and the variance-covariance matrix from the model. RESULTS: The advantages of this new approach over the traditional methods are two-fold: 1) The variance of Q-TWiST can be estimated directly, and 2) It is possible to restrict or to compare parameters across different health states. CONCLUSION: Q-TWiST represents a unique quantitative method to simultaneously evaluate the risks and benefits of treatment on successive health states. We present a single survival model that addresses data censoring, covariate adjustment and variance estimation. This new approach provides opportunities for expanded use of the Q-TWiST method beyond the clinical trial such that application to observational studies is possible.

Health Related Quality of Life Based Patient Reported Outcomes: Session 1

DETERMINING THE MINIMALLY IMPORTANT DIFFERENCE OF THE OVERACTIVE BLADDER QUESTIONNAIRE (OAB-Q)
Coyne KS1, Matza L1, Kopp Z2, Jumadilova Z2, Thompson C1, Khullar V3
1The MEDTAP Institute at UBC, Bethesda, MD, USA; 2Pizer Inc, New York, NY, USA; 3Imperial College, School of Medicine, London, UK

OBJECTIVE: The Overactive Bladder Questionnaire (OAB-q) assesses symptom bother and health-related quality of life (HRQL) among patients with overactive bladder (OAB) and has been shown to be reliable, valid, and responsive. The purpose of this analysis was to establish the instrument’s minimally important difference (MID), which is the smallest difference in score that patients perceive as beneficial. METHODS: Data were from four clinical trials (1 US, 3 international), totaling 3426 patients. All trials involved 12 weeks of tolterodine treatment. Distribution-based (e.g., effect size, standard error of measurement, SEM), and half-standard deviation (half-SEM) approaches were used. RESULTS: The mean age of the 4 trial samples ranged from 58.7 to 61.3 years. Patients were predominantly female (65.0%) and Caucasian (87.8%). At baseline, half-standard deviation of the OAB-q Symptom Bother subscale ranged from 9.4 to 9.9, and SEM ranged from 7.0 to 7.28. Half-standard deviation of the HRQL subscales (Coping, Concern, Sleep, and Social Interaction) ranged from 9.8 to 13.6, with SEM ranging from 6.8 to 8.3. The OAB-q subscales had significantly greater OAB-q change scores associated with greater patient perceived treatment benefit and satisfaction. The difference between change scores of patients perceiving “no benefit” and “little benefit” ranged from 6.9 to 16.8 for all scales except Social Interaction (4.3 to 7.8), with the majority of differences greater than 10 points (possible subscale scores range from 0 to 100). Greater OAB-q change scores were associated with greater improvements in micturition diary variables. CONCLUSIONS: Multiple methodologies provide strong justification for recommendation of a 10-point MID for all OAB-q subscales. This MID may be conservative for some subscales, however a uniform MID is recommended to facilitate instrument use and interpretation.

OBJECTIVES: Clinical trials often include condition-specific measures that lack the attributes required for economic evaluation. This study reports on the investigation of models for estimating EQ-5Dindex (a generic, utility-weighted index of health status) from FACT-G and QLQ-C30 (two widely used cancer-specific measures.) METHODS: As part of an observational study in colorectal cancer FACT-G, QLQ-C30 and EQ-5D were administered to 223 patients in an NHS hospital one week after discharge following surgery. Alternative models for estimating EQ-5Dindex from the items of both cancer-specific questionnaires were evaluated. Items were treated as both continuous data or were dichotomised. These were entered as independent variables in a series of stepwise linear regression analyses with EQ-5Dindex as the dependent variable. Criteria for comparing model performance were pre-specified and included explained variance (r²), Pearson correlation coefficient (r) and mean absolute difference. RESULTS: More than 55% of the variance was explained by the model which employed continuous items and the estimated EQ-5Dindex correlated well with actual scores (r > 0.65). Items from FACT-G and QLQ-C30 performed equally well in this model. Less variance was explained by the model using dichotomised items (r² < 0.40) and the correlation between estimated and actual EQ-5Dindex was less pronounced (r < 0.45). Items from QLQ-C30 performed better than those from FACT-G in this model, but altering the cut-point to dichotomise the items had an appreciable effect on explained variance. Despite moderate to good correlation between actual and estimated EQ-5Dindex in both models, the mean absolute difference between these scores gave some cause for concern. CONCLUSIONS: Despite differences in data item content, both QLQ-C30 and FACT-G generate viable estimates of EQ-5D. This is an important finding that allows for the conversion of data from studies where direct comparison is otherwise impossible. However, the technique requires further refinement so as to improve its performance.

"FAMIDIAL STUDY": ANALYSIS OF THE DIFFERENCES BETWEEN DIALYSIS PATIENTS AND THEIR CAREGIVERS (FAMILY CARERS, NURSES AND DOCTORS) ON DIALYSIS PATIENTS’ HRQOL AND OF THE FAMILY CARERS’ HRQOL AND BURDEN
Alvarez-ude F1, Rebollo P2, Valdes C2, Estebanez C1
1Hospital General de Segovia. Institute “Reina Sofia” for Nephrological Research, Segovia, Spain; 2Hospital Universitario Central de Asturias and Institute Reina Sofia for Nephrological research, Oviedo, Asturias, Spain

OBJECTIVE: To assess the agreement between patients and caregivers (Familiar-FAM, Nurse-NUR and Physician-PH) on patients’ HRQoL and to evaluate HRQoL and burden of FAM. METHODS: 221 patient-carrier pairs stratified by age and gender were randomly selected from 14 dialysis units: 132 hemodialysis–69 peritoneal dialysis. Patients’ HRQoL was evaluated by patient and FAM, NUR and PH using EQ-5D (dimensions and Visual Analogue Scale-VAS). Patients and FAM answered the SF-36 (PCS-MCS) and Duke-UNK Functional Social Support (FSS), FAM also answered Caregiver Burden Interview of Zarit (ZS); PH, the co-morbidity Index of patients; and NUR, patient’s dependence in daily activities using the Barthel Scale (BS). RESULTS: The correlation coefficients between the VAS of the patients and their carriers were 0.42 (FAM), 0.49 (NUR) and 0.30 (PH); NUR and PH scored VAS higher than patients (p < 0.01). The agreement (Kappa) between EQ-5D dimension scores varied between moderate for Mobility (0.56-FAM; 0.55-NUR; 0.47-PH) and Self-Care (0.48-FAM;
VARIABILITY IN QOL QUESTIONNAIRES AND THE HANDLING OF MISSING DATA IN PATIENTS WITH NON-SMALL CELL LUNG CANCER TREATED WITH CHEMOTHERAPY

Johnson LL¹, Miller KL¹, Miller DP¹, Colwell HH¹, Bhattacharyya SK², Deeter RG², Mathias SD¹
¹Ovation Research Group, San Francisco, CA, USA; ²Amgen, Inc, Thousand Oaks, CA, USA

OBJECTIVES: To review quality of life (QoL) questionnaire use and methods for handling missing data in published chemotherapy studies [2000–2005 (Jan)] of non-small cell lung cancer (NSCLC). METHODS: We conducted a detailed search of the published literature on quality of life, NSCLC, and chemotherapy, excluding studies with limited total sample size (<30). Thirty-five studies covering 10 QoL questionnaires and 17 chemotherapy drugs were reviewed. RESULTS: Most studies were conducted in Europe (primarily Italy and the U.K.), and the average sample size was 137 patients per treatment arm (range: 16, 406). QoL questionnaires were typically self-administered at baseline and every 3 weeks thereafter until disease progression (roughly 2 years duration). The EORTC QLQ-C30 and EORTC QLQ-LC13 (QLQ-LC13) questionnaires were most often used in Europe (15 of 24 European studies); however, the FACT-L questionnaire was commonly used in the U.S. (7 of 13 U.S. studies). The EORTC QLQ-C30/QLQ-LC13 questionnaires were most often used in Europe (15 of 24 European studies); however, the FACT-L questionnaire was commonly used in the U.S. (7 of 13 U.S. studies). The EORTC QLQ-C30/QLQ-LC13 (QLQ-LC13) questionnaires were most often used in Europe (15 of 24 European studies); however, the FACT-L questionnaire was commonly used in the U.S. (7 of 13 U.S. studies). The EORTC QLQ-LC13 combination detected significant differences for QoL outcomes in 15 out of 15 studies which reported EORTC domain scores, while the FACT-L detected differences in 3 of 7 studies. Both questionnaires contain similar domains, but the EORTC QLQ-C30/QLQ-LC13 includes more items on chemotherapy symptoms and FACT-L includes items on patients’ attitude towards their cancer. Several studies examined potential non-random bias due to missing data resulting from death or disease progression. One study detected conflicting results when including (versus excluding) missing data from their analyses. Another study observed different QoL results when comparing cycle-to-cycle data at similar time points rather than study completion data using the last observation carried forward. CONCLUSIONS: Agreement between patients and carriers was moderate for “objective” dimensions of HRQOL and low for “subjective” ones. HRQol of FAM is slightly worse than that of the general population. The burden of FAM depends on perceived social support, patient’s age and health status of carriers and patient.

A10 Abstracts

THE COST-EFFECTIVENESS OF DRUG-ELUTING STENTS BASED ON THE A SYNTHESIS OF THE RESULTS OF 15 RANDOMISED CONTROLLED TRIALS

Hawkins NJ, Sculpher M
University of York, York, UK

OBJECTIVES: To inform coverage decisions regarding drug-eluting coronary stents by assessing their cost-effectiveness for particular high-risk patient subgroups. METHODS: The analysis compared the cost-effectiveness of the two drug eluting stents (DES) available in the UK (Cypher and Taxus) and bare-metal stents (BMS). A review was undertaken of randomised trials including one or both of the DES. Individual patient data were available for some of these trials. The model parameters extracted from the trials were further revascularisation rates, undertaken for clinical reasons, either percutaneous interventions or bypass surgery; during follow-up. Three subgroups were considered on the basis of their higher baseline risk of repeat revascularisation: small vessels, long lesions and diabetics; a fourth subgroup without these risk factors was also considered. The evidence synthesis was implemented as a Bayesian hierarchical logistic model. A probabilistic decision model was developed. RESULTS: Evidence from 15 trials was incorporated into the synthesis, 5 of which supplied individual patient data. The probability of repeat percutaneous interventions was lowest with Cypher and highest with BMS. Both DES had lower probabilities of subsequent bypass surgery than BMS, and the probability for Taxus was slightly lower than for Cypher. Cost-effectiveness results were sensitive to the stent prices. At current UK prices (£908 BMS, £1,300 Taxus, £1,341 Cypher), Taxus is dominated by Cypher in diabetic patients and subject to extended dominance in all other subgroups. Cypher has an incremental cost per QALY gained, compared to BMS, of £13,759, £23,086, £13,740 in patients with small vessels, long lesions and diabetics, respectively. In patients with no risk factors, this increases to £35,865. CONCLUSIONS: Given existing list prices, Cypher is likely to be considered more cost-effective than BMS and Taxus in patients at high baseline risk of further revascularization. This conclusion may change following any changes in the relative stent prices.

PREDICTING THE COST-EFFECTIVENESS OF THE ABT-578 COATED DRIVER CORONARY STENT (ENDEAVOR) IN DE NOVO NATIVE CORONARY ARTERY LESIONS

Remák E¹, Hutton J¹, Oliver E¹, Brasseur P¹
¹The MEDTAP Institute at UBC, London, UK; ²Medtronic Ltd, Watford, UK; ³Medtronic AG, Tolochenaz, Switzerland

OBJECTIVES: Endeavor combines the Driver stent, the drug ABT-578 and a PC polymer into a new drug-eluting coronary artery stent (DES) system. As results from long-term trials of Endeavor are not yet available, economic modeling techniques must be employed to evaluate the cost-effectiveness of this new treatment option. The study undertook the first assessment of the cost-effectiveness of Endeavor DES, compared to the Driver bare-metal stent (BMS), in the treatment of de novo lesions in native coronary arteries. METHODS: A Markov model was developed simulating the results of the ENDEVOR II trial at nine months follow-up, and extrapolating to five years using the BENESTENT I trial. No difference was assumed between the outcomes with the two stents after the first year. The events modeled were target vessel revascularisations, AMIs, cere-