400–600 million in direct medical costs and up to $2 billion from treatment-induced reduction in infant mortality. **CONCLUSIONS:** Feeding extremely preterm infants with an exclusively human-milk based diet (fortified with Prolacta HMF™) is highly cost-effective in the prevention of NEC as compared to feeding human-milk fortified with bovine-milk based supplements.

INDIVIDUAL'S HEALTH – Patient-Reported Outcomes Studies

PIH19

IMPACT OF SEVERITY DIFFERENCES ON LOGICAL INCONSISTENCIES IN VALUES FOR EQ-5D HEALTH STATES

Jalundhwala Y[1], Shaw JW[2], Pickard AS[2], Walton S[2], Busschbach J[3]

[1]University of Illinois at Chicago, Chicago, IL, USA, [2]University of Illinois at Chicago, Chicago, IL, USA, [3]Erasmus MC, Rotterdam, The Netherlands

OBJECTIVES: The values assigned to EQ-5D health states may be inconsistent with their logical severity. A logical inconsistency occurs when a health state is logically more severe than another but receives a lower value. The objective of this research was to determine whether logical inconsistencies are related to differences in the severity of paired health states. **METHODS:** One hundred twenty-six Dutch university students used visual analog scales to value the 243 EQ-5D health states. Logical inconsistency rates were estimated for 55 pairs of health states over 1,000 randomly generated vectors of 11 states. Multiple logistic regression was used to model the logical inconsistency of paired health state values as a function of between-state differences in the severity of specific dimensions. **RESULTS:** Over the 1,000 health state vectors, mean inconsistency rates for paired low-, moderate-, and high-severity states were 0.011, 0.220, and 0.045, respectively. Average inconsistency rates for paired moderate-severity states differed from the average rates for paired low-severity states ($p < 0.001$) and high-severity states ($p < 0.001$). For each EQ-5D dimension, the probability of logical inconsistency was positively related to the similarity of the paired states with respect to health problems. The probability of logical inconsistency was higher when both states had moderate or extreme problems in a given dimension than no problems. **CONCLUSIONS:** The probability of logically inconsistent valuations increases with the similarity of the paired states and is higher when both states are moderately severe than otherwise. These findings have implications for the selection of health states in population-based valuation studies.

PIH20

THE PSYCHOMETRIC PROPERTIES OF THE PERCEPTIONS OF CARE ADJECTIVE CHECK LIST–REVISED (PCACL-R) EVALUATED IN A LARGE UK MATERNITY CARE POPULATION

Redshaw M[1], Martin CR[2]

[1]University of Oxford, Oxford, UK, [2]University of the West of Scotland, Ayr, UK

OBJECTIVES: To develop and validate a novel assessment tool for the assessment of perceived care quality of women's experiences of maternity care. The current study reports the psychometric properties of the developed Perceptions of Care Adjective Check List–Revised (PCACL-R). **METHODS:** Data were collected from a UK national survey of women's experiences of maternity care (n = 2960). Confirmatory factor analysis, convergent validity analysis, convergent validity analysis, predictive validity and internal consistency approaches were used to evaluate the psychometric properties of the tool. **RESULTS:** Confirmatory factor analysis demonstrated a excellent fit to a bi-dimensional structure entirely consistent with the negative and positive valencing of adjectives in the measure, consistent with the use of the tool as a two (negative/positive) sub-scale tool. The PCACL-R revealed good convergent validity and excellent divergent validity characteristics. Discriminant validity was assessed against measures of maternal deprivation, partner status and type of delivery and revealed good discriminant validity of the PCACL-R. Internal consistency characteristics of the two PCACL-R sub-scales were acceptable. The predictive validity of the PCACL-R was also excellent. **CONCLUSIONS:** The PCACL-R is recommended as a valid, reliable, respondent acceptable and easy to administer instrument to assess women's experiences of their maternity care.

PIH21

VALIDATION AND COMPARISON OF EUROQOL AND SHORT FORM 6D IN CHINESE CHRONIC PROSTATITIS PATIENTS

Zhao FL[1], Yue M[2], Wu JH[2], Yang H[3], Wang T[3], Li SC[1]

[1]University of Newcastle, Callaghan, N.S.W., Australia, [2]Hospital of PLA, Beijing, Beijing, China, [3]The First People's Hospital of Yunnan Province, Kunming, Yunnan, China

OBJECTIVES: Generic, preference-based Health-related Quality of Life (HRQoL) instruments are increasingly used in decision-making process of clinicians and policy-makers. However, no such HRQoL instrument has been validated or used in Chronic Prostatitis (CP), in which a goal of patient management is to improve the HRQoL. We therefore aimed to assess and compare the psychometric properties of EuroQol (EQ-5D) and Short Form 6D (SF-6D) among CP patients in China. **METHODS:** Consenting CP patients were interviewed using EQ-5D and SF-6D. Convergent and discriminative construct validities were examined with five and two a priori hypotheses respectively. Sensitivity was compared using receiver operating characteristic (ROC) curves and relative efficiency (RE) statistics. Agreement between instruments was assessed with intra-class correlation coefficients (ICCs) and Bland-Altman plot, while factors affecting utility difference were explored with multiple liner regression models. **RESULTS:** In 268 CP patients (median age 32 years) who completed the interview, mean ± SD EQ-5D and SF-6D utility scores were comparable at 0.74 ± 0.13 and 0.75 ± 0.09 respectively. Five of the seven hypotheses for construct validity were fulfilled in both instruments. The areas under ROC all exceeded 0.5 ($P < 0.001$), suggesting that the instruments are sensitive to detect clinically relevant differences in CP patients, with SF-6D having 9.7–19.9% higher efficiency at detecting difference in CP symptom severity. Despite no significant difference in paired comparisons of utility scores between the 2 instruments, lack of agreement was observed with low ICC (0.218–0.630) and Bland-Altman plot analysis. CP symptom severity significantly ($P < 0.05$) influenced differences in utility scores between EQ-5D and SF-6D. **CONCLUSIONS:** EQ-5D and SF-6D are demonstrated to be valid and sensitive HRQoL measures in Chinese CP patients, with SF-6D showing better HRQoL dimension coverage, greater sensitivity and lower ceiling effect. Further research is needed to determine other psychometric properties, such as longitudinal response and reliability.

PIH22

CONSTRUCT VALIDITY OF THE BENIGN PROSTATIC HYPERPLASIA IMPACT INDEX

Angalakuditi M[1], Seifert RF[2], Hayes C[1], Viktrup L[1]

[1]Eli Lilly & Co, Indianapolis, IN, USA, [2]Lewin Group, 13 Innovus, Falls Church, VA, USA

OBJECTIVES: To demonstrate the construct validity of the Benign Prostatic Hyperplasia Impact Index (BII) in the assessment of men with Benign Prostatic Hyperplasia (BPH). **METHODS:** Data from two Tadalafil multi-center, double-blind, placebo-controlled Phase II clinical trials of men 45 years of age or older with moderate to severe BPH and evidence of bladder obstruction (N = 281; N = 1053) were included in this analysis. Measures included BII, International Prostate Symptom Score (IPPS), IPSS Quality of Life Index (IPSS QOL), Lower Urinary Tract Symptoms Global Assessment Question (LUTS-GAQ), uroflowmetry measure peak flow rate (Qmax) and postvoid residual volume (PVR). Data from each study were analyzed separately. Pearson correlation coefficients were computed between the BII scores across visits and the other measures. T-tests and general linear modeling compared BII scores of subjects with global ratings of improvement versus no improvement, and between subjects taking Tadalafil versus placebo. **RESULTS:** There were strong correlations of BII with subjective measures (IPSS, IPSS QOL, LUTS-GAQ) and low correlations of BII with objective measures (Q_{max}, PVR) were observed in this analysis. There were also statistically significant differences in BII at the End-of-Study Visit between subjects who improved versus subjects with no improvement (Studies 1 and 2, $p < .0001$) and between subjects taking Tadalafil versus subjects taking placebo (Study 1, $p = 0.0088$; Study 2, $p = 0.0092$). **CONCLUSIONS:** Results demonstrate that BII is reliable and has construct validity. BII is well correlated with subjective measures and less correlated with objective measures, which is confirmatory of literature findings. It is a valuable measure of the impact of BPH symptoms on health and functioning that can be useful in determining treatment options and outcomes for men with BPH.

PIH23

OVERVIEW OF THE DEVELOPMENT PHASES OF THE OXFORD ANKLE FOOT QUESTIONNAIRE FOR CHILDREN

Morris C[1], Doll H[1], Theologis T[2], Wainwright A[2], Davis N[2], Churchman D[3], Willet K[2], Fitzpatrick R[1]

[1]University of Oxford, Health Services Research Unit, Oxford, Oxfordshire, UK, [2]University of Oxford, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Oxford, Oxfordshire, UK, [3]Isis Innovation Ltd, Oxford, UK

OBJECTIVES: Ankle and foot problems are a common reason for children to present in clinic. The Oxford Ankle Foot Questionnaire for Children was designed to supplement clinical assessment methods to evaluate the effectiveness of interventions. **METHODS:** The development of the questionnaire was conducted in three phases. First the items were devised through focus groups with children affected by foot and ankle problems, and their parents. Second, test versions of child and parent (caregiver) questionnaires were evaluated to enable scales to be developed; three scales were identified and shown to be valid and reliable: *Physical* (6 items), *Emotional* (4 items), and *School & Play* (4 items). Finally, a prospective study was conducted with trauma and elective patients to assess how the scale scores change over time and/or with treatment. **RESULTS:** In our prospective study for the assessment of trauma patients, mean changes in percentage scores were as expected large and all effect sizes were large (>0.8). For elective patients, the mean improvement in scores and effect sizes were moderate. The Minimal Detectable Change (MDC_{90}) ranged from 6 to 8 on the percent scores. The Minimal Important Difference (MID) for elective patients ranged from 7 to 17%. Half the standard deviation of baseline domain scores (often equated to MID) ranged from 11 to 18 on the percent scale. **CONCLUSIONS:** The results from the study demonstrated the longitudinal validity and responsiveness of the domain scales. The Questionnaire uniquely offers an expedient and budget conscious means to evaluate the effectiveness of the treatment of children's foot or ankle problems from both the child and parent/caregivers perspective. The questionnaire has broad utility both in routine clinical settings, or applied research to evaluate treatment programmes and interventions used in paediatric orthopaedics, trauma and rheumatology.