutic context, their views of themselves, and their expectations of treatment. However, such nonspecific or contextual factors are seldom measured in research trials or in clinical contexts. PROMIS methodology was used to develop the Healing Encounters and Attitudes Lists (HEAL), a set of measures of nonspecific factors that may influence healing or symptoms improvement with treatment. The HEAL measures can be administered as computerized adaptive tests (CAT) or brief static scales, and are applicable to a broad range of medical conditions and treatments. The HEAL include: Patient-Provider Connection, perceptions of the Healthcare Environment, Treatment Expectancy, Positive Outlook, Spirituality, and Attitudes toward Complementary/Alternative Medicine (CAM). The HEAL are currently under review for inclusion in PROMIS. In this report, we present information on concurrent and predictive validity of the HEAL in a sample of persons initiating new treatment for chronic pain in Pittsburgh, Minneapolis, and Los Angeles.

Methods: Participants were persons with chronic pain who had recently (within one month) started an integrative/CAM treatment (n=109) or conventional medicine treatment (n=100) for pain. All participants completed informed consent procedures before joining this online assessment study. The HEAL measures and selected PROMIS Health Status Measures were administered at baseline and 6 and 16 weeks later. Participants also completed the clinical global impression of improvement (CGI) at the 6 and 16 week follow-up assessments.

Results: The average age of participants was 47.5 (sd=15), 75% were female, 23% were non-white or multiracial, and 48% had less than a 4-year college degree. The most frequently reported treatments were physical therapy, acupuncture, medication management, chiropractic, and injections. In the overall sample, PROMIS Pain Intensity and Pain Interference improved from baseline to the 16-week assessment [t(195)= 7.7, p<0.001, t(195) = 4.74, p<0.001, respectively]. On the CGI, 67% reported improvement (somewhat better or much better) at 6 and 16 weeks post baseline. HEAL Positive Outlook was inversely associated with PROMIS Depression (-.70, p<0.01) at baseline, suggesting concurrent validity but not complete overlap. In regression models, baseline HEAL Positive Outlook and Spirituality accounted for 17.7% of the variance in PROMIS Pain Interference at 6 weeks. HEAL Patient-Provider Connection and Attitudes toward CAM accounted for 9.4% of the variance in clinical global impression of improvement at 6 weeks. Outcomes at 16 weeks were also associated with baseline HEAL scores. For example, Baseline HEAL Treatment Expectancy and Baseline Pain Intensity accounted for 42% of the variance in 16 week Pain Intensity. Treatment Expectancy contributed 2% additional variance beyond that accounted for by Baseline Pain Intensity.

Conclusions: Several HEAL measures at baseline were associated with pain treatment outcomes in persons initiating conventional or CAM treatments. Use of HEAL in research trials may allow a greater understanding of contributors to outcomes, and use of HEAL in clinical settings may provide useful information for enhancing interpersonal aspects of treatment settings.

Incorporating PROMIS Symptom Measures into Primary Care Practice: A Randomized Trial

David Haggstrom, MD[1,2], Nina Talib, PhD[2], Patrick Monahan, PhD[2], Jacob Kean[3]

VA Center for Health Information & Communication[1], Regenstrief Institute, Indiana University School of Medicine[2], University of Utah[3]

Objective: Symptoms account for more than 400 million clinic visits annually in the U.S. Among the most prevalent, chronic, disabling, and undertreated symptoms is the SPADE pentad sleep, pain, anxiety, depression, and low energy/fatigue. The objective of this trial was to determine the effectiveness of providing PROMIS symptom scores to primary care clinicians.

Methods: Patients were enrolled from 4 academic primary care clinics. Those with 1 SPADE symptom on a 5-item screener and who provided consent then completed a 20-item PROMIS questionnaire (i.e., 4-item scales for each of the SPADE symptoms). A PROMIS T-score of 50 is the population norm and each 10 point increase represents one standard deviation worse. Thus, a T-score of 55 (effect size of 0.5 worse than population norm) was operationally defined as clinically significant. The 300 trial participants were randomized to a feedback group (n=151), in which their physician received a graph of their symptom scores, or a control group (n=149). At 3-month follow-up, the PROMIS questionnaire was re-administered. Also, patients were asked if their symptoms had been discussed and/or treated at the baseline visit and whether they still desired treatment. The baseline visit note was reviewed for documentation of SPADE symptoms as well as symptom-specific diagnostic or treatment actions. The primary trial outcome was 3-month improvement in the composite SPADE T-score (mean of 5 individual symptom scores), and secondary outcomes included individual symptom T-scores, symptom documentation, diagnostic, and treatment actions in the EMR note; patient-recalled discussion and treatment of symptoms; and residual desire for treatment at follow-up.

Results: There were no differences between study arms in the baseline characteristics of participants who were 72% female, 49% African-American, with a mean age of 49 years, and a moderately severe composite PROMIS T-score of 58.3. Most patients (84%) had multiple clinically significant (T-score 55) SPADE symptoms. Both groups demonstrated moderate symptom improvement, with a nonsignificant trend favoring the feedback compared to control group [composite T-score improvement of 3.5 vs. 2.4 (effect size, 0.46 vs. 0.37), P = 0.16]. Symptoms present at baseline persisted at 3-month follow-up 70% of the time. Pain was documented the most often (81%), whereas fatigue was documented the least often (16%). Except for pain, the other 4 symptoms were documented <50% of the time. The proportion of symptoms not discussed at the baseline visit was lowest for pain (12%), intermediate for sleep and fatigue (22% each), and highest for depression (35%) and anxiety (36%). Residual desire for treatment at 3 months ranged from 23% for depression to 40% for pain. Both discussion/treatment at the baseline visit and residual desire for treatment increased with PROMIS symptom severity. However, there were no differences between feedback and control group patients in these outcomes, nor in diagnostic or treatment actions by the physician.

Conclusions: Assessing and providing PROMIS symptom scores to physicians immediately prior to seeing the patient is feasible in the primary care setting. However, changing symptom outcomes may require additional training, visit time, team support, or other incentives.

Computer-based Administration of Health Literacy and Functional Literacy Measures in Physical Rehabilitation Populations

Elizabeth Hahn, MA[1], Noelle Carlozzi[2], Susan Magasi[3], David Tulsky[4], Sofia Garcia[1], Jin-Shei Lai[1], Joy Hammel[2], Alex Wong[5], Ana Miskovic[1], Sara Jerousek[1], Arielle Goldsmith[1], Kristian Nitsch[1], Allen Heinemann[1]

Northwestern University[1], University of Michigan[2], University of Illinois at Chicago[3], University of Delaware[4], Washington University[5]

Objective: Literacy skills are critical for adults to function effectively in their daily lives. Functional literacy focuses on the ability to read, write and speak in English, and to perform quantitative tasks; health literacy focuses on obtaining, processing and understanding health information and services to make appropriate health decisions. Health literacy may be significantly worse than functional literacy because of the unfamiliar context and vocabulary of the health care system. The purposes of this project were: 1) to validate computer-based administration of functional and health literacy measures in physical

rehabilitation populations, 2) to evaluate associations between functional and health literacy, and 3) to evaluate associations between health literacy and self-reported health.

Methods: English-speaking community-dwelling adults with spinal cord injury, stroke or traumatic brain injury completed the NIH Toolbox Oral Reading Recognition Test (word recognition; interviewer-guided performance) and Picture Vocabulary Test (vocabulary knowledge; self-administered), Health Literacy Assessment Using Talking Touchscreen Technology (Health LiTT; comprehension of prose, document and quantitative health information; self-administered), and an assessment of overall health status (poor-excellent; self-administered). Touchscreen computers were used with adaptive accessories such as a rollerball mouse, and study staff were available to provide assistance as needed. Correlational and analysis of variance methods were used.

Results: Approximately 200 individuals were enrolled in each of the three physical rehabilitation groups (total n=604). Participants in the spinal cord injury and traumatic brain injury groups were predominantly male and Non-Hispanic White, with mean ages of 46 and 40, respectively. The stroke group was slightly older (mean age, 56 years), with equal numbers of women and men, and 48% Non-Hispanic Blacks. About one-third in each group had a high school or lower education, and about one-third were currently married/in a committed partner relationship. The majority of participants (92%) were able to use the computer to complete some or all of the questions. Study staff primarily provided assistance due to an individuals physical limitations. Missing data for the literacy measures were less than 4%. The stroke group had the lowest levels of health literacy and functional literacy, and the poorest overall health. Health LiTT was strongly correlated with the two functional literacy measures (Oral Reading Recognition, r=0.62; Picture Vocabulary, r=0.65). Mean Health LiTT scores increased across increasingly higher levels of self-reported health, within each injury group (spinal cord injury, p=0.004; stroke, p=0.004; traumatic brain injury, p=0.122).

Conclusions: A diverse sample of participants with spinal cord injury, stroke or traumatic brain injury were able to complete computer-based administration of functional and health literacy measures. Lower health literacy was associated with poorer health. As evidence accumulates about the impact of low literacy on health outcomes, a self-administered health literacy-specific measure may be useful to estimate the size of the population at risk from low health literacy, to identify vulnerable patients in clinical settings, to enhance the effectiveness of the client-provider relationship in rehabilitation practice, and to use in testing interventions.

Understanding the Need for Assistance When Completing Measures of Patient-Reported Outcomes in Huntington Disease

Elizabeth Hahn, MA[1], Nancy Downing[2], Julie Stout[3], Jane Paulsen[2], Becky Ready[4], Siera Goodnight[5], Jin-Shei Lai[1], Jennifer Miner[5], Noelle Carlozzi [5]

Northwestern University[1], University of Iowa[2], Monash University[3], University of Massachusetts[4], University of Michigan[5]

Objective: Health-related quality of life is critically important for people with Huntington disease (HD). Motor, cognitive and psychiatric abnormalities begin more than a decade before diagnosis (prodromal HD) and progressively worsen throughout the disease course. Unfortunately, the symptoms that are characteristic of HD may also impair the ability of affected individuals to complete patient-reported outcome (PRO) measures. The purpose of this report is to describe the extent and type of assistance required by HD participants to complete PRO measures, and to examine the impact of such assistance on PRO scores.

Methods: The Huntington Disease Quality of Life (HDQLIFE) mixed-methods study developed a PRO measurement system for HD. A sample of individuals across the prodromal and diagnosed disease severity spectrum completed PROs from Neuro-QoL and PROMIS identified as relevant to HD, and newly developed HDQLIFE item pools. Participants were asked to report the amount and type of assistance they received in completing the questionnaires. Responses to the assistance questions were used to create three groups: no assistance, computer assistance only, and assistance with survey questions with or without computer assistance. Chi-square tests, Fishers exact tests and analysis of variance methods were used to compare sociodemographic, clinical and PRO data across HD diagnostic groups and assistance groups, and to compare participants who were excluded from this report to those who were included. Multivariable linear regression models were created for PROs, and least-squares means were estimated to determine the effects of assistance on PROs after adjustment for other covariates (gender, age, race/ethnicity and diagnostic group).

Results: Among 532 enrolled participants, 56 (10.5%) did not complete the assistance questions and were excluded from analyses. The excluded group had larger proportions with late stage disease, racial/ethnic minority status, low education and single marital status, and poorer cognitive, motor and independence functioning compared to those retained (n=476). Sociodemographic, clinical and PRO variables differed between prodromal HD (n=191), early stage HD (n=186) and late stage HD (n=99) individuals retained for analysis. The majority of the late stage HD individuals (78%) received some assistance completing questionnaires compared to 29% of early stage HD and <1% of prodromal HD individuals. Those who received assistance had higher proportions with late stage disease, were older, had less education, had a lower proportion who were married/partnered, and had poorer functional and cognitive skills. Before and after adjustment for gender, age, race/ethnicity and diagnostic group, those who received assistance had poorer scores on some PROs. Participants who received assistance completing the PRO surveys took about twice as long to complete them compared to those who did not receive assistance.

Conclusions: Participants who received assistance with survey completion were older, more advanced in disease, and had greater clinician-rated cognitive and functional impairment than those who did not receive assistance. This emphasizes the potential impact of HD on the ability to complete PROs independently. These findings suggest that when using PROs, clinicians and researchers should be aware of the necessity of assistance, and should plan the best strategies for assessing PROs in people with HD.

Introduction to the PROMIS-Preference (PROPr) Summary Score

Janel Hanmer, MD, PhD[1], Barry Dewitt[2]

University of Pittsburgh[1], Carnegie Mellon University[2]

Objective: A single preference-based summary score provides a simple method to compare health-related quality of life across groups and across time, and is necessary for applications like cost-effectiveness analyses and cost-utility analyses. The goal of this project was to create a generic, societal, preference-based summary score for PROMIS (adult). Generic means that it should capture a core set of domains important to most people, regardless of their specific health conditions. Societal means that the score should represent the aggregated preferences of the US society as a whole. Preference-based means that expressed preferences are used to give relative value to different levels of health, drawing heavily from multi-attribute utility theory.

Methods: First, we used a combination of expert (n=9) and community (n=50) input to select a set of PROMIS domains which were comprehensive and structurally independent. Second, we tested several methods to present a PROMIS domain for valuation in a community sample (n=118); we developed a method in which two representative items are selected from a domain to construct health states for valuation. Third, the health states were valued by an online panel (n=983) using the standard gamble technique. Fourth, single-attribute utility scores were constructed for each domain. Fifth, the single-attribute scores were combined into a multi-attribute score such that Dead = 0 and Full Health = 1.0.

Results: Seven adult PROMIS domains were selected for PROPr. The online panel was representative of the US adult population by age, gender, race, ethnicity, income, and education. Single-attribute functions were estimated as cubic functions. These single-attribute functions can be used as cardinal measures of utility in each domain. The multi-attribute function is a multiplicative function. It is scaled such that Dead = 0 and Full Health = 1.0, allowing PROPr scores to be used for calculating aggregated indices of morbidity and mortality (such as quality-adjusted life years). In the online panel sample, PROPr scores are correlated to the EuroQol-5D and Health Utilities Index. The PROPr score changes by age, sex, and presence of chronic conditions are consistent with prior measures.

Conclusions: PROPr was carefully constructed with input from a variety of measurement experts as well as community members. The final scoring algorithm was estimated using a large representative sample of the US noninstitutionalized population. PROPr is the first score to directly link single-attribute utility functions to health domains as measured by Item Response Theory (IRT). As such, PROPr should gain many of the advantages of an IRT-based descriptive system, including flexible administration of items from item banks. A PROPr score can be estimated if there are score estimates from each of these 7 PROMIS domains: Cognitive Function Abilities Subset v2.0, Emotional Distress Depression v1.0, Fatigue v1.0, Pain Interference v1.1, Physical Function v1.2, Ability to Participate in Social Roles and Activities v2.0, and Sleep Disturbance v1.0.

Characterizing Patients with Patient Acceptable Symptom State (PASS) after Selected Foot and Ankle Orthopedic Procedures using Patient Reported Outcomes

Jeff Houck, PT, PhD[1], Michael Anderson, MD[2], A. Samuel Flemister, MD[2], Irvine Oh, MD[2], Benjamin DiGiovanni, MD[2], John Ketz, MD[2], Judy Baumhauer, MD[2]

George Fox University[1], University of Rochester[2]

Objective: To determine whether selected patient reported outcomes(PROs) characterize which patients are satisfied with their current level of symptoms and activity (Patient acceptable symptom state(PASS)) after specific foot and ankle orthopedic procedures.

Methods: Patients treated with three of the most frequent current procedural terminology(CPT) codes (n=359) were provided for call backs 2 months to 2 years after their surgery. Patients reached on the phone (n=73; age 58.4 (14.2), follow up 12.4 (7.2) months) for CPT codes 28725-Arthrodesis subtalar (n=26), 28750-Fusion of Big Toe Joint (n=22), and 27691-Transfer tendon ant/post tibial through interosseous space (n=25) completed the following: answered the PASS question, rated their global level of functioning, declared whether their surgery was successful or not, and completed PROMIS physical function (PF), pain interference (PI), and depression scales(Dep). T-tests and two-way ANOVAs were used to compare PASS No to PASS Yes individuals depending on the scale. Receiver operator curve analysis was used to derive cut offs useful in characterizing patients that are likely to identify as PASS Yes/No. The association between PASS Yes/No and success after surgery was also assessed (Chi square Test). Logistic regression was used to determine if a set of PRO variables yielded higher accuracy in classifying PASS Yes/No participants and the influence of co-variates (i.e. age and length of follow up).

Results: PASS Yes(n=56) as compared to PASS No (n=16) participants were characterized by significantly higher PROMIS PF 8.9 (95% CI 13.5 to 4.3; p <0.01), lower PROMIS PI 8.3 ((95% CI 4.3 to 13.5; p <0.01), and lower PROMIS Dep 7.7 (95% CI 2.4 to 13.0; p <0.01). The global rating of function was also significantly different (p<0.01) for participants that were PASS Yes (7.2 (1.4)) compared to PASS No (4.2 (2.0)). The ROC analysis yielded significant (p<0.01 for all scales) area under the curves(AUC) for PROMIS and Global Rating Scales (AUC ranged from 0.73 to 0.88). Cut offs meeting the pre-determined accuracy of 95% sensitivity and specificity showed that PROMIS PF of <33.2, PI of >65.5, Dep >60.3 and global rating of function of <4.5 accurately identified PASS NO participants. In contrast, PROMIS PF of >46.4, PI of <49.9, Dep <33.2 and global rating of function of >7.5 with identified PASS Yes classification. PASS Yes/No classification correlated with patients judgement of whether their surgery was successful or not (Chi-square X2=103.4, p<0.1). Logistic regression showed that the combination of PROMIS PI (p=0.04), Global rating of function (p<0.01), and Age (p =0.07) achieved 88.2% accuracy for classifying participants as PASS Yes/No.

Conclusions: Selected PROs were successful in characterizing patients with acceptable symptoms and activity levels after specific foot and ankle orthopedic procedures. The specified cut offs are useful for clinical decision making and informing patient engagement strategies following common foot and ankle orthopedic procedures.

Evaluating Change in Joint Function: The HOOS JR, the KOOS JR, and the PROMIS PF

Man Hung, PhD, Maren W Voss, Jerry Bounsanga, Yushan Gu, Michael B Anderson, Jeremy Gililland, Chris Peters, Charles L Saltzman, Tom Greene, Chris Pelt

University of Utah

Objectives: Patient reported outcome (PRO) measures are useful in improving patient care and have been incorporated into the Centers for Medicaid and Medicare Services comprehensive care for joint replacement model. The HOOS JR and KOOS JR were developed in short-forms that reduce patient burden in assessments. The PROMIS Physical Function (PF) instrument is developed using item response theory, facilitating a computer adapted administration format that can shorten the number of items required while maintaining precision. All three measures may be useful assessment tools in evaluating outcomes for joint conditions while minimizing the burden of test-taking. Responsiveness is the ability of a measure to detect change over time and is an essential component of outcome measures used to understand treatment response. The purpose of this study was to evaluate the responsiveness of three measures commonly used to assess PROs in joint care.

Methods: 983 patients visiting an academic orthopaedic clinic between 2014 and 2017 were evaluated for joint conditions. The HOOS JR, KOOS JR, and PROMIS PF measures were administered electronically prior to clinic visits. Change scores were calculated between a baseline visit and three and six-month follow-up visits. Three-month follow-up was calculated in two formats, as 90 days plus or minus 10 days and as 90 days and beyond. Similarly, the six-month follow-up was calculated as both 180 days plus or minus 10 days and as 180 days and beyond. Change scores were anchored by a patients self-report of noticeable change. Patients reporting no change or worsening symptoms were excluded from the responsiveness analysis.

Results: The average age of the sample was 61.03 years (SD = 12.33, Range = 18-90 years) and predominantly White (n=875, 89.0%). All changes from baseline scores were significant at the 3-month, >3-month, and >6-month follow-up ($p<0.05$). All three measures showed large effect sizes, ranging from 0.80-1.20 at each time-point. The standardized response mean was large for each measure and at each time-point (Range = 1.06-1.53).

Conclusions: Our study demonstrates the responsiveness of the HOOS JR, KOOS JR and PROMIS PF in a population of joint/adult reconstruction patients. The PROMIS PF was consistently the most responsive instrument for patients with joint conditions. Though we are not able to evaluate the measures based on specific procedure codes due to a limited sample size, our analysis indicates these three measures are responsive to treatment changes in this population. The HOOS JR, KOOS JR and PROMIS PF are useful clinical assessment tools that may be selected for use based on the needs of joint/adult reconstruction patient population.

Correspondence of Performance Based Physical Tests and PROMIS Physical Function Scores in Community Dwelling Elderly

Ryan P. Jacobson, DPT, PCS[1], Michael Bass, MS[2], Chris DeSilva, BS[3], Judith F. Baumhauer, MD, MPH[3], Jeff Houck, PT, PhD[1]

George Fox University[1], Northwestern University[2], University of Rochester[3]

Objective: To establish the correspondence between the Patient-Reported Outcome Measurement Information System (PROMIS) Physical Function (PF) scale and benchmarks for common physical performance tests used clinically.

Methods: Forty-six individuals over 65 years old (77+/-4.6 years) participated in a single data collection. The session included completion of the PROMIS PF v1.0 computer adaptive test using Assessment Center (www.assessmentcenter.net) and completion of the modified Physical Performance Test (mPPT). The mPPT is a battery of six tests including gait speed, 5 times sit to stand (5xSTS), stair climbing, picking up a penny, putting on a coat, placing a book on a shelf, and turning in place. Tests are timed with a stop watch. The mPPT is used to classify elderly individuals frailty level based on physical function. Analysis included univariate correlation between PROMIS PF T-scores and mPPT total score (as well as with each test). Multivariate analysis modeled age and mPPT tests for the best predictor(s) of PROMIS PF. In an additional analysis, PROMIS PF model parameters were also obtained for items that matched selected mPPT battery tests (gait speed - PFC38, PFM23, PFB5r1; stairs climbing - PFA21, PFC32, PFM21; 5xSTS - PFA15, PFC41, PFA41). Significant univariate correlations allowed association of the specific performance based test score with the no difficulty model parameter for each PROMIS PF item.

Results: Of the 6 tests on the mPPT, 5 of 6 showed significant univariate correlations with PROMIS PF T-scores (r values ranged from 0.32 to 0.64), with gait speed showing the highest correlation. The best multivariate model that predicted PROMIS PF was age (p=0.03), gait speed (p<0.01), and stair climbing (p=0.02). Using the equations of the line for gait speed and 5xSTS, a 1.6 m/s gait speed and 14 second 5xSTS corresponded to no difficulty for the selected PROMIS items. For stair climbing, a 9 second time corresponded to no difficulty with the most similar item (PFA21 normal pace stairs) and a 5 second time corresponded to no difficulty for a more difficult item (PFC32 5 flights of stairs), but an even more difficult item (PFM21 10 story building) resulted in an unrealistic estimate of stair climbing time (i.e. <1 second).

Conclusions: This data suggests sufficient alignment between performance based tests and PROMIS PF T-scores for clinical evaluation and goal setting for the items evaluated (i.e. gait speed, stair climbing, and 5xSTS). However, only PROMIS PF items that showed similarity in task to gait speed, stair climbing or 5xSTS demonstrated strong correspondence with those mPPT tests. The strength of the correlations suggests clinical usefulness but not enough accuracy to replace performance testing. Overall these findings improve the clinical application of PROMIS PF by anchoring T-scores to performance based tests of function. Ability to estimate expected physical performance based on patient report significantly adds to the scope and impact of PROMIS for ongoing clinical assessment and prevention.

Real-time short form development: Applying item information to obtain live score-level reliability estimates for evaluating expected short form performance

Michael A. Kallen, PhD, MPH, Brickell T, Lange RT, French L, Carlozzi NE

Northwestern University

Objective: Best practices for short form development suggest a combination of content expertise and detailed item-level psychometrics be used to identify a subset of an item banks items that meet targeted psychometric characteristics. We developed an item information-based approach for item selection that uses an interactive framework to provide on the spot estimates of score-level reliability for varied item sets, enabling a real-time assessment of the psychometric properties of a developed short form.

Methods: The new Traumatic Brain Injury Caregiver Health-Related Quality of Life (TBI-CareQOL) Measurement System Caregiver Strain item bank contains 33 items; a 6-item short form was targeted for development. Score-level-specific item information values were estimated along a theta measurement continuum from -2.8 to +2.8 (T-score metric: 22 to 78); those values were entered into a spreadsheet for subsequent manipulation. Formulas to calculate score-level test information, expected standard error, and expected reliability were then embedded in the spreadsheet, as was a formula to calculate score-level Spearman-Brown predicted reliability when reducing a 33-item measure to a 6-item measure. Item content and supporting item psychometrics (e.g., slope, thresholds, location, and average information) were also presented in the spreadsheet. This interactive spreadsheet was projected during an in-person meeting of content and measurement experts who identified preferred items and could observe, in real time, the expected psychometric performance of a selected item set. This allowed the experts to evaluate the overall performance of preferred item sets, as well as test the impact of potential item replacement strategies.

Results: A final set of 6 items was identified as best meeting the combined goals of appropriate content coverage and reliable assessment across a wide range of the measurement continuum. Emphasis was given to including content that was identified as being clinically important (Caregiver Strain 1 SD above the mean). Estimated score-level reliabilities for the new 6-item short form were: .80 from theta -1.6 to -1.2 (T-score 34 to 38), .90 from theta -0.8 to +1.6 (T-score 42 to 66), .80 from theta +2.0 to =2.4 (T-score 70 to 74), and .70 at theta +2.8 (T-score 78).

Conclusions: This approach to short form development follows best practices use of content and psychometric expertise. It employs item information values to determine expected score-level reliabilities, executing calculations in real time, and thereby providing content and measurement experts on the spot psychometric performance details of the items proposed for inclusion in a short form.

Feasibility and Implementation of PROMIS CAT as Part of Standard of Care in an Orthopaedic Spine Population

Raj Karia, MPH, Heather Gold, Alissa Link, Rachel Lebwohl, Devin Mann, Thomas Errico, James Slover, Joseph Zuckerman, Michael Cantor

NYU Langone Medical Center

Background/Objectives: Patient-reported outcomes (PRO) are defined as reports coming directly from patients about how they feel or function in relation to a health condition and its therapy. Although disease-specific PROs measures have been developed and are well validated to assess the orthopaedic spine population, they are not widely used in routine clinical care. One of the challenges to widespread adoption is the increase in patient burden as the time needed to complete traditional PRO measures can be considerable. In this study, we evaluated the feasibility and implementation of the recently developed Patient-Reported Outcomes Measurement Information System (PROMIS) computer adaptive tests (CAT) as part of standard care in a high-volume othopaedic spine practice. Our objective was to determine the feasibility of implementing PROMIS CAT along with traditional PRO measures in the electronic health record (EHR).

Methods: A multidisciplinary team of clinicians and experts in information technology, user experience, medical informatics and population health worked together to develop a PRO web application. This application integrates with the EHR to allow patients to complete the PROMIS CAT (Physical Function, Pain Interference), PROMIS Pain Intensity Short Form, the EQ-5D, the Oswestry Disability Index (ODI), the Neck Disability Index (NDI) and the Scoliosis Research Society questionnaire (SRS-22) for patients visiting a university-based orthopaedic spine clinic. All questionnaires were collected electronically, using a tablet computer in the patient waiting area prior to the physician encounter. PRO measure scores then

were seamlessly integrated into the patient EHR after completion. Completion rates and time to complete the PRO measures were recorded.

Results: Over a two month period, 743 patients were administered the tablets and 84% of patients were able to complete once the PRO web application was started. The median time required to answer all items in the PROMIS CAT, EQ-5D, plus the NDI, ODI or SRS-22 was 8 minutes and 54 seconds. Feedback from patients completing the application was positive and a high percentage of patients reported that it was very user-friendly. Although a relatively high completion rates was achieved once the application was started, only 37% of all patients coming into the clinic accessed the PRO assessment.

Conclusions: It is feasible to develop and implement technology that integrates PROMIS CAT into the EHR at a large academic orthopaedic spine clinic with broader enterprise-wide rollout planned. A PRO application that includes PROMIS CAT was can be very useful to reduce patient burden. Challenges in widespread adoption of patient-reported outcomes are not solely technology driven and further exploration to streamline the approach to their implementation in the clinical setting is necessary. Further investigation is also needed to understand the advantages and disadvantages on using PROMIS CAT compare to traditional PRO measures.

Linking ASCQ-Me and PROMIS Pain Measures

San Keller, PhD, MS[1], Michelle Langer, PhD[1], Manshu Yang, PhD[1], Ellen Werner, PhD[2]

American Institutes for Research[1], National Heart, Lung and Blood Institute[2]

Objective: We developed the Adult Sickle Cell Quality of Life Measurement System (ASCQ-Me) to address the need for a reliable and valid method to document the patient-reported outcomes (PRO) of adult sickle cell disease (SCD) care and to complement the Patient Reported Outcome Measurement Information System (PROMIS). Pain is the predominant symptom of SCD and the most common reason for seeking treatment. Because ASCQ-Me and PROMIS both have pain short form scales, we undertook a linking study to determine the PROMIS pain score equivalent to any ASCQ-Me pain score.

Methods: Data were collected from 476 adults with SCD at seven geographically diverse sites across the U.S. Categorical confirmatory factor analysis (CCFA) using the item response theory (IRT) graded response model was used to evaluate the combined PROMIS and ASCQ-Me for unidimensionality. Overall model fit was evaluated using the root mean square error of approximation (RMSEA): an RMSEA value 0.06 was considered to reflect good fit and values above 0.10 were considered to reflect poor fit. The S-X2 statistic was used to assess item-level fit, for which a nonsignificant result ($p > 0.05$, adjusted for multiple comparisons using the Benjamini-Hochberg procedure) indicated adequate model fit. A fixed item parameter calibration was used to link the two instruments.

Results: ASCQ-Me and PROMIS Pain summed-scale scores correlated at 0.79, suggesting content similarity sufficient to link their scores. Cronbachs alpha for the two scales combined was 0.96. An initial CCFA model of a single dimension fit reasonably well (RMSEA=0.07). No items exhibited significant misfit, but 6 item pairs, were flagged for local dependence – likely reflecting similarities within scales in how questions were phrased. Thus, a bifactor model was used in which two specific factors corresponded to the scale-specific items. This model fit well (RMSEA=0.06) and had nonsignificant values for all S-X2 statistics. Factor loadings () were much higher on the overall factor than on the specific factors. The explained common variance for the overall factor was 0.84, supporting sufficient unidimensionality for the combined scales. Thus, a fixed item parameter calibration was used to link the two instruments. ASCQ-Me item parameters were calibrated using the graded response model with PROMIS item parameters fixed to PROMIS bank values. A summed score to T-score table on the PROMIS metric for the ASCQ-Me scale was computed based on these linked item parameters. The linked PROMIS T-scores from the ASCQ-Me data had a mean of 58.01 (SD=8.24), which is not significantly different from the actual PROMIS T-scores on the short-form (mean=57.56; SD=8.66).

Conclusions: Analyses indicated items from the PROMIS and ASCQ-Me pain short forms measured a unidimensional construct and scores from the two scales could be linked together. A summed score translation table was provided to allow for those with data from the ASCQ-Me pain short form to estimate the PROMIS pain short form score equivalents.

Assessing the SPADE Symptom Pentad Using PROMIS 4-Item Measures: Prevalence and Minimally Important Difference

Kurt Kroenke, MD[1], Nina Talib, PhD[1], Patrick Monahan, PhD[2], Jacob Kean, PhD[3], David Haggstrom, MD[3]

Regenstrief Institute[1], Indiana University School of Medicine[2], VA Center for Health information and Communication[3]

Objective: The 5 SPADE symptoms (sleep, pain, anxiety, depression, and low energy/fatigue) are among the most prevalent and disabling symptoms in clinical practice. Notably, these symptoms represent 5 of the 7 domains assessed by the PROMIS profiles, signifying their cross-cutting importance across most medical and mental disorders. In this study, PROMIS 4-item measures were used to assess the prevalence of SPADE symptoms in primary care and the correspondence between PROMIS and brief non-PROMIS legacy symptom measures.

Methods: 300 primary care patients participating in a clinical trial completed a 20-item PROMIS questionnaire consisting of the 4-item symptom scale for each of the 5 SPADE symptoms that are part of the PROMIS 29-item profile. Participants also completed the 2-item Pittsburgh Insomnia Rating Scale, the 3-item PEG pain scale, the 2-item GAD-2 anxiety scale, the 2-item PHQ-2 depression scale, and the 4-item SF-36 vitality scale. Participants also completed a 5-item screener rating the severity of each SPADE symptom on a 0 to 10 numeric rating scale (NRS); those with at least 1 symptom with a NRS 4 were eligible for the trial. A PROMIS T-score of 50 is the population norm and each 10 point change represents one standard deviation (SD). Thus, a T-score of 55 (effect size of 0.5 worse than population norm) was operationally defined as a clinically significant threshold. Pearson correlations between PROMIS and non-PROMIS measures were calculated. The association between PROMIS T-scores and NRS scores was analyzed. Linear regression was used to determine the PROMIS T-score change for each 1-point change in the corresponding non-PROMIS measure. The minimally important difference (MID) was examined by estimating the PROMIS T-score change for a 0.5 and 0.2 SD change (moderate and small effect sizes) in the non-PROMIS measure.

Results: Participants largely had multiple SPADE symptoms: the proportion with 0, 1, 2, 3, 4 and 5 threshold-level symptoms (T-score 55) was 5%, 11%, 13%, 18%, 21%, and 31%, respectively. High co-occurrence rates were similar for all 5 SPADE symptoms. Symptom severity was moderately high, with mean T-scores ranging from 56 for depression to 62 for pain. There was a strong linear association between NRS and PROMIS scores. An optimal NRS screening cutpoint was 5 in that the mean PROMIS T-score at this cutpoint exceeded 55 for each SPADE symptom. Correlations were high (0.70-0.86) between each PROMIS scale and its corresponding non-PROMIS scale. A 0.5 SD change in the non-PROMIS measure corresponded to T-score changes on the sleep, pain, anxiety, depression and fatigue scales of 3.9, 4.8, 4.1, 4.1, and 3.9, respectively, and a 0.2 SD change corresponded to T-score changes of 1.6, 1.9, 1.7, 1.6, and 1.6, respectively.

Conclusions: SPADE symptoms are prevalent in primary care and PROMIS 4-item scales are useful for screening and severity assessment. To minimize respondent burden, a 2-step screening process might use single-item NRS scores, proceeding to PROMIS scales for NRS scores 5. PROMIS symptom T-score changes of 1.6 and 4.0 represent small and moderate effect sizes and thereby a possible MID range.

PROMIS: American sign language

Poorna Kushalnagar, PhD[1], David Cella, PhD[2], Samuel Atcherson, PhD[3]

Gallaudet University[1], Northwestern University[2], University of Arkansas[3]

Objective: To culturally and linguistically validate PROMIS in American Sign Language (PROMIS-ASL) for use with deaf/hh adults.

Methods: We used the standard procedures developed at the U.S. National Center for Health Statistics Cognitive Survey Laboratory to culturally adapt and translate items from the Patient Reported Outcomes Measurement Information System (PROMIS). We describe procedures led by a team of primarily deaf/hh and hearing investigators. The PROMIS-ASL version will undergo rigorous validation and psychometric procedures. Analyses conducted on data from the enrollment and 10-day retest assessments will include convergent validation of the PROMIS-ASL version. Using multidimensional exploratory or confirmatory factor analyses, we will identify or confirm the items most likely to comprise each subset. Once the PROMIS-ASL version is finalized, we will compute test-retest reliability using ICC (intraclass correlation coefficient) from two-way random effects ANOVA models.

Results: We produced an accessible patient reported outcomes measure in American Sign Language (PROMIS-ASL) and a culturally appropriate set of items that are relevant to the experiences of DHH users of accessible technology and services.

Conclusions: The final PROMIS-ASL product will be distributed for public use.

Using PROMIS to Assess Quality of Life for Children with Brain Tumor

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

Northwestern University[1], Lurie Childrens Hospital[2]

Objective: Health-related quality of life (QOL) of childhood brain tumor (BT) 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 BT by comparing how it deviates from that of the US pediatric general population, which can help investigators better understand the impact of BTs on QOL. This study aims to evaluate QOL reported by BT using PROMIS - Anxiety, Depression, Fatigue, Peer Relationship, Mobility, and Upper Extremity Function.

Methods: Data from 145 children aged 8-17 (mean age=12.8yr; 54.2% boys; 79.1% white) with BTs from Chicago and Boston metropolitan areas were analyzed. Of them, averaged years since diagnosis=4.3 (18% were <=1 year post-diagnosis), 70.4% received surgery, 78.4% received chemotherapy, 55.8% received radiation, 21.6% received all these three types of treatment while 4.6% did not receive any of them. Patients completed PROMIS measures during their clinical visits. Parents rated their childrens QOL using a single item (poor/fair/good/very good/excellent). T-test and ANOVA were used to compare group differences.

Results: At the group level, BT were less anxious (mean Tscore=43; range: 31-72), less depressed (mean Tscore=45; range: 32-72) and less fatigued (mean Tscore=44; range: 25-74) but similar mobility (mean Tscore=48; range: 21-62), upper extremity function (mean Tscore=49; 26-57) and peer relationships (mean Tscore=50; range: 17-66) when compared to the pediatric general population (mean Tscore=50). Patients who received chemotherapy (versus not) were more depressed (p<0.05); patients who received radiation (versus not) had worse peer relationship (p<0.05). Patients with <=1 year post-chemotherapy had worse anxiety, depression, fatigue, mobility and upper extremity function than those with >1 year post-chemotherapy (p<0.05). Patient with <=1 year post-radiation were more anxious than those with >1 year post-radiation. There was no significantly different scores between types of treatment. Patients who were rated fair/poor QOL had significantly poorer (p<0.001) scores on all domains, except anxiety, than those with other ratings.

Conclusions: At the group level, children with brain tumor did not demonstrate worse QOL when compared to their peers. However, wide ranges of scores were noted, in which patients with inferior QOL were averaged out by those with superior scores. Length since last treatment, especially chemotherapy, negatively impacted BTs QOL. Treatment outcomes were influenced by treatment protocols which were led by diverse characteristics of brain tumors (e.g., sizes and locations). Our next step is to examine influential factors to patients with more inferior QOL which could facilitate targeting interventions to improve QOL of patients with brain tumors.

Interpretation of the PROMIS Physical Function CAT compared to the IKDC-SKF after ACL Reconstruction

Andrew D. Lynch, PT, PhD[1], Volker Musahl, MD[1], Bryan Galvin, ATC[1], Alicia Oostdyk, MPH[2], Freddie H. Fu, MD[1], Lan Yu, PhD[1], David R. Flum, MD, MPH[1], Danielle Lavallee, PhD, PharmD[1], Paul Pilkonis, PhD[1], James J. Irrgang, PhD, PT, ATC, FAPTA[1]

University of Pittsburgh[1], UPMC Center for Sports Medicine[2]

Objectives: To determine the psychometric properties of the PROMIS Physical Function Computer Adaptive Test (PFCAT) and IKDC Subjective Knee Form (IKDC-SKF) in individuals undergoing ACL reconstruction. We evaluated test-retest reliability and minimal detectable change (MDC) in a stable cohort; and responsiveness (effect sizes), minimal clinically important differences (MCIDs), and thresholds for the patient acceptable symptom state (PASS) up to 24 months after reconstruction.

Methods: Reliability and MDC: 140 individuals who had ACL reconstruction at least 2 years prior participated. Participants completed the PROMIS PF Item Bank and the IKDC-SKF at baseline. Participants completed the PFCAT, IKDC-SKF, and a 15-point global rating of change question (GRC) 1 and 3 months after baseline. Participants were identified as stable if they responded hardly any worse (-1), about the same (0), or hardly any better (1) on the GRC. We assessed test-retest reliability with an intraclass correlation coefficient (ICC), and the MDC was derived from the ICC and the standard deviation of the baseline scores.

Responsiveness and PASS: 170 individuals with ACL injury scheduled for surgery participated. Baseline data was collected before surgery, and follow-up 12 and 24 months after surgery. Responsiveness was assessed using Cohens effect sizes. We operationally defined a GRC score of a good deal better (5) or greater as having made a clinically important change. We created a receiver-operator characteristic curve and used the Youden Index to identify the change score with the best balance of sensitivity and specificity for differentiating between those who had made an important change (i.e. the MCID). Similar methods were used to estimate the score on each measure that differentiates between those who indicated that their symptoms were acceptable (i.e. PASS-Yes).

Results: Stable cohorts of 93 and 83 subjects respectively were identified at 1- and 3-months follow-up. Generally, the ICCs for the PROMIS PFCAT were lower (0.55 - 0.68) than the ICCs for the IKDC-SKF (0.85 - 0.90). The ESs for the PFCAT at 12 and 24 months were substantial, but smaller than the corresponding ESs for IKDC-SKF. For the PFCAT, PASS thresholds were estimated to be 53.4 at 12 months (0.96 AUC, 95% CI: 0.91 to 1.0) and 54.3 at 24 months (0.79 AUC, 95% CI: 0.64, 0.96). For the IKDC-SKF, PASS thresholds were estimated to be 82.2 at 12 months (AUC 0.95, 95% CI: 0.89 to 1.0) and 81.6 at 24 months (0.95 AUC, 95% CI: 0.89 to 1.0) Conclusion: The PROMIS PFCAT demonstrates excellent responsiveness after ACL reconstruction and moderate test-retest reliability. The limited questions in the PF item bank that reflect demanding physical functioning may have contributed to poorer reliability and measurement error in comparison to the IKDC-SKF. The IKDC-SKF demonstrates excellent test-retest reliability and responsiveness. MDC and PASS values are similar to previous reports for the IKDC-SKF.

Individuals with Knee Impairments Identify Items (PROMIS) Pain Interference and Physical Function Item Banks

Andrew D. Lynch, PT, PhD, Nathan E. Dodds, BS, Lan Yu, PhD, Paul A. Pilkonis, PhD, James J. Irrgang, PhD, PT, ATC, FAPTA

University of Pittsburgh

Background: The content and wording of the Patient Reported Outcome Measurement Information System (PROMIS) Physical Function and Pain Interference item banks have not been qualitatively assessed by individuals with knee joint impairments. The purpose of this investigation was to identify items in the PROMIS Physical Function and Pain Interference Item Banks that are irrelevant, unclear, or otherwise difficult to respond to for individuals with impairment of the knee and to suggest modifications based on cognitive interviews.

Methods: Twenty-nine individuals with knee joint impairments qualitatively assessed items in the Pain Interference and Physical Function Item Banks in a mixed-methods cognitive interview. Field notes were analyzed to identify themes and frequency counts were calculated to identify items not relevant to individuals with knee joint impairments.

Results: Issues with clarity were identified in 23 items in the Physical Function Item Bank, resulting in the creation of 43 new or modified items, typically changing words within the item to be clearer. Interpretation issues included whether or not the knee joint played a significant role in overall health and age/gender differences in items. One quarter of the original items (31 of 124) in the Physical Function Item Bank were identified as irrelevant to the knee joint. All 41 items in the Pain Interference Item Bank were identified as clear, although individuals without significant pain substituted other symptoms which interfered with their life.

Conclusions: The Physical Function Item Bank would benefit from additional items that are relevant to individuals with knee joint impairments and, by extension, to other lower extremity impairments. Several issues in clarity were identified that are likely to be present in other patient cohorts as well.

Trends in PROMIS Scores in the Early Post-operative Period Following Llateral Aankle Ligament Rreconstructive Techniques

Ashlee MacDonald, MD[1], Jeff Houck, PT, PhD[2], John Ketz, MD[1], Judith Baumhauer, MD, MPH[1], Irvine Oh, MD[1], Adolph Flemister, MD[1]

University of Rochester Medical Center[1], George Fox University[2]

Introduction/Objective: Lateral ankle ligament injuries are common conditions accounting for 25% of musculoskeletal injuries. When conservative management fails and chronic instability ensues, operative treatment is often sought. Though surgical outcomes are generally good following lateral ankle ligament reconstruction, literature suggests current scoring systems for evaluating outcomes and monitoring progression have deficiencies. Patient Reported Outcomes Measurement Information (PROMIS) scores have recently been established as a method of monitoring patient outcomes. The purpose of this study was to evaluate the trends in post-operative PROMIS physical function (PF), pain interference (PI), and depression scores in patients undergoing lateral ankle ligament reconstruction.

Methods: PROMIS scores were prospectively obtained from all patients evaluated in our foot and ankle clinic between February 2015 and October 2016. Using ICD-9/10 and CPT codes, a total of 111 patients who underwent lateral ankle ligament reconstruction were identified. After meeting exclusion criteria (less than three-month follow-up, incomplete PROMIS scores or multiple surgeries), 55 patients were included. PROMIS PF, PI, and depression were evaluated at each post-operative visit. Changes in scores were calculated as compared to baseline pre-operative scores and compared at each follow-up time point using two-way ANOVA. Differences in reconstruction type in patients undergoing allograft (A), modified Brostrm-Gould (BG), or modified Brostrm-Gould augmented with fibertape (BG+FT) were also evaluated.

Results: The average follow-up was 27.05 weeks (range 12-60.1 weeks). 11 patients had >9 months follow-up. Changes in PF were significantly different from baseline at all time-points except for 8-12 week follow-up. PF was significantly worse at 2 and 4-6 week follow-up, and significantly better at >12 weeks follow-up ($p < 0.01$). PI significantly improved from baseline beginning at 8-12 week follow-up ($p = 0.02$). Depression was unchanged from baseline at 2 weeks and 4-6 week follow-up, then significantly improved thereafter ($p < 0.01$). Though not significant, when comparing reconstruction types, there was a trend towards slower improvement in PF in those with BG+FT (n=15), compared to A (n=17, $p = 0.07$) and BG (n=21, $p = 0.051$) at 8-12 weeks. Two patients had other types of reconstruction and were not included in this analysis.

Conclusion: Patients undergoing lateral ankle ligament reconstruction demonstrate significant improvements in PF, PI, and depression PROMIS scores compared to baseline. Patients reached baseline PF at 8-12 weeks follow-up, and significantly improved beyond >12 weeks. PI scores were significantly improved from baseline beginning at 4 weeks follow-up. Depression scores also significantly improved at 8-12 weeks follow-up. BG+FT showed a trend of slower improvement in PF, though not significant. Though longer follow-up is needed, the significant improvements in PF, PI, and depression following lateral ankle ligament reconstruction in our study provides data that can be used for pre-operative counseling and monitoring progression post-operatively.

The Road to Recovery for Bunion Surgery: Data Analytic Plots to Target Patient Progress

Ashlee MacDonald, MD[1], Jeff Houck, PT, PhD[2], Judith Baumhauer, MD, MPH[1]

University of Rochester Medical Center[1], George Fox University[2]

Introduction/Objective: Patient reported outcomes (PROs) can provide information on individual patients progress throughout a treatment course and additionally, with common surgeries, powerful numbers can be generated to provide data analytic curves to provide a recovery road map for patients and surgeons. Those who deviate negatively from the predicted path may have a complication and early intervention can be initiated. Those who deviate positively have the potential to need less physical therapy, early return to sports or work. Hallux valgus (HV) is a common condition of the foot with 4.4 million patients seeking care yearly and surgery is equally common. The purpose of this study was to determine if PROMIS PROs can be used to construct data analytic curves for HV surgery.

Methods: PROMIS scores were prospectively obtained from patients evaluated in a specialty foot and ankle clinic between February 2015 and November 2016. Using ICD-9/10 and CPT codes, a total

of 65 patients with hallux valgus who underwent a bunionectomy by a single surgeon were identified. Those with less than two-month follow-up, multiple procedures during the follow-up period, as well as incomplete PROMIS assessment scores at any time point were excluded, resulting in 34 patients. Using a previously described method, bunionectomy-specific pre-operative cut-off values to achieve and fail to achieve minimally clinically important differences (MCID) in PF with 95% specificity and 95% sensitivity were determined. We then stratified patients based on their pre-operative PF T-scores as above or below the MCID cut-off. PF was evaluated using two-way ANOVA at 4 follow-up time periods and pre-operative cut-offs (above or below MCID cut-off) as factors to establish data analytic curves based on pre-operative scores.

Results: Bunionectomy-specific PF cut-off for 95% specificity of exceeding MCID was 39.6 and 50.2 for 95% sensitivity for failing to achieve MCID. Patients were stratified based on PF T-scores above (n = 13) or below (n = 21) the MCID cut-off of 50.2. Data analytic curves were generated for above the PF cut off and below PF cut off. (Figure 1) Pairwise comparison demonstrated that those starting with a T-score above the bunionectomy specific cut-off had significantly better PF pre-operatively (p <0.01) and again at 6-12 week follow-up (p = 0.02). There were no differences at 1 week or 3-4 week follow-up time points.

Conclusion: This data confirms pre-operative PROMIS PF scores are significant post-operative predictors. While patients with pre-operative scores below the bunionectomy-specific cut-off met MCID changes in PF, their T-scores were significantly lower at 6-12wk follow-up than patients with high pre-operative T-scores. Although longer term follow-up is desirable, this short term follow up suggests a significant clinical impact of using PROMIS scores for pre-surgical decisions as well as provides a road map for recovery for patients and surgeons.

Health-related Quality of Life in the Longitudinal Research on Aging Drivers (LongROAD): PROMIS-29 Profile

Thelma J. Mielenz, PhD and the LongROAD Research Team

Columbia University

Objective: To describe baseline health-related quality of life in a new prospective cohort study of older adult drivers with the aim of understanding the many factors involved in keeping older adults driving safely.

Methods: Five study sites (California, Colorado, Maryland, Michigan and New York) recruited 2,990 older drivers (ages 65-79) and collected health measures both via self-report and performance-based, medications, vehicle inspection and driving data. Follow-up by phone at 1 and 3 years and another in-person at 2 years. This study will focus on the eight domains that the PROMIS-29v2 covers and will be summarized descriptively along with baseline characteristics of the LongROAD cohort, including: Emotional distress (anxiety and depression), fatigue, pain (interference and intensity), physical function, sleep disturbance and ability to participate in social roles and activities.

Results: These current results are from a preliminary dataset. Pain intensity will be added along with the full final cohort. The first domain in the table, Ability to Participate in Social Roles and Activities, is interpreted as a T-score of 60 is one SD better than average (more ability to participate) and a T-score of 40 is one SD worse than average (less ability to participate). This preliminary cohort has one SD better thus more ability to participate.

Conclusions: This study will report the final baseline PROMIS-29 Profile for this new older driver cohort. Once the follow-up data is complete, we plan to look at the trajectories of health-related quality of life for each of the eight domains when older adults stop driving by urban and rural status. We also plan to explore if our PROMIS-29 Profile v2.0 responses can be translated into quality-adjusted life-years.

Association of Four Neuro-Qol Scales and Clinical Performance Measures Obtained in a MS Clinic

Deborah M. Miller, PhD, Lyla Mourany, MS, Stephen Rao, PhD, Jay Alberts, PhD, Malory Weber, MBA, David Schindler, Francois Bethoux, MD, Robert Bermel, MD

Cleveland Clinic

Objective: To determine the association between four Neuro-Qol scales, Upper Extremity, Lower Extremity, Cognition, Depression and four related performance measures administered in the clinical setting of a multiple sclerosis (MS) comprehensive care center.

Background: In 2015 the Cleveland Clinic Mellen Center established a learning health system as a component of our routine care. The method for routine collection of patient health status is called the MSPT. It includes three components. The first component includes the twelve CAT administered Neuro-Qol scales. The second is an iPad-based neuroperformance assessment tool that was designed to simulate the Multiple Sclerosis Functional Composite (MSFC). This MSFC analog includes the Walking Speed Test (WST), the manual dexterity test (MDT), the cognitive Processing Speed Test (PST) and the Low Contrast Visual Acuity Test (LCVAT). The third component includes patient reported demographics, disease characteristics and clinical measures including the PHQ-9, a measure of depression. This analysis provides an additional assessment of the psychometric properties of these Neuro-Qol measures beyond the initial validation testing in 161 clinic attenders. We also assessed the previously established severity classifications for the Neuro-QOL Upper Extremity and Mobility scales.

Methods: Established Mellen Center patients complete the MSPT as part of their routine follow-up visits. We assessed the relationship between the Neuro-Qol Mobility score and the WST score, the Neuro-Qol Upper Extremity score and MDT score, the Neuro-Qol Cognition score with the PST and the Neuro-Qol Depression scale with the PHQ-9. The Spearmans Rho was used to assess the correlation due to the non-parametric distribution of the data. The Kruskal-Wallace test was used asses the performance if the previously established cut points for level of severity for the Neuro-Qol Mobility and Upper-Extremity scores.

Results: Between September 2015 and December 2016, 1245 patients completed the first administration of the Neuro-Qol and clinical performance tests as a component of their clinical encounters. The majority, 70.5%, were female, Caucasian, 86.9%, with a relapsing remitting course of MS, 74.5%. There were moderately strong and statistically significant correlations between the Neuro-Qol scales and the performance tests. The correlation between the Neuro-Qol Mobility score and the WST score was Rho = -0.59, $p < 0.0001$, between the Neuro-Qol Upper extremity and MDT was Rho = -0.55, $p < 0.0001$, between the Neuro-Qol Cognition score and the PST was Rho = 0.24, $p < 0.0001$ and between Neuro-Qol Depression scale with the PH-Q-9 was Rho = 0.76, $p < 0.0001$. The scores for the previously established cut scores for no problems, mild problems, moderate problems and severe problems for Neuro-Qol Upper Extremity was KW = 237.86 $p < 0.0001$ and for the Lower Extremity with WST was KW = 253.47 $p < 0.0001$

Conclusions: While there were statistically significant correlations between the Neuro-Qol scores and the performance measures, the moderate size of the correlations indicate that each type of assessment offers unique information. These data also support the interpretation guidelines for meaningful differences in scores for the Neuro-Qol Upper Extremity and Mobility scores.

Documenting Sickle Cell Disease Health Related Quality of Life (HRQOL) in the Electronic Medical Record

Lynne Neumayr, MD[1], Marsha Treadwell, PhD[1], Priya Shah[2], Michael Rowland[1], Kacie Smith[1], Shanda Robertson[1], Anna-Barbara Moscicki, MD[3]

UCSF Benioff Children's Hospital Oakland[1], University of California, Berkeley[2], University of California, Los Angeles[3]

Background: Sickle cell disease (SCD) is a genetic disorder associated with ongoing hemolysis, anemia, and ischemic re-perfusion injury that is characterized by severe pain episodes, acute chest syndrome, stroke, organ dysfunction and increased risk of infection. In 2014, the NHLBI recommended hydroxyurea (HU) for SCD patients 1+ years of age with genotypes HbSS or HbSB0 to prevent complications and improve HRQOL. Scores on HRQOL measures have been shown to be lower in children with SCD than national norms in research studies but have not been routinely incorporated into an electronic medical record (EMR) within clinic encounters. As part of an effort to harmonize EMR data collection of patient-centered outcomes across the University of California medical settings, we incorporated the PROMIS

Patient Global Health (PGH-7) measure into the EMR at UCSF Benioff Childrens Hospitals. The PGH-7 total score is a composite score of 7 questions about general, physical, mental and social health each scaled from 1 to 5. This brief, reliable measure has been validated for use in pediatric chronic disease.

Objective: This study aimed to build the PGH-7 survey within the EMR and test its feasibility for clinical use, compare PGH-7 scores between patients with SCD prescribed HU and not prescribed HU, and identify other modulators of HRQOL being collected as part of a survey funded by the HRSA Pacific Sickle Cell Regional Collaborative.

Design/methods: Consent was obtained from patients 8+ years of age and their families at the UCSF Benioff Childrens Hospital Oakland Sickle Cell Center and the PGH-7 survey was administered in written form. The PGH-7 survey was built into the EMR as a flowsheet and clinicians entered patients responses.

Results: To date, 99 patients have completed the survey including: 58 with HbSS, 20 with HBSC, 10 with HbSB+ thalassemia, 6 with HbBeta0 thalassemia and 5 with other genotypes. The average age of 50 females and 49 males was 13.5 4.1 years. Average total PGH-7 score was 26.3 4.7 with a mean scores of 3.5 1.1 in physical health and 4.0 1.2 in mental health. A previous study of 3635 children 8 to 17 years old, 25% with chronic diseases reported a mean overall score of 28.7 4.7 ($p < .001$), physical health 4.2 .9 ($p < .001$) and mental health 4.2 1.0 (p=ns). There was a negative correlation at trend level between age and PGH-7: r= -0.19 (p=.06). Genotype, gender and annual family income were not predictors of the PGH-7. 61% of the 64 HbSS and HbSB0 patients were prescribed HU. The mean PGH-7 score for the HU group was 27.0 4.7 vs. 25.2 4.9 (p=ns) in those not prescribed HU.

Conclusions: We demonstrated the feasibility of building the PGH-7 questionnaire into the EMR and collecting data during clinic visits. The average overall PGH-7 score and physical health in patients with SCD was lower than normative data. We plan to collect HRQOL data annually and provide the survey within the patient portal to allow for direct entry by patients.

A Pilot Study of the Feasibility and Variability of the PedsPCF Item Bank as a Screening Tool for Minimal Hepatic Encephalopathy

Daniella Ohnemus[1], Katie Neighbors, MPH[2], Lisa Sorensen, PhD[1], Jin-Shei Lai, MD[1], Estella Alonso, MD[1]

Northwestern University[1], Lurie Children's[2]

Objective: Despite the need for clinical monitoring for signs of minimal hepatic encephalopathy (MHE) in children with portal hypertension, few reliable screening methods exist. The PROMIS Pediatric Cognitive Function (PedsPCF) item bank, a 43-item parent- and child-report questionnaire, could be a useful tool for cognitive screening during clinical follow-up. The present study aims to determine the feasibility and variability of the PedsPCF item bank, as well as to explore its correlation with other neurocognitive tests and clinical indicators of portal hypertension.

Method: Pediatric patients with portal hypertension were identified from the population of patients seen at Lurie Childrens Liver Clinic via a search of the Organ Transplant Tracking Registry (OTTR). A battery of neuropsychological tests (Conners Continuous Performance Test, D-KEFS Trail Making and Color-Word Interference Tests, and Grooved Peg Board Task) was administered to patients, along with a collection of cognitive functioning and health-related quality of life surveys (PedsPCF, Behavior Rating Inventory of Executive Function, and PROMIS Pediatric-25 Profile) for both patients and their parents.

Results: A total of 10 patients were enrolled in the pilot study. All who enrolled completed at least 90% of the full battery within 2 hours. Preliminary results suggest that the full PedsPCF item bank tracks with other neuropsychological tests and surveys. Although the Behavior Rating Inventory of Executive Function (BRIEF) may be more sensitive than the PedsPCF, as evidenced by lower scores relative to the mean, applications from the PedsPCF (e.g., Short Form and computerized adaptive testing) make it more practical for implementation in clinical settings. Combining the PedsPCF Short Form with one neuropsychological test may identify the same patients as the BRIEF in a shorter testing period. The Grooved Peg Board Task appears to offer the highest sensitivity of the neuropsychological tests administered.

Conclusions: The battery shows promise in terms of feasibility and variability. The PedsPCF Short Form in conjunction with the Grooved Peg Board Task may hold the most promise in terms of sensitivity and feasibility for screening in the clinical setting.

Practical Considerations for Use of the NIH Toolbox in Down Syndrome Research

Alison Pritchard, PhD, Matthew Ryan, MS, Crystal Thomas, BA, George Capone, MD

Kennedy Krieger Institute

Individuals with Down Syndrome are at increased risk for sleep-related breathing disorders (SRBD; Shott & Donnelly, 2004); however, the extent to which SRBD contributes to cognitive and psychological decline in this population remains unknown. This ongoing study is designed to assess the association between SRBD and depression with functional decline among adults with Down Syndrome. In this study individuals with Down Syndrome as well as symptoms of major depression and recent functional decline are compared with age-, gender-, and BMI-matched individuals with Down Syndrome without psychopathology or functional decline. Additionally, those participants with depression and functional decline who agree to participate in standard-of-care treatment with psychotropic medication will be followed longitudinally in order to determine the value of this treatment in improving symptomatology, cognition, and adaptive functioning.

In addition to psychophysiological measures (salivary cortisol, EEG) and sleep studies, all participants receive neuropsychological testing at baseline. Neuropsychological measures included in the present study are as follows: NIH Toolbox Dimensional Change Card Sort, Flanker Task, Picture Sequencing Memory, Picture Vocabulary, and 9-Hole Pegboard (all subtests administered via Assessment Center); selected subtests of the NEPSY, Stanford Binet V, and the CANTAB.

Thus far, 8 adults (50% male, 50% controls) ages 18–31 years (mean = 24 years) have participated in the baseline testing phase of this study. One of the 8 participants evidenced such severe functional decline that he was unable to participate in testing, despite considerable effort on the part of the examiner. The average IQ of participants, as estimated by the Stanford Binet V, was 47 (SD = 0).

In this 5-minute rapid-fire presentation, we hope to offer several qualitative considerations relevant to use of the NIH Toolbox with this particular population of adults with Down Syndrome, as follows: 1. Routing participants with significant cognitive limitations to the appropriate start point for each subtest can be challenging. Accurately reporting adults with Down Syndromes level of education, which in many cases involved completion of a high school diploma or certificate, led to overestimation of the appropriate level of difficulty for start points for these individuals, particularly on the Picture Vocabulary and Picture Sequencing subtests. Reading achievement level may better estimate the appropriate start point for these individuals. For the purposes of the present study we utilized a universal entry point of 1st grade for all participants. 2. The pace of stimulus presentation and that required for responding was too fast for several of our study participants who tend to be motorically slow to begin with. This was a particular problem on the Flanker task. 3. Some participants, due to limited cognitive abilities and/or level of functional decline, were unable to successfully complete the practice items, particularly on the Picture Sequencing Memory task, where only 4 participants (3 of them controls) were able to successfully complete the practice items. 4. As a result of the built-in ceilings, the NIH Toolbox subtests were still able to be completed relatively quickly with this population.

Qualitative Assessment of the Reliability and Practicality of PROMIS Questions from Stroke Survivors and Caregivers

Mathew Reeves, PhD, Garrett Reichle, Taylor Seaton, Emily Downes, Amanda Woodward, Michele Fritz

Michigan State University

Background: Further information on the reliability and practicality of PROMIS questions used in specific populations is needed. Practical experience collecting PROMIS data through phone interviews in stroke survivors and caregivers is also limited.

Objective: To examine interviewer observations on stroke survivor and caregiver responses to PROMIS short forms in terms of question comprehension, context, reliability, respondent burden and interviewer burden.

Methods: Data were collected as part of the Michigan Stroke Transitions Trial (MISTT), an ongoing clinical trial designed to test the efficacy of social work case management for stroke patients returning home. Stroke survivors and caregivers completed two comprehensive telephone interviews conducted 7- and 90-days after being discharged home. Short form versions of the PROMIS Global Health Scale (10

questions), PROMIS Self-Efficacy for Managing Medications & Treatments (5 questions), and PROMIS Informational Support (5 questions) were evaluated. Three experienced interviewers provided independent assessments of participants responses to each of the 20 questions. The following 5 domains were evaluated and graded on a 1 to 4 scale: 1) Respondent Comprehension (1= complete understanding, 4= no or little understanding), 2) Question Context (1= no or little trouble applying question, 4= a lot of trouble applying question), 3) Overall reliability of respondent answers (1= very reliable, 4= no or little reliability), 4) Respondent burden (1= no or little difficulty, 4= a lot of difficulty), and 5) Interviewer burden (1= no or little difficulty, 4= a lot of difficulty). Lower scores represent more favorable responses. Mean scores for the interviewer assessments of the 5 domains were calculated for the 20 PROMIS questions (100 total assessments). Mean scores of >2 were considered unfavorable.

Results: Of the 100 individual assessments made by each interviewer, there was perfect agreement on 18, good agreement on 52 (score range=1), and fair agreement on 30 (score range=2). Overall, 87 assessments had favorable mean scores of 2 and 13 had unfavorable mean scores of >2 (range: 2.33 to 3.0). Five of these 13 unfavorable scores were related to the Question Context domain, with 3 of the 5 from the Informational Support short form (mean scores= 2.3, 2.3, and 3.0). With respect to specific PROMIS questions, two had consistently unfavorable domain scores. Mean scores for 4 of the 5 domains related to the question CURRENT level of confidence I can list my medications, including the doses and schedule were unfavorable (range 2.3 to 2.6). Mean scores for all 5 domains related to the question I have someone to give me the information if I need it were unfavorable (range 2.3 to 3.0).

Conclusion: Interviewers were in general agreement that most PROMIS questions had good comprehension and reliability with limited burden. However, respondents struggled with question context, particularly for the informational support short form, indicating that they had difficulty applying the questions despite understanding them. Two PROMIS questions appeared to have limited reliability and practicality when applied to this population of stroke survivors and caregivers.

Concordance of Responses Between PROMIS Self-Efficacy for Managing Daily Activities and Activities of Daily Living in the Michigan Stroke Transitions Trial

Mathew Reeves, B.V.Sc., PhD, FAHA, Taylor Seaton, Michele Fritz, Amanda Woodward

Michigan State University

Background: The Michigan Stroke Transitions Trial (MISTT) is a randomized clinical trial testing case management strategies in acute stroke patients discharged from hospital. Patient-reported outcomes are collected by phone interview at 90-days. Two scales are used to measure patient function: the PROMIS Self Efficacy Activities of Daily Living as a computer adaptive test (CAT) and questions on Activities of Daily Living (ADL) adapted from the Health and Retirement Study (HRS) as a short form. Although the scales ask similar questions, the PROMIS scale asks about ones confidence to do activities while the HRS questions ask about ability. Although conceptually it is important to differentiate between confidence and ability, the interview burden from potentially redundant questions and their impact on data quality is a concern especially for respondents recovering from acute stroke.

Objective: To examine the agreement in responses between PROMIS Self Efficacy for Managing Daily Activities questions and the HRS-based activity of daily living (ADL) questions.

Methods: For analysis, interviews with complete responses to both PROMIS and HRS scales were used (n=78). The original 5 PROMIS response options were grouped into 3 categories: quite or very confident, somewhat confident, and not at all confident. Likewise, the original 5 HRS response options were also grouped into 3 categories: no difficulty, some difficulty, and a lot of difficulty. Cross-classification tables were created and the concordance was quantified by calculating overall agreement (%) and chance corrected agreement using the weighted kappa statistic.

Results: Of the 35 questions in the PROMIS Self Efficacy Activities of Daily Living scale question bank, 16 described a specific activity that could be directly compared to one of the 5 HRS questions. Because the PROMIS questions were administered as a CAT, only 3 questions had sufficient sample size (>40 responses) to allow meaningful comparisons. The PROMIS question, how confident are you to go shopping and run errands was compared to the HRS ADL question how much difficulty do you have doing all your grocery shopping. The overall agreement for this comparison was 65%, with weighted kappa = 0.45, indicating moderate agreement. A second PROMIS question, how confident are you to lift and carry groceries was compared to the same HRS ADL question on difficulty doing all your grocery shopping. The overall agreement was 60%, with weight kappa = 0.31, indicating fair agreement. The last comparison looked at the PROMIS question, how confident are you to perform your household chores with the HRS

ADL question how much difficulty do you have doing work around the house or yard. The overall agreement was 75%, with weighted kappa = 0.60, indicating moderate agreement.

Conclusion: For the three comparisons examined, the concordance between the PROMIS and HRS ADL questions ranged from fair to moderate agreement. The moderate concordance suggests that these concepts are related, however the concordance is not high enough to suggest redundancy. Asking questions on both confidence and ability, although similar, appear to be important to understand the full range of stroke-related disabilities in the early recovery period.

Alternative Approaches to Addressing Non-Normal Distributions in the application of IRT Models

Steven P. Reise, PhD, Ron Hays, PhD

UCLA

When estimating logistic item response theory (IRT) model parameters, it is commonly assumed that the latent trait is normally distributed in the population. When the latent trait is either skewed (e.g., skewed-normal) or non-parametric (i.e., the distribution does not follow any well-known parametric function), item and person parameter estimates are to some degree inaccurate and misleading. We estimated an IRT graded response model (GRM), assuming normality, based on a sample of 1,498 individuals who responded to 29 items from the PROMIS Anger item bank. We then contrasted this business-as-usual model with four alternative models that do not assume normality: (a) a zero-inflated mixture, (b) a log-logistic, (c) a Ramsay curve, and (d) a heteroskedastic-skew model. We found that the zero-inflated mixture and log-logistic models offer viable, potentially more valid, alternatives to the GRM. In contrast, our analyses suggest that the Ramsay curve and heteroskedastic-skew models may not yield accurate results in the present data. The main reason, we believe, is that both approaches attempt to estimate a non-normal latent trait distribution, but the Anger data do not appear to adhere to any easily estimable unimodal skewed distribution. This may be due to Anger being a quasi- or unipolar trait.

Feasibility of Home Health Care Patients' Self-Administration of the PROMIS Global Health Survey

Jennifer S. Riggs, PhD, RN[1], Elizabeth A. Madigan, PhD, RN, FAAN[2], Eugene Nuccio, PhD[3], Jacqueline Gillis, MS[1], Zabrina Gonzaga, MSN, RN[4]

Abt Associates, Inc.[1], Case Western Reserve University[2], University of Colorado Anschutz Medical Campus[3], Lantana Consulting Group[4]

Objective: The Outcome and Assessment Information Set (OASIS) is the standardized data collection instrument required in home health for adult non-maternity patients. The OASIS Field Test is a Centers for Medicare & Medicaid Services (CMS) supported study to establish reliability and validity of items for potential future use in the OASIS. Part of this study is a trial of the Patient Reported Outcomes Measurement Information System (PROMIS) v.1.0 Global Health scale to determine feasibility for using patient-reported outcomes in home health.

Methods: Twelve Medicare-certified home health agencies across Colorado, Massachusetts, Ohio and North Carolina were selected as study sites. The study team trained clinicians at each site who then conducted home visits to patients who consented to participate, collecting assessment data and offering patients the opportunity to complete the PROMIS survey as a paper or web-based option.

Results: Preliminary results are reported for patients who completed surveys at home health admission (n = 117) and a subgroup (n = 43) who also completed surveys at home health discharge. Virtually all patients participating agreed to complete the PROMIS survey and preferred the paper option. At home health admission, patients rated overall health and physical health similarly (good 33%, 34%; fair 37%, 37% respectively). Quality of life and mental health were also rated similarly (good 43%, 43%; fair 23%, 19% respectively). Almost two thirds of patients reported satisfaction with social activities and relationships (good 47%, very good 20%) and slightly more than half reported they carried out social activities and roles well (good 37%, very good 20%). When asked to report the extent they were able to carry out usual physical activities, one third (35%) of patients responded only a little. About half of home health patients reported being bothered by emotional problems (sometimes 39%, often 9%, always 3%) and almost three

fourths reported substantial fatigue (moderate 55%, severe 16%, very severe 1%). Among the four fifths of patients who reported having pain at admission, 50% rated their pain as 5 or higher on a 0 10 scale. At home health discharge, more than half of patients reported improved pain (58%) and improved ability to carry out physical activities (51%), while less than a quarter of patients (23%) reported improvement in fatigue. When compared to the US population age 65 and older (M = 50.5, SD = 9.6, N = 1396), home health patients reported significantly worse physical health (M = 38.5, SD = 4.9, N = 106; P <.001). Additionally, home health patients (M = 45.7, SD = 6.9, N = 108) reported significantly worse mental health than the US population age 65 and older (M = 53.3, SD = 8.6, N = 1394; P <.001).

Conclusions: A self-administered PROMIS survey is feasible among cognitively intact home health patients and provides valuable insight into patients perception of their health and quality of life. Based on these preliminary findings, further testing in a larger home health patient sample, as well as in other post-acute care settings, is warranted.

Cross-Sectional Investigation of the Effects of Psychiatric Service Dogs on War Veterans with PTSD

Kerri Rodriguez, MS, Marguerite O'Haire, PhD

Purdue University

Objective: The objective of this research was to empirically evaluate the effects of psychiatric service dogs on standardized assessments of posttraumatic stress disorder (PTSD) symptomology, depression, quality of life, social health, and sleep quality in military members diagnosed with PTSD. To achieve this objective, we compared two groups: (1) individuals receiving usual care who had been placed with a psychiatric service dog and (2) individuals receiving usual care while on the waitlist. Our hypothesis was that military members with PTSD who have been placed with a service dog will show decreased PTSD symptom severity and better overall mental health and wellness compared to those on the waitlist to receive a PTSD service dog.

Methods: A pragmatic efficacy trial was conducted with 141 post-9/11 military veterans with PTSD (113 male, 28 female) including (n = 76) with a service dog and (n = 66) on the waitlist to receive a service dog. The primary outcome was longitudinal change on the PTSD Checklist (PCL-C) including up to three assessments during the waitlist period and two assessments after being placed with a service dog. A cross sectional survey compared secondary outcomes with the NIH Patient Reported Outcomes Measurement Information System (PROMIS) adult short forms of depression (8A), anxiety (8A), anger (5A), ability to participate in social activities (8A), social isolation (8A), companionship (6A), and sleep disturbance (8A). The human-animal bond relationship was also measured with those with a service dog via the Monash Dog Owner Relationship Scale (MDORS).

Results: Mixed model analyses revealed clinically significant reductions in PTSD symptoms from baseline following the receipt of a service dog, but not while receiving usual care alone. Though clinically meaningful, average reductions were not below the diagnostic cutoff of 50 for PTSD on the PCL. Regression analyses revealed significant differences with medium to large effect sizes among the service dog group, including lower depression (Cronbachs =.91), anxiety (=.79), and anger (=.50) as well as lower social isolation (=.60) and a greater ability to participate in social activities (=.73). No difference was found across groups in subjective sleep disturbance (=.42). With respect to human-animal bond attachment, there was no significant correlation between the MDORS and any outcome variables (all ps >.053) nor between the length of time with a service dog and any single-time point measure (all ps >.12). Follow up data analysis is currently underway for objective physiology via salivary cortisol awakening response.

Conclusions: The results indicated that compared to usual care alone, the provision of psychiatric service dogs was effective in conferring clinically significant reductions in the PCL, though not below the clinical diagnosis cutoff of PTSD. Although significant group differences were found in mental and social health variables such as depression, anxiety, and social isolation, their provision does not appear to be a stand-alone treatment. This research suggests that service dogs may serve as a complementary treatment option as part of an integrative treatment milieu for PTSD among military veterans.

Performance of PROMIS and Legacy Measures among Advanced Breast Cancer Patients and their Caregivers

Alla Sikorskii, PhD[1], Gwen Wyatt, PhD, RN, FAAN[1], David Victorson, PhD[2], Vered Hankin, PhD[2]

Michigan State University[1], Northwestern University[2]

Objectives: The purpose of this research was to compare PROMIS Profile-29 and the corresponding legacy measures of HRQOL including symptoms and functioning and share the experience of using them in a symptom management trial.

Methods: PROMIS and legacy measures were collected at intake into a randomized trial of reflexology delivered by friend or family caregivers to women undergoing treatment for advanced breast cancer (n=256). Post-intervention data were collected at week 5, and follow-up data were collected at week 11 after randomization to caregiver-delivered reflexology or attention control. The ability of two sets of measures to discriminate patient and caregiver subgroups including trial arms was assessed using effect sizes and p-values from t-tests of significance.

Results: Both patients and caregivers had an average age of 55 years and similar numbers of comorbid conditions. The distribution in patient and caregiver samples was highly comparable between PROMIS and legacy measures for physical function, pain, mental health, depression, social roles and fatigue. The majority of correlations between legacy and PROMIS measures exceeded 0.6 for both patients and caregivers. Shorter measures such as PROMIS Profile-29 and single item measures of patient symptoms demonstrated virtually the same known groups validity and similar responsiveness as longer measures of symptoms and functioning.

Discussion: Evidence of validity and reduced respondent burden support the use of shorter measures of HRQOL in clinical and general populations. Many of the shorter measures are appropriate for research and clinical application of interventions to improve HRQOL.

Relationships Between Psychosocial and Physical Function Scores using PROMIS, ASCQ-Me or Older SCD-Specific Measures: the SHIP-HU Study

Wally R. Smith, MD[1], Donna K. McClish PhD[1], Richard Lottenberg, MD[2], Ian Chen, MD[3], Marsha Treadwell, PhD[4], India Sisler, MD[1], Shirley Johnson, BA[1], Daniel Sop, BS[1]

Virginia Commonwealth University (VCU)[1], University of Florida[2], Eastern Virginia Medical School[3], Children's Hospital of Oakland, UCSF[4]

Objective: We tested the relationships between biological and psychosocial function measured using PROMIS and ASCQ-Me and compared them with scores on previously validated psychosocial and medical skill set PROs felt important in SCD.

Methods: All patients (n=214, 53% female) were enrolled in the Start Healing in Patients with Hydroxyurea (SHIP-HU) Clinical Trial. At baseline, 79% were on HU. Baseline physical function assessment included PROMIS Global health physical summary (GPHYS), Physical health (PHYS), Fatigue (FATG), and Pain behavior (PAINB), as well as ASCQ-Me scales for sleep/Wake disturbance (WAKE), Stiffness (STIFF), Sleep impact (SLEP), Pain crisis frequency (PAINF), Pain crisis severity (PAINS), and Pain impairment (PAINI). Baseline psychosocial assessment included the PROMIS Global mental health (GMH) and satisfaction with social roles (SSR) scales, as well as the ASCQ-Me emotional distress (EMOTD) and social functioning (SF) scales. Analyses consisted of correlations with other PROs also completed at baseline: the Smith-Bovbjerg Sickle Cell Stress scale (STRESS), the Social support (SS), Sickle Cell Self Efficacy (SCSES), and Medical Knowledge and Skills (MEDSKILL) scales of the Transition Intervention Program - Readiness for Transition assessment, and three subscales of the Coping Strategies Questionnaire (Emotion-focused, Active and Behavioral (COPEEMOT, COPEACT, COPEBEH).

Results: The strongest correlations of physical domain PROs were with STRESS, SCSES, and COPEEMOT. Physical domain PROs did not correlate with MEDSKILL, COPEBEH, or (usually) COPEACT. Psychosocial domain scores from ASCQ-Me and PROMIS strongly correlated with many scores from older psychosocial measures –STRESS, SS, SCSES, and Emotion-focused Coping. But they did not correlate with Medical Knowledge/Skills, Active Coping, or Behavioral Coping.

Conclusions: In SHIP-HU, Relationships were found between physical domain PROs and stress, self-efficacy, and only emotion-focused coping. Strong relationships were found between ASCQ-Me and

PROMIS psychosocial PROs when they were compared to previously validated and reported psychosocial PROs. Along with these relationships, the relative brevity of the newer measures and the ability to compare PROMIS scores to national norms lends credence to their utility for future SCD research.

Comparison of Single-level and Multilevel Transforaminal Epidural Steroid Injections Utilizing PROMIS Outcomes Data

Nicole Strong DO, Benjamin Strong MD, Rajeev Patel MD, David Speach MD, John Orsini MD, David Mitten MD, Christopher DaSilva, Clifford Everett MD, MPH

University of Rochester Medical Center

Objective: To compare single-level and multilevel TFESI outcomes utilizing the PROMIS domains: Physical Function, Pain Interference, and Depression.

Methods: A single-center prospectively collected database was searched by CPT codes 64483 and 64484 to identify patients who underwent single-level or multilevel transforaminal epidural injections (TFESIs) respectively. Patients who had follow-up between 91-150 days since the date of their final injection were selected. The minimal clinically important difference (MCID) was calculated for both single-level and multilevel injections for each PROMIS domain using the distributive method, defined as the standard deviation of the change. The single-level (SL) and multilevel (ML) groups were then compared using PROMIS domain scores for Physical Function (PF), Pain Interference (PI), and Depression (D).

Results: A total of 266 patients were identified (168 SL, 98 ML). There were 126 male (78 SL, 48 ML), 140 female (90 SL, 50 ML), $p=0.688$. The mean age (years) was 54.6 (SL), 58.1 (ML), $p=0.066$. The mean follow-up was 117 days (SL), 121 days (ML), $p=0.076$.

The baseline PROMIS scores were: PF 36.3 (SL), 37.5 (ML), $p=0.111$; PI 65.4 (SL), 64.5 (ML), $p=0.244$; D 52.3 (SL), 51.9 (ML), $p=0.723$. The change in PROMIS scores were: PF 0.8 (SL), 2.8 (ML), $p=0.037$; PI -2.4 (SL), -4.7 (ML), $p=0.041$; D -0.4 (SL), -2.3 (ML), $p=0.061$. The PROMIS MCID values were: PF 3.6 (SL), 3.8 (ML); PI -4.4 (SL), -4.0 (ML); D -4.2 (SL), -3.9 (ML). The percentage of patients who met MCID were: PF 25.0% (SL), 38.8% (ML), $p=0.018$; PI 34.5% (SL), 48.0% (ML), $p=0.031$; D 28.0% (SL), 30.6% (ML), $p=0.647$.

Conclusion: At 3-5 month follow-up, multilevel TFESIs demonstrated significantly greater improvement in PROMIS Physical Function and Pain Interference domains compared to single-level injections. Additionally, multilevel injections had a greater proportion of patients who achieved MCID for both Physical Function and Pain Interference domains.

Establishing PROMIS Minimal Clinically Important Difference (MCID) Values after Single-level and Multilevel Transforaminal Epidural Steroid Injections

Nicole Strong DO, Benjamin Strong MD, Rajeev Patel MD, David Speach MD, John Orsini MD, David Mitten MD, Christopher DaSilva, Clifford Everett MD, MPH

University of Rochester Medical Center

Objective: To establish PROMIS minimal clinically important difference (MCID) values for single-level and multilevel transforaminal epidural steroid injections (TFESIs) at 1-3, 3-5, and 5-7 month follow-up time periods.

Methods: A single-center prospectively collected database was searched by CPT code 64483 Transforaminal Epidural Block Unilaterally or Bilaterally, First Level Injected. Patients with the additional CPT code 64484, Transforaminal Epidural Block Unilaterally or Bilaterally, Additional Level Injected, were then separated, which resulted in the two cohorts of: single-level only injections and multilevel injections. Patients were further subdivided into three groups: 30-90, 91-150, and 151-210 day follow-up since the date of their last injection. Baseline PROMIS scores: Physical Function (PF), Pain Interference (PI), Depression (D) were compared to PROMIS scores at final follow-up. The MCID was calculated for each PROMIS domain by the distributive method as of the standard deviation of the change. The percentage of patients who achieved the MCID in each PROMIS domain was then calculated.

Results:

Single-level Group - Demographics Group (months) Total Patients Men Women Mean Age (years) Mean Follow-up (days) 1-3 305 145 160 57.8 56 3-5 168 78 90 54.6 117 5-7 87 36 51 54.6 178

Single-level Group - Outcomes Group Baseline PROMIS Score Mean PROMIS change MCID % Met MCID (months) PF PI D PF PI D PF PI D PF PI D 1-3 36.2 65.3 51.7 1.37 -2.48 -1.31 3.7 -3.7 -4.2 26.2 37.7 30.5 3-5 36.3 65.4 52.3 0.84 -2.45 -0.38 3.6 -4.4 -4.2 25.0 34.5 28.0 5-7 36.8 64.9 50.8 1.14 -2.28 -0.36 3.6 -3.7 -3.4 24.1 29.9 23.0

Multilevel Group - Demographics Group (months) Total Patients Men Women Mean Age (years) Mean Follow-up (days) 1-3 161 82 79 58.6 60 3-5 98 48 50 58.1 121 5-7 45 27 18 55.2 179

Multilevel Group - Outcomes Group Baseline PROMIS score Mean PROMIS change MCID % Met MCID (months) PF PI D PF PI D PF PI D PF PI D 1-3 36.2 65.1 51.6 2.44 -2.98 -1.74 3.9 -3.9 -4.2 29.2 37.9 31.1 3-5 37.5 64.5 51.9 2.78 -4.66 -2.33 3.8 -4.0 -3.9 38.8 48.0 30.6 5-7 37.8 65.1 53.2 1.36 -2.56 -1.78 4.2 -3.9 -4.6 26.7 35.6 33.3

Conclusion: MCID values were calculated for PROMIS domains (PF, PI, D) following single-level and multilevel TFESIs at 1-3, 3-5, and 5-7 month follow-up. Combined results after TFESIs indicated the PROMIS MCID for Physical Function ranged between 3.6 and 4.2, Pain Interference between -3.7 and -4.4, Depression between -3.4 and -4.6. This analysis provides a benchmark to understand what is meaningful improvement when using PROMIS as a patient reported outcome measure following interventional spine procedures.

Change in CRP, Cortisol, and Quality of Life Among Cancer Survivors Following 26-Weeks of Exercise

Heather P. Tarleton[1], Jeanette Ricci[1], Arash Asher, M.D.[2]

Loyola Marymount University[1], Cedars-Sinai Samuel Oschin Comprehensive Cancer Institute[2]

Purpose: Post-treatment fatigue, stress, depression and poor motivation collectively act as barriers to participation in physical activity and can lead to poor quality of life (QoL) among cancer survivors. This study examined the effect of combined aerobic and resistance training (CART) on QoL and on biomarkers associated with stress.

Methods: Cancer survivors (n=33) enrolled in group-based CART for three 60-minute sessions per week for 26-weeks. Participants completed the NIH PROMIS and SF-36 surveys and cortisol and c-reactive protein (CRP) were assessed using volunteered blood specimens.

Results: After 26 weeks of CART, social role satisfaction (7.9 14.9%), fear/anxiety (2.3 14%), fatigue (2.1 14.5%) and physical function (1.5 9.6%) improved with increasing attendance at CART sessions. Pain interference, cortisol levels, and CRP levels decreased with increasing attendance at CART sessions. Participants with higher comorbidity at baseline appeared to experience the greatest improvements in fatigue (r = 0.555, p = 0.005), anxiety/fear (r = 0.547, p = 0.005), and social role satisfaction (r = 0.611, p = 0.002). Participants within two years of treatment completion experienced greater improvements in psychosocial wellbeing. After CART, participants reported on the SF-36 that their health was much better than a year ago.

Conclusions: Group-based CART supports improvements in the self-efficacy and QoL of cancer survivors and decreases in biomarkers of stress and inflammation. Oncologists should consider recommending that their patients begin CART as soon as medically feasible following the cessation of cancer treatment.

ASCQ-Me and PROMIS in the Sickle Cell Disease Implementation Consortium Needs Assessment and Registry

Marsha Treadwell, PhD,[1], Ellen Werner, PhD, MA[2]

Children's Hospital of Oakland, UCSF,[1] National Heart, Lung and Blood Institute[2]

Background: The Sickle Cell Disease Implementation Consortium (SCDIC) is the first research program within the National Heart, Lung and Blood Institute to use Implementation Research to promote the adoption of research findings into healthcare. The SCDIC will identify and address barriers to quality care in sickle cell disease (SCD) within its eight funded sites across the U.S. for individuals with SCD ages 15 45 years. The SCDIC is in its first phase, conducting a needs-based community assessment of the barriers to care for youth and adults with SCD and creating case report forms (CRFs) from standard measures for a registry that will include clinical and utilization information, and patient reported outcomes (PROs). In Phase II, we will use implementation research methods to design intervention studies that will be conducted to address issues identified from the needs assessments. The SCDIC Registry will also continue throughout Phase II.

Objective: To identify a uniform system of common data elements, including PROs, for data collection within the SCDIC needs assessment and registry.

Methods: In the first year of the SCDIC, representatives from the eight sites met in a series of face-to-face and video-conferences to discuss objectives of the needs assessment and registry data collection. We defined domains of interest from the needs assessment that would inform the design of the Phase II intervention studies. We identified measures from the PhenX Toolkit, PROMIS and ASCQ-Me that will be used across sites, along with qualitative data collection strategies to conduct the needs assessment. Similarly, we defined common data elements and identified measures from the PhenX Toolkit, PROMIS, ASCQ-Me and Neuro-QOL for the SCDIC Registry. We generated potential research questions that could be answered from the data that will be collected from the 2400 patients to be enrolled in the registry.

Results: The domains that we considered of core importance for the needs assessment and associated measurement included patient experiences of care (ASCQ-Me Quality of Care), patient health metrics (ASCQ-Me Pain Episode Frequency and Severity, PROMIS Pain Interference), demographics and SCD genotype from the PhenX Toolkit, as well as SCD Self Efficacy from PhenX. The needs assessment is currently underway with a target enrollment of n = 526. Common data elements for the registry include measurement from PhenX (SCD Core Tier 1 and 2 demographics), PROMIS (Pain Quality, Depression, Fatigue), Neuro-QOL (Cognitive Function) and ASCQ-Me (SCD Medical History Checklist; Emotion, Pain, Sleep and Social Functioning Impacts; and Pain Episode Frequency and Severity).

Conclusions: The comprehensive HealthMeasures.net systems, as well as the PhenX Toolkit provide us with the foundation for understanding patient needs, and for developing and evaluating strategies that take a multi-modal, multi-sector approach to address the longstanding and pervasive disparities that youth and adults with SCD face in accessing quality healthcare and achieving optimal outcomes.

Multi Center Research: Cognitive Data in the SEARCH for Diabetes in Youth Study

Greta N. Wilkening, Psy.D.[1], Anna Bellatorre[1], Ralph DAgostino[2], Lisa Testaverde[1], Leora Henkin[2], Ken Wilson[1], Dana Dabelea[1]

University of Colorado Anschutz Medical Campus[1], Wake Forest School of Medicine[2]

SEARCH for Diabetes in Youth (SEARCH) is a multi-center population-based study of youth with diabetes diagnosed at age <20 years. Initiated in 2000, SEARCH was designed to estimate the prevalence, incidence, clinical presentation and evolution of diabetes. SEARCH has 5 centers; Washington, Ohio, Colorado (and Navajo Nation), California, and the Carolinas. Funding is provided by the Centers for Disease Control and Prevention (CDC) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

A selected sample of 2294 youth representing diverse racial and ethnic backgrounds was eligible for study thru 2019 in the longitudinal cohort study. The cohort has been followed for up to 16 years. Study visits have included surveys (QOL, work and educational history, psychological adjustment), physical exam (cardiovascular and neuropathy measures, retinopathy screening) and laboratory investigations. Cognitive assessment was added for the current ongoing visit, as studies have demonstrated changes in cognition in patients with diabetes.

There are challenges to completing cognitive assessment as part of a complex lengthy visit, across multiple sites, with limited resources. These challenges include time limitations, lack of trained examiners, acceptability to subjects and staff, technological issues, Wi-Fi availability, cultural and linguistic variability.

The NIH Toolbox-iPad (Picture Vocabulary, Flanker Inhibitory Control and Attention, Dimensional Change Card Sort, Picture Sequence Memory, List Sorting Working Memory, Processing Speed tests) was chosen for assessment. This measure, developed for longitudinal assessment, covers areas of concern in the evaluation of individuals with diabetes, is brief, and was designed to be administered by research staff. Data quality issues persist when testing is administered by examiners without experience in psychological assessment. Though the NIH Toolbox has been used in multi-site studies a search of publications on the NIH Toolbox web site, and through Ovid, and Google Scholar identified 7 empirical multi-site studies published from 2008-2016, with one incorporating a clinical sample.

SEARCH study staff were trained over 2 full-days including direct observation and rating of administration. Knowledge of specifics of administration was initially tested via written examination, with quarterly re-assessment. Common errors (>10%) in understanding testing procedure or purpose, or how to handle specific issues (e.g. visual loss) have been discussed via conference calls. New staff are trained on-site, and assessed via direct on-line observation. Raters are required to demonstrate competence prior to completing study assessments with participants. Cooperation, compliance and complications during assessment are reported to the coordinating center.

For the ongoing cycle, between March and December, 2016, 273 SEARCH visits were completed, with cognitive data available on 252. Completeness of data varies from 90-100% across sites. The standard deviations of test performances (Fully Adjusted T-scores) are within anticipated range (8-12), e.g. good data coherence. Examiners, initially anxious about cognitive assessment, report comfort in administration, and no participant resistance or distress at completing these measures.

The NIH Toolbox can be taught to examiners nave to psychological testing, adherence to testing requirements can be effectively monitored, the Toolbox is accepted by subjects and examiners and initial data suggest that the scores are consistent with anticipated results.

Using PROMIS Global Health in Dyadic Analysis in the Michigan Stroke Transitions Trial

Amanda Toler Woodward, PhD, Michele C Fritz, BSc LVT, Mathew J. Reeves, PhD

Michigan State University

Objective: A growing body of research is examining the interpersonal aspects of stroke recovery within the context of patient-caregiver dyads. One aspect of this is examining the level of congruence between patients and caregivers on their assessment of patient functioning and the extent to which agreement or disagreement influences other patient and caregiver outcomes. The purpose of this study is to examine the agreement between stroke patients and caregivers in the assessment of the patients physical and mental quality of life using the PROMIS Global Health short form.

Methods: Data were obtained from the Michigan Stroke Transitions Trial (MISTT), an ongoing randomized clinical trial testing a social work case management intervention in stroke patients and caregivers discharged from three Michigan hospitals. Phone interviews are conducted with patients and caregivers 7 days and 90 days after discharge home. The 10-item PROMIS Global Health short form is administered to both patients and caregivers in reference to the patient. We generated summary scores using PROMIS guidelines. The Global Physical Health subscale was created by summing responses for four questions: 1) overall health, 2) ability to carry out every day physical activities, 3) level of pain, and 4) level of fatigue. Similarly, summary scores for the Global Mental Health subscale were created by summing responses for four questions: 1) overall quality of life, 2) overall mental health including mood and ability to think, 3) satisfaction with social activities and relationships, and 4) frequency of emotional problems. Raw scores were converted to standardized T scores with higher scores indicating better functioning. Pearson correlations were calculated to compare T scores between patients and caregivers. Paired t-tests were used to test whether the mean difference in T scores between patient and caregiver assessment of patient physical and mental health were equal to zero.

Results: Data are limited to 7 day interviews complete in both patients and caregivers (n=72 dyads). Patient and caregiver scores were significantly correlated on both the physical health (r=.54, p<.000) and mental health (r=.46, p<.001) subscales indicating nonindependece of data and the need to focus on the dyad rather than the person as the unit of analysis. The mean patient assessment of physical health was 45.3 (SD=8.02) and the mean caregiver assessment was 44.3 (SD=8.08). For mental health, the mean patient score was 46.9 (SD=7.49) and for caregivers was 45.2 (SD=9.2). Paired t-test results indicate

no significant difference between patient and caregiver assessment of patient physical (t=1.08, p=.285) or mental health (t=1.66, p=.101).

Conclusions: This preliminary analysis illustrates the potential use of the PROMIS Global Health short form within the context of dyadic analysis. Results indicate agreement on average within dyads on the assessment of patient physical and mental health. Future analysis will determine whether agreement is maintained at 90 days after discharge and if there are consequences for other study outcomes.

Author Index

Joseph, 15
Zukerman

Joseph, 23