tion, pain and overall HRQL with patients reporting more morbidity. The results are consistent with the published literature. Future research should focus on intra-rater reliability, responsiveness and validity.

### PCN24

**HANDLING MISSING PATIENT REPORTED OUTCOMES (PRO) DATA FOR PREMATURE WITHDRAWALS FROM CLINICAL TRIALS OF SEVERE DISEASES**

Crawford B1, Massaro J2, Dhawan R1, Gupta S1

1Mapi Values, Boston, MA, USA; 2Boston University, Boston, MA, USA.

PRO data have now become an integral part of clinical trials to evaluate the efficacy of treatment for severe and terminal illness. However, it is often fraught with missing data due to illness, death and early termination of a trial (prior to survey or study efficacy). OBJECTIVES: To evaluate different methods of data imputation when trials are terminated early and there is missing data due to deteriorating health. METHODS: Using data from a large cancer (multiple myeloma) trial which included both missing data due to illness and early termination due to study termination or non-illness-related reasons, several statistical methods were evaluated. A total of 598 subjects completed the EORTC QLQ-C30 at least one post-baseline timepoint and study termination or non-illness-related reasons, several statistical methods were evaluated. A total of 598 subjects completed the EORTC QLQ-C30 at least one post-baseline timepoint and were available for analysis. After setting PRO scores of all subjects died to the worst possible scores, remaining missing data were imputed and analyzed by either multiple imputation (M = 4) using a generalized estimating equation technique, Sun and Song method for censored data, and Pattern-Mixture models. Global Health was the primary PRO endpoint with all other scales adjusted for multiplicity using the Hochberg-Benjamini method. RESULTS: All methods found similar results, although the multiple imputation method found the most number of scales/symptoms significant (N = 10). Sun and Song and the Pattern-Mixture model each found the same four scales/symptoms significantly different between groups—Global Health, Cognitive Functioning, Emotional Functioning and Dyspnea. CONCLUSIONS: All methods are useful approaches to handle missing data imputation and analysis. The multiple imputation method appears to be less conservative, finding ten significant differences versus four with the other methods. The Sun and Song approach provides an insight into what the treatment differences would have been had all the subjects stayed in the study. The Pattern-Mixture model was the most complex method and did not provide any additional information over the other methods.

### PCN25

**RELIABILITY AND VALIDITY OF BRAZILIAN PORTUGUESE HEALTH UTILITIES INDEX (HUI) QUESTIONNAIRES FOR ASSESSING THE HEALTH STATUS OF SURVIVORS OF CHILDHOOD CANCER**

Horsman JR1, Shimoda S2, Furlong W1, Barr RD2

1McMaster University, Hamilton, ON, Canada; 2Centro de Tratamento e Pesquisa Hospital do Câncer, São Paulo, SP, Brazil.

OBJECTIVES: To assess inter-rater reliability and discriminative validity of Brazilian Portuguese HUI questionnaires for a sample of survivors of childhood cancer. METHODS: A sample of consecutive patients attending the long-term follow-up clinic at Centro de Tratamento e Pesquisa Hospital do Câncer in São Paulo was recruited. Self- and proxy-assessment versions of self-complete, one-week health-status recall HUI questionnaires were used. A questionnaire was completed independently by each patient (self-assessment), and nurse and physician (proxy-assessments). HUI single-attribute utility scores (n = 14) and health-related quality of life (HRQL) scores (n = 2) were derived using standard coding algorithms and published utility functions. Single measure, one-way random effects models were used to calculate intra-class correlations (ICC) of inter-rater agreement. Analysis of variance (ANOVA) assessed the statistical significance (p < 0.05) of differences in mean utility scores for patients in various diagnostic groups (n = 15). RESULTS: HUI data for 138 cancer survivors (45.7% female) were collected from the three types of assessors. Patient mean age at diagnosis was 6.3 (min = 0.2, max = 17.4) years and 22.8 (13.4, 40.2) years at survey. There was substantial to almost perfect agreement (ICC > 0.85, p < 0.001) between all pairs of raters for all types of utility scores. ANOVA detected statistically significant (p < 0.013) differences among the diagnostic groups, in means of HUI2 and HUI3 HRQL utility scores, HUI3 vision (p < 0.0005), HUI3 ambulation (p = 0.002), HUI2 mobility (p = 0.001) and HUI2 self-care (p < 0.045). CONCLUSIONS: There was substantial or better agreement in all single-attribute and HRQL utility scores between pairs of patient, nurse and physician assessors. Differences in mean HRQL among diagnostic groups is evidence of discriminative validity. The Brazilian Portuguese HUI questionnaires should be considered for future PRO-studies of morbidity and HRQL in survivors of cancer in childhood in Brazil. Future research on the measurement properties of the Brazilian Portuguese HUI questionnaires should focus on concurrent validity and other health problems.