**PMR8**

**DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF A CANINE DERMATITIS QUALITY OF LIFE QUESTIONNAIRE: RESULTS FROM THREE CLINICAL TRIALS**

Postonio-Reeves N1, Brohan K2, Taskoll S2, Panter C2, Amodie D2, Gasper S1

1Zoetis, Florham Park, NJ, USA, 2Adelphi Values Ltd, Manchester, UK

OBJECTIVES: Atopic dermatitis is a common allergic skin disease in canines, charac-
terized by itching and scratching potentially leading to hair loss, excoriation and scalbing. Atopy can substantially impact both the owner and canine’s quality of life (QoL). The objective was to develop a modified 30-item questionnaire completed by owners to assess QoL in canines with atopic dermatitis and their owners. METHODS: A 34-item draft questionnaire was developed through review of the literature and clinical input. This measure was included in three clinical trials of APOQUEL®. A single open-label continuation study in the United States (n=44) and two ran-
domized, controlled, blinded clinical trial in Australia (n=124; n=226). Questionnaire data from each trial was individually subject to item response, dimensionality, and scale analysis to define the final version of the questionnaire. Quality of completion was high (maximum 4.9% missing data). Across all studies, seven items consistently displayed ceiling effects (range: 23.5%-86.3%) and four displayed floor effects (range: 21.6%-70.4%). These results, together with the item discrimination index, confirmatory and exploratory factor analyses (CFA, EFA), item-total correlations and internal consistency, were used to guide item reduction. This resulted in a final model on which variable groups validity and responsive-
tiveness to change were examined. Overall, owners of dogs on APOQUEL demonstrated a positive mean change on six items across all three datasets, indicating improvement in these aspects of QoL. CONCLUSIONS: This study describes the development and psychometric evaluation of a new questionnaire of atopic dermatitis and their owners. Findings from each study, along with consideration of the clinical relevance were used to support item reduction decisions. The final questionnaire then underwent scale-level psychometric evaluation. Consistent improvements in QoL were seen across all three studies.

**PMR9**

**DATA QUALITY WITH MIXED-MODE ADMINISTRATION OF THE SHORT FORM-36**

Bresing H1, Paciorek A1, Carroll P2, Wilson LS1, Litwin MS3, Misakowski CI

1University of California San Francisco, San Francisco, CA, USA, 2University of California, San Francisco, CA, USA, 3University of California Los Angeles, Los Angeles, CA, USA

OBJECTIVES: To examine the effects of survey mode of administration on data qual-
ity when using a mixed-mode (i.e., paper-mode vs. web-mode) approach. METHODS: A cross-sectional vector of 4,836 participants active in the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE) registry, self-selected to complete the Medical Outcomes Short-Form 36 instrument. Data quality parameters com-
pared over time: rate, data completeness at the item and computable scale score; differences in missing items; item response times; and rates of agreement by mode. RESULTS: The majority of participants, 4,376 (90%), opted for paper-mode. Overall response rate was 77%, paper-mode 76% versus 88% web-mode (p < .01). Paper-mode respondents were older, 70.7 (SD 8.8) versus 68.8 (SD 8.5) (p < .001). Paper-mode participants had significantly more missing items (mean 6.7 vs. 10, p < .0001) and significantly lower proportion of computable scale scores for bodily pain, general health, vitalit,y social functioning, and mental health (p < .05). The social functioning scale had the largest difference for the proportion with missing data, 92.3% vs. 95.9% (p = .00). Mean scale scores were lower among paper-mode participants for all eight scales and the physical and mental component summary scores (p < .001). Significantly lower mean scale scores with small differences in effect size were noted for paper versus web-mode for the physical component summary, 47.5 vs. 50.2 (D = .30), physical functioning, 80.2 vs. 88.2 (D = .36), and vitality, 64.8 vs. 71.2 (D = .75). Overall, for the SF-36, the task force recommended four scales as not achievable. Of these, the web-mode, four scales did not achieve a floor value: Vitality (10–100); mental health (28–
100); bodily pain 23–100; and general health (15–100). CONCLUSIONS: We observed significant differences in data quality with mixed-mode survey administration but effect size differences were small which suggests that mixed-mode administration did not introduce significant measurement differences.

**PMR10**

**ALGORITHMS TO ESTIMATE HEALTH UTILITIES FROM TOTAL JOINT ARTHROPLASTY DISEASE-SPECIFIC MEASURES**

Odum SM1, Troyer JI2

1UNC-Chapel Hill, Charlotte, NC, USA, 2UNC-Charlotte, Charlotte, NC

OBJECTIVES: Orthopedic researchers studying total joint arthroplasty (TJA) often use disease-specific measures (DSMs) of patient health as cost-effectiveness outcomes. However, for cost-utility analysis, health utility scores are needed. The objective was to develop regression algorithms to map five common disease specific TJA outcome measures to three preference-based utility scores. METHODS: An online survey was completed by 438 total hip arthroplasty (THA) patients and 550 total knee arthroplasty (TKA) patients. THA patients completed the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC®), Harris Hip Score (HHS), and the Hip Disability and Osteoarthritis Outcomes Score (HOOS). Knee patients completed the Knee Activity Scale Measure or the Knee, Ear, Nose and Throat (K-ENST), Knee Self-Efficacy and Osteoarthritis Outcomes Score (KOOS). All patients completed three prefer-
ence-based questionnaires, the SF-6D, EQ-5D and HUI-3, and responses were used to cal-
culate utility scores. RESULTS: The efficacy algorithms for the SF-6D, EQ-5D, and HUI-3 were estimated for each possible pairing of utilities and DSMs, and prediction perform-
ance was considered to pick the best DSM/utility match. RESULTS: For THA, the regression model with HOOS subscores most precisely estimated an EQ-5D utility. For the HOOS/EQ-5D model, the MAE was 0.06. The SF-6D model, the MAE was strictly nearer than zero. The best performing TKA model mapped the KSS to the EQ-5D. The MAE was 0.05. The SF-6D was 0.06 and the MAE also approached 0.00 for the KSS/EQ-5D. These prediction errors are small which indicates that these models can accurately estimate the EQ-5D utilities. However, they should be used in conjunction with average patient characteristics and not with individual level data. CONCLUSIONS: Clinicians and researchers can input their disease specific data into these models to estimate health utilities to consider the cost-effectiveness of osteoarthritis-related interventions relative to interventions for very different diseases and conditions.

**PMR11**

**EQUATION OF ELECTRONIC AND PAPER ADMINISTRATION OF PATIENT REPORTED OUTCOME MEASURES: A SYSTEMATIC REVIEW AND META-ANALYSIS**

Wild D1, Muehlhausen Wo1, O Donohoe P2, Doll H1

1UCON, Oxford, UK, 2CRF Health, London, UK

OBJECTIVES: To conduct a systematic review and meta-analysis of the equivalence between electronic study in the administration of patient reported outcome mea-
ures. METHODS: A systematic literature review was undertaken in which 1,997 records were identified from which 73 studies met our inclusion criteria. The data were entered into a random effects meta-analysis. Pooling errors and mean difference were estimated. The modifying effect of particular study character-
istics was explored by calculation of pooled values for studies grouped by: mode of administration, year of publication, study design, and time interval between administra-
tion. The random effects pooled correlation coefficient was 0.875 (95% CI 0.856 to 0.883). Correlations were available for 56 studies, with average values still highly variable (P = 91.56). After excluding 20 studies with outlying values the I2 was 56.63, with an overall random effects pooled correlation coefficient of 0.870 (95% CI 0.856 to 0.882). In terms of factors that might explain the heterogeneity, there was a statisti-
cally significant difference in pooled correlation estimates between years, with a tendency for agreement to be greater in more recent studies (fixed p<0.001, random p<0.001). Paper vs. hand-held administration was significantly higher than paper vs web agreement (0.894 vs 0.855), with non-overlapping 95% CI. CONCLUSIONS: The present study supports the previous a previous manuscript (Gwaltney et al 2008) showing that written assessments administered on paper are the same as written assessments on an electronic device. This study has confirmed this con-
clusion and it also broadened our conclusions to computer and paper administration and breaking the comparison down by hand-held vs internet administration. The results suggest that scores obtained by different modalities are directly comparable.

**PMR12**

**DEVELOPMENT OF THE BEHAVIOR RATING INVENTORY OF EXECUTIVE FUNCTION (BRIEF) IN FIVE LANGUAGES**

Acquadro C1, Vasarri S2, Paciorek A3, Misakowski CI

1University of California, San Francisco, San Francisco, CA, USA, 2University of California, San Francisco, CA, USA, 3University of California Los Angeles, Los Angeles, CA, USA

OBJECTIVES: The Behavior Rating Inventory of Executive Function (BRIEF) is the most widely used questionnaire developed for parents and teachers of school-age children to assess executive function behaviors of children and adolescents in the school and home environments. It is composed of 86 items organized in eight clinical scales (Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/ Organize, Organization of Materials, Monitor) and two validity scales (Inconsistency and Negativity). It is used to evaluate children/adolescents (5-18 years old) with a wide spectrum of developmental and acquired neurological conditions. The objec-
tive of this study was to develop the BRIEF in four Balto-Slavic languages (Bulgarian, Latvian, Lithuanian and Serbian) and one Uralic language (Estonian). METHODS: The methodology was used: (1) Clarification of concepts with the develop-
ners, (2) Translation from English to backward translation, (3) Cross-cultural adapta-
tion of the forward translations, one back-translation by an English-speaker fluent in the target language, and review by the developer. RESULTS: The translation process was not reveal any problems with conceptual equivalence. Concepts assessed were cross-culturally relevant. The main difficulties consisted in finding conceptual equivalents of the original items with strong idiomatic content. For instance, the most problematic items were items 18, 55 and 71. Item 71 (‘Lies around the house, doing nothing productive) was translated as written assessments on an electronic device. This study has confirmed this con-
clusion and it also broadened our conclusions to computer and paper administration and breaking the comparison down by hand-held vs internet administration. The results suggest that scores obtained by different modalities are directly comparable.