compare guideline-adherent and routine clinical practice. Value of information analysis will be employed to identify areas for future research. We are applying this translational approach to various current and developing examples of PM in cancer: 1) trastuzumab for human-epidermal growth factor receptor-2 positive BC; 2) gene-expression profiling to identify patients who will benefit most from adjuvant treatments in BC; 3) cytotoxic P450 5D6 testing to select patients for adjuvant tamoxifen therapy in BC; and 4) testing for Lynch Syndrome in CRC patients and their family members to inform treatment and preventative interventions. This research will develop evidence-based information for patients, providers, industry, researchers and policymakers to objectively assess how PM can be beneficial and efficient in improving cancer outcomes.

PODIUM SESSION III: RESEARCH ON THE USE OF UTILITY MEASUREMENT

ASSOCIATION BETWEEN UTILITY AND TREATMENT AMONG PROSTATE CANCER PATIENTS

Jaydevappa B, Chipare S, Malkowicz B
University of Pennsylvania, Philadelphia, PA, USA

OBJECTIVES: To analyze the association between utility, treatment and generic and prostate-specific health-related quality of life (HRQoL) among prostate cancer patients. METHODS: In this longitudinal cohort study we recruited 201 (645 yrs) newly diagnosed prostate cancer patients from urology clinics of an urban academic hospital. Participants completed Quality of Wellbeing (QWB-SA), generic (SF-36), and prostate-specific (UCLA-PCI) HRQoL surveys prior to treatment and up to 24 months post-treatment. Baseline and demographic data were obtained via medical chart review and utility scores were computed using QWB-SA. To analyze the relationship between treatment and utility we used linear mixed effects models, after adjusting for covariates. Similar models were used to examine association between generic and prostate-specific HRQoL and utility. RESULTS: Mean baseline utility was comparable between radical prostatectomy (RP) and external beam radiation therapy (EBRT) groups (0.73 vs. 0.69, p = 0.1750). Mixed effects models indicated that RP was associated with higher utility at 24 month (OR = 1.12, p = 0.027), after controlling for covariates. RP was associated with improved functioning for role physical, role emotional, vitality, mental health and bodily pain and impaired urinary function. Higher scores on generic health subscales were indicative of higher utility. Also, for prostate-specific HRQoL, higher scores on bowl function, sexual function, urinary bother and bowel bother were associated higher utility. CONCLUSIONS: Treatment appears to have significant association with post-treatment utility. Thus, utility assessment provides an entry into economic evaluation. ICERs.

RELIABILITY OF HEALTH UTILITIES INDEX (HUI) SCALING: PARENT AND PATIENT INTER-RATER AGREEMENT ACROSS TWO CLINICAL TRIALS OF TREATMENT FOR ACUTE LYMPHOBlastic LEUKEMIA (ALL) IN CHILDHOOD

Horsman JB, Rae CS1, Furlong W2, Silverman LB2, Sallan SE2, Athale U2, Barr RD3
1Health Utilities Inc, Dundas, ON, Canada; 2McMaster University, Hamilton, ON, Canada; 3Harvard Medical School, Boston, MA, USA

OBJECTIVES: To assess differences in the reliability of HUI Mark 3 (HUI3) health-related quality of life (HRQL) utility scores for patients between self and parent assessments across two Dana-Farber Cancer Institute (DFCI) clinical trials for treatment of ALL during childhood. METHODS: Patients were enrolled in either the DFCI 95-001 or 00-001 trials, and were ≥12 years of age at the time of HUI survey. Patients, parents, and blind to each other, completed HUI questionnaires at each of 5 trial phases: induction; CNS prophylaxis; intensification; and post-treatment. Reliability was assessed in terms of inter-rater agreement of individual scores and summaries in mean scores. Agreement was quantified for the single-way mixed model intra-class correlation coefficient (ICC). An ICC of 0.41–0.60 represents moderate reliability, 0.61–0.80 good reliability, and 0.81–1.00 very good reliability. Mean differences of ≥0.03 are clinically important. Statistical significance was set at p < 0.05. RESULTS: The number of patient and parent paired assessments varied by assessment phase for both the 95-001 (minimum = 29, maximum = 30) and the 00-001 (minimum = 28, maximum = 54) trials. ICCs in the two trials ranged from 0.49 (p = 0.05) to 0.88 (p < 0.05). There was substantial overlap of ICC 95% confidence bounds across the two trials at each of the five assessment phases. There was no significant difference (p > 0.06) between patient-parent pairs of scores at any assessment phase in either trial. The difference between trials in mean parent-parent scores was 0.503 and insignificant (p > 0.05) for each of the 5 assessment phases. CONCLUSIONS: Agreement between parent and parent scores was moderate or better for all assessment phases in both trials. There were no important differences in mean patient and parent scores for any of the assessment phases of the two trials. Inter-rater reliability of scores was similar across the two trials. Parental assessments provide acceptable and consistent estimates of HRQoL for children.

CONCEPTUAL PAPERS & RESEARCH ON METHODS – Clinical Outcomes Methods

EVALUATING CLINICIAN REPORTED OUTCOME (CRO) ENDPOINTS FOR FDA REGULATORY APPROVALS

Nixon A, Gallop K
Oxford Outcomes Ltd, Oxford, Oxon, UK

OBJECTIVES: Clinician reported outcomes (CROs) are the most commonly observed endpoints in FDA approved product labels. Little work has been done to adequately scrutinize in terms of their suitability as endpoints. This study evaluates four widely used CROs in order to assess their suitability as endpoints for regulatory approvals. METHODS: Published evidence on the Karnofsky Performance Status