[1]MarksMan Healthcare Solutions LLP (HEOR and RWE Consulting), Navi Mumbai, India, [2]MarksMan Healthcare Solutions, Navi Mumbai, India

OBJECTIVES: Existing evidence on the psychometric properties of physical functioning and pain outcomes measures used in the chronic low back pain (CLBP) trials are not been systematically reviewed. The objective of our study was to evaluate the methodological quality of studies that evaluated psychometric properties of functioning and pain outcome measures for CLBP using Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN) check list. **METHODS:** We searched Pubmed and EMBASE databases from inception to June 2015 with specific key words. Longitudinal cohort and cross sectional studies which included at least one assessment of psychometric property of outcomes measure in CLBP patients were included. Studies published in English language and on humans were included. Studies published as reviews, editorials and case reports were excluded. Two reviewers independently performed study selection, data extraction and quality assessment procedures; disagreements between reviewers were resolved through discussion. **RESULTS:** A total of 32 studies met our inclusion criteria. In this systematic review only one third (34%) of instruments were tested for all psychometric properties and showed mixed methodological quality according to COSMIN check list. Among all instruments Quebec Back Pain Disability Scale had showed excellent reliability (Cronbach's alpha coefficient 0.96) and test-retest reliability (ICC=0.92) for physical functioning assessment. Pain Catastrophising Scale had showed excellent reliability (α=0.92) and high degree test-retest reliability (ICC=0.842). More than half of the instruments showed excellent/good reliability and half showed fair validity. Responsiveness was assessed in only nine studies and had all showed fair quality. **CONCLUSIONS:** Overall we found moderate methodological quality for most of the measures to advise tools use based on psychometric properties. Further research is needed to investigate the psychometric properties of all outcome measures used in CLBP research.

PRM149

QUALITY OF LIFE INSTRUMENTS USED TO MEASURE JOB-RELATED STRESS: A SYSTEMATIC REVIEW

Han A, Bae S

Ewha Womans University, Seoul, South Korea

OBJECTIVES: Although job-related stress has been reported to critically influence quality of life (QoL), few studies reported the optimum instruments to measure job-related QoL. The aim of the study was to investigate the type and frequency of instrument used to assesses the level of stress or quality of life related work. **METHODS:** A systemic search of PubMed was employed to find all relevant full texts for recent 5 years focusing on human participants. Search keywords included "job-related stress or work-related stress", and "Quality of life". Information regarding the methods used to quantify stress, instruments used to measure quality of life, demographic and professional characteristics of the target population, as well as the sample size was gathered. **RESULTS:** 50 studies were found during study period, of which 28 of them examined the job related stress and the QoL. Of those 28 studies, health care workers were most frequently studied (19), and diverse methods were used to quantify the level of stress, of which JCQ (Karasek's job content questionnaire), was most frequently used, followed by ERI(Siegrist's Effort-Reward Imbalance) and JSS(Job Stress Scale). Frequently, however, stress level was not quantified at all (5 out of 28 studies). Regarding QoL instruments, SF-12/36 were most frequently used (11 studies), followed by PROQoL(The Professional Quality of Life Scale) and WHOQoL(The World Health Organization Quality of Life)., yet few studies provided explanation regarding the selection of the methods used. **CONCLUSIONS:** Diverse methods have been used to evaluate the impact of job-related stress on the QoL, without proper justification. Methodologically robust studies could enhance the importance of job-related stress on the QoL.

PRM150

COMPARISON OF TWO METHODS FOR ESTIMATION OF WORK LIMITATION SCORES FROM HEALTH STATUS MEASURES

Anatchkova M[1], Fang H[1], Kini N[1], Ware J[1], Bjorner JB[2]

[1]University of Massachusetts Medical School, WORCESTER, MA, USA, [2]University of Copenhagen, Copenhagen, Denmark

OBJECTIVES: To compare two methods for estimation of Work Limitations Questionnaire scores (WLQ, 8 items) from the Role Physical (RP, 4 items) and Role Emotional scales (RE, 3 items) of the SF-36 Health survey. These measures assess limitations in role performance attributed to health (emotional, physical, or both) and different breadth of impact (work vs. work and other activities). We compared WLQ estimates based on an item response theory crosswalk (Method1) and a regression imputation (Method 2). Such estimates can expand the information from studies using only the SF-36 measure, and can inform future data collection strategies. **METHODS:** We used data from two independent cross-sectional panel samples (Sample1, n=1382, 51% female, 72% Caucasian, 49% with preselected chronic conditions, 15% with fair/poor health; Sample2, n=301, 45% female, 90% Caucasian, 47% with preselected chronic conditions, 21% with fair/poor health). Method 1 used previously developed and validated IRT based calibration tables. Method 2 used regression models to develop aggregate imputation weights as described in the literature. We evaluated the agreement of observed and estimated WLQ scale scores from the two methods and their ability to discriminate among known groups of patients. **RESULTS:** Estimated scores from the two methods were strongly correlated (r=.99). Estimated and observed scale scores had strong correlations (r= .68 for RE and r=.76 for RP). Observed and estimated WLQ from both methods successfully differentiated between levels of self reported general health and between patients with and without chronic conditions. For both methods the estimated WLQ means from SF36RP score were closer (and not statistically different for Method1) to the observed WLQ means than estimated WLQ scores from the SF36 RE scale. **CONCLUSIONS:** Our results suggest that both methods provide useful WLQ estimates for group level analysis. Method 1 appears slightly more accurate than Method 2, but is computationally more complex.

PRM151

QUALITY OF LIFE ISSUES IMPORTANT TO PATIENTS WITH HAEMATOLOGICAL MALIGNANCIES

Salek S[1], Kamudoni P[2], Oliva E[3], Ionova T[4]

[1]University of Hertfordshire, Hatfield, UK, [2]Institute for Medicines Development, Cardiff, UK, [3]Hematology Unit, Reggio Calabria, Italy, [4]Multinational Center for Quality of Life Research, St Petersburg, Russia

OBJECTIVES: To date the focus of both development and application of HRQoL instruments has been on research/clinical trials, but their application in clinical practice has been neglected. Therefore a PRO was conceptualized for use in clinical practice. The aim of this study was to explore the views of patients with haematological malignancies (HM) concerning the quality of life issues important to them. **METHODS:** Ethics approval was obtained from Bristol Ethics Committee. Adult patients (with myeloid or lymphoid neoplasm or acute leukemia according to 2008 WHO classification at any stage of disease) capable of reading/understanding English and able to give informed written consent were recruited into the study. This qualitative study employed semi-structured interviews (face-to-face or by phone) and survey (online) with open-ended questions for data collection. Interviews were tape recorded and transcribed verbatim. Content analysis was carried out using Nvivo9. **RESULTS:** 57 patients (male=37; female=20) with mean age of 61.8 (sd=9.1) and mean duration of disease of 3.7 (sd=3.1) were recruited into the study. 14 of the 17 interviews were face-to-face and remaining 40 were online surveys. Of the 57, 46 were multiple myeloma, 8 MDS, one NHL and 2 AML. Issues reported by patients were grouped into two broad categories: symptoms; and QoL. The symptoms were grouped into 5 categories: breathlessness; fractures; spinal problems; anaemia and black urine. 21 QoL issues were reported by patients with the highly prevalent items being 'feeling angry', 'avoid crowds', 'strain on relationships', 'unable to do any activity' and 'doing things slowly'. These QoL issues appear to be unique to such patients. **CONCLUSIONS:** The findings of this study indicate that there are certain symptoms and QoL issues that are specific to patients with HM and thus deserve a dedicated recognition. This also demonstrates the depth and breadth of the impact of HM on physical and psychosocial functional behavior.

PRM152

VALUATION OF THE EQ-5D-5L WITH COMPOSITE TIME TRADE-OFF FOR THE GERMAN POPULATION

Klose K[1], Graf von der Schulenburg J[2], Greiner W[1]

[1]School of Public Health, Bielefeld University, Bielefeld, Germany, [2]Leibniz Universität Hannover, Hannover, Germany

OBJECTIVES: The current value set for the German version of the EQ-5D-3L was generated more than 10 years ago and there is also a demand for a new valuation study, because the questionnaire has been changed from a 3-level to a 5-level version. The aim of this study was to derive a value set for the EQ-5D-5L for Germany. **METHODS:** The study design followed, as one of the first countries, the improved EQ-5D-5L valuation protocol developed by the EuroQol Research Foundation. Data were collected in computer assisted personal interviews with members of the German general population. Participants were asked to each value a randomly selected block of 10 EQ-5D-5L health states using composite time trade-off (cTTO). Econometric modelling was used to estimate values for all 3,125 possible health states described by the EQ-5D-5L. The most appropriate model was determined in terms of logical consistency, goodness of fit and parsimony. **RESULTS:** The values for 86 selected EQ-5D-5L health states were elicited using cTTO from a representative general population sample (n=1,159). Random effects model with main effects represented by 20 dummy variables and a constant performed better than the other models. This model showed high consistency in the sense that the higher the dimensional level the lower the coefficient for all parameter estimates (p-value<0.05). The predicted values ranged from -0.4675 to 0.9876. **CONCLUSIONS:** The improved valuation protocol combined with intensive interviewer training and close data monitoring showed a high feasibility and acceptability to respondents as well as interviewers in Germany and resulted in high data quality. The EQ-5D-5L preference values estimated in this study might facilitate the use of the EQ-5D-5L in a range of applications like in cost-utility analysis for health care policy and clinical assessment in Germany.

PRM153

MEASURING THE IMPACT OF MULTIPLE SCLEROSIS: ENHANCING THE PERFORMANCE OF THE MSIS-29 AND MSWS-12

Cano S[1], Cleanthous S[1], Marquis P[2], Hobart J[3], Naoshy S[4], Mikol D[4], Petrillo J[4], Steiner D[4], Watson C[4]

[1]Modus Outcomes, Stotfold, UK, [2]Modus Outcomes, Newton, MA, USA, [3]Plymouth University, Plymouth, Devon, UK, [4]Biogen, Cambridge, MA, USA

OBJECTIVES: ASCEND is a phase 3, randomized, double-blind, placebo-controlled trial designed to assess whether natalizumab slows disability progression in secondary progressive multiple sclerosis (SPMS). The objective of this study was to use Rasch Measurement Theory (RMT) methods to evaluate the Multiple Sclerosis Impact Scale (MSIS-29) and MS Walking Scale (MSWS-12), patient-reported outcome instruments assessing the impact of MS, and to explore an optimized scoring structure based on empirical post-hoc analyses. **METHODS:** Baseline blinded data from ASCEND (n=889) were analyzed. In stage 1, RMT methods examined: scale-to-sample targeting, item fit, local dependency, and reliability. In stage 2, a post-hoc revision of the scoring structure (MSIS-29 and MSWS-12) and conceptual grouping of items (MSIS-29 only) was conducted and re-evaluated. **RESULTS:** Stage 1 showed adequate scale performance for the MSIS-29 except for item misfit (6 Physical items; 2 Psychological items), suggesting more than two clinical concepts. The MSWS-12 performed psychometrically well except for disordered thresholds (2/12), item misfit (3/12) and mis-targeting (person location range: -5.01 to 5.77; item location range: -2.99 to 4.10). For stage 2, two MSWS-12 items were rescored. The revised MSWS-12 showed improved scale performance (i.e., response categories and item fit), but not targeting. The MSIS-29 was re-categorized into three conceptually clearer sub-scales: 'Symptoms'; 'Psychological'; 'Limitations'. The revised MSIS-29 scoring structure improved targeting and some item misfit (person location range;