METHODS

METHODS—Quality Of Life/Patient-Reported Outcomes/Utility Studies

PMD15

RELIABILITY AND VALIDITY OF THE EQ-5D SELF-REPORT QUESTIONNAIRE IN TAIWAN POPULATION

Chang Tj, Tarn YH
National Defense Medical Center; Taipei, Taiwan

OBJECTIVE: To assess the reliability and validity of a Taiwan traditional-Chinese version of the EQ-5D instrument in general population. METHODS: The survey was conducted as a postal survey in a sample of 12,923 persons in Taiwan in December 2002 using a standardized questionnaire containing the EQ-5D, the Short Form 12 Health Survey (SF-12) and a demographic profile. To assess the validity of the EQ-5D, or visual analogue scale (EQ-VAS), the SF-12 scores and other socio-demographic variables were examined using Pearson correlation coefficient. Test-retest reliability was analyzed in a subgroup of 302 respondents measured twice within one month, using kappa and agreement method. RESULTS: The general survey response rate amounted 1647 of the total 12,923 was 12.7% (Female: 50.4%; mean age: 42.4 years). The response rate for test-retest reliability was 61.3%. Subjects reporting moderate or extreme problems for EQ-5D dimensions generally had lower SF-12 scores than those without such problems. Subjects with more chronic health problems had lower EQ-5D value and EQ-5D VAS score. The responses on EQ-5D mobility, self-care, usual activity and pain/discomfort dimension had higher correlation with SF-12 PCS than MCS; however, the responses on EQ-5D anxiety dimension had higher correlation with the SF-12 MCS than PCS. Agreement for test-retest reliability of the EQ-5D items ranged from 76% to 99% (n = 180, interval: 30 days). CONCLUSIONS: The traditional Chinese EQ-5D self-classifier appears to be a valid and reliable measure of the health status in Taiwan general population. These data provide a basis for further studies of the Chinese EQ-5D instrument.

PMD16

INVESTIGATING THE CEILING EFFECT IN THE EUROQOL IN THE U.S GENERAL POPULATION

Bharmal M, Thomas III J
Purdue University, West Lafayette, IN, USA

OBJECTIVES: The presence of a ceiling effect in the EuroQol (EQ-5D) was investigated. METHODS: Data used were from the 2000 Medical Expenditure Panel Survey (MEPS) that included the 5-item EQ-5D, the EuroQol visual analogue scale (EQ-VAS), and the SF-12. Individuals ≥18 years were included in the analysis. Respondents who reported no problems on all of the EQ-5D dimensions were categorized by SF-12 physical component (PCS) and SF-12 mental component scores (MCS). Those who scored at the mean or higher (better health) and those below the mean (worse health) on PCS or MCS were compared on socio-demographic characteristics and self-reported medical conditions. Data were analyzed using SAS 8.2 and STATA 8. An alpha of 0.05 was required for significance. RESULTS: A total of 11,248 individuals were included in the study and 5,104 (45.38%) reported no problems on all of the EQ-5D dimensions. However, such individuals reported diabetes, asthma, high blood pressure, coronary heart disease, angina, myocardial infarction, stroke, emphysema, or joint pain. There was no significant difference in the number of chronic medical conditions among respondents categorized in better health or worse health based on MCS. However, respondents classified in better health based on PCS were less likely to report presence of one or more medical conditions. They also were more likely to be male, younger, more educated, and employed. The mean PCS score among individuals reporting no problems on EQ-5D was significantly lower for respondents reporting any of the above listed medical conditions except asthma and emphysema than those not reporting any medical condition. The mean EQ-VAS score also was significantly lower for respondents reporting any of the above medical conditions except asthma and stroke than those with no medical conditions. CONCLUSIONS: The EQ-5D appears to have a ceiling effect. The one-item EQ-VAS was more sensitive than the five-item EQ-5D in detecting respondents with medical conditions.

PMD17

USING AN ARTIFICIAL NEURAL NETWORK TO PREDICT UTILITY SCORES FROM SF-36 DATA

McEwan P1, Kind P2, Dixon S3, Currie CJ4
1Cardiff University, Cardiff, Wales, United Kingdom; 2Outcomes Research Group, York, United Kingdom; 3Sheffield University, Sheffield, South Yorkshire, United Kingdom; 4University of Wales College of Medicine, Cardiff, Wales, United Kingdom

OBJECTIVES: Preliminary studies have generated utility scores from SF-36 data using linear regression models, providing variable results. Of particular importance is the ability of these models to overcome floor effects. The objective of this study was to determine if an improvement in the accuracy of this modeling could be achieved using neural networks. METHODS: Data on 12,268 subjects were abstracted from the Health Outcomes Data Repository (HODaR) in Cardiff, UK, and split into training, validation and test sets. A single layer, feed-forward neural network was constructed and trained using data from 6268 respondents. A validation set containing data from 3000 respondents was used to find the optimal network structure. For comparative purposes, a linear regression model was then fitted to the same data, and both the regression model and neural network were evaluated using the independent test set data containing 3000 respondents. RESULTS: The following results related to the minimum and maximum utility scores, lower and upper quartiles, and median and mean values, respectively. Actual results from the survey were as follows: −0.48, 1.0, 0.58, 0.88, 0.72, and 0.67, respectively. The “trained” neural network gave the following values −0.38, 1.0, 0.53, 0.87, 0.72, and 0.68, respectively; the multiple regression model gave 0.01, 1.0, 0.47, 0.85, 0.69, and 0.67, respectively. Correlations between actual and predicted utility were 0.82 for the neural network and 0.80 for the regression model. CONCLUSION: In early analysis, the modest improvement in correlation between actual and predicted utility scores obtained via the neural network was primarily due to this class of model being more reliable in mapping lower utility values (<0.5). The bimodal distribution of EQ5D data (UK scoring algorithm) complicates standard regression modelling, favouring the use of these more flexible, data driven neural networks. It is hypothesised that these latter methods are more appropriate.

PMD18

REAL-LIFE PATIENT REPORTED OUTCOMES: DATA MINING OF CONSUMER SPONTANEOUS REPORTS FOR TWO STATINS

Reynolds MW, Ross S, Fahrbach K, Frame D, James K
Metaworks, Inc, Medford, MA, USA

OBJECTIVE: The objective of this study was to examine how consumer spontaneous Medwatch reports can be used to
compare patient-reported outcomes in two statins (Statin A and Statin B). Medwatch is the Food and Drug Administration’s (FDA) program for reporting serious reactions with drugs and other medical products. Consumer reports are an increasing number of those drug safety reports and through March 30, 2003, make up more than 27% of the total Medwatch reports.

METHODS: The QSCANTM software was used to compare the consumer spontaneous reports of two statins (Statin A and Statin B), using a signal detection methodology, the proportional reporting ratio (PRR). The two statins were also individually compared to the class of statin drugs. RESULTS: Statin A had 32,371 consumer reports of adverse events, and Statin B had a total of 21,895 consumer reports through the first quarter of 2003. Report dates were comparable for the two statins. Compared to Statin B, Statin A had a significantly higher proportion of reports for arterial occlusion, cardiac arrest, increased blood cholesterol, coronary artery surgery, transient ischemic attack, and aggravated condition. Statin B had a significantly higher PRR for burning sensation (within nervous system disorders), diabetes mellitus, and decreased blood pressure. The significantly elevated PRRs for Statin B persisted when compared to all other statins, whereas Statin A appeared similar to the rest of the statin class for these particular adverse event reports. CONCLUSIONS: Using drug safety signal detection data mining technology of consumer-driven spontaneous reports, it is possible to identify possible areas of drug differentiation between two drugs or a single drug and an entire drug class. Consumer Medwatch reports are spontaneous drug adverse event reports that are voluntarily submitted by the consumer and may be an informative source of real-life patient reported outcomes.

PMD19
TRAINING REGULATORY AGENCIES IN PRO EVALUATION: THE WORKMATS EXPERIENCE
Acquadrò C1, Lobo-Luppi L1, Chassany O2
1Mapi Research Institute, Lyon, France; 2Hôpital Saint-Louis, Paris, France

A survey of European Regulators conducted by the ERIQA Group evidenced the need of training on HRQL and PRO. To meet this need ERIQA in collaboration with the Cochrane Group evidenced the need of training on HRQL and PRO. To succeed possible using this FOCUS methodology for meeting the training needs of the regulators said that this training might change their way of evaluating files and 90% might recommend the session to colleagues encouraged us to plan sessions in 2004 in other European countries and the US.

PMD20
PRO INSTRUMENT DEVELOPMENT FOR RARE DISORDERS: A FOCUS-ED APPROACH
Gold KF1, Tran KT1, Stephens JM1, Kimura A2
1Abt Associates Inc, Bethesda, MD, USA; 2Transkaryotic Therapies Inc, Cambridge, MA, USA

OBJECTIVES: In the absence of a disease-specific instrument, existing generic profiles may not be sensitive enough to measure health-related quality of life in certain disease populations. In addition, instruments developed for other diseases may have some face validity but are not ideal. The goal of this research is to develop an approach for rare disease instrument development given the challenges of available sample and resources.

METHODS: The functional outcomes for clinical understanding scale (FOCUS)” focuses” on aspects of rare diseases not captured in validated instruments. A list of outcome domains is first generated for the rare disorder. A validated instrument (“anchor”) for a disease with some symptoms common to the rare disease of interest is then identified. New items are developed in a similar scale to the anchor instrument in order to allow joint scoring. Data are collected to assess the validity and reliability for 1) the anchor instrument; 2) new items (FOCUS) and 3) the anchor instrument conjoined with the new items. RESULTS: A FOCUS instrument was developed for Mucopolysaccharidosis II (MPSII; Hunters Syndrome (HS)), a rare genetic disorder in children. The childhood health assessment questionnaire (CHAQ) was selected as the “anchor” instrument. Items of the HS-FOCUS were generated for MPSII indicators not covered by the CHAQ. Both the CHAQ and the HS-FOCUS were tested in MPSII patients and parents. Five of eight FOCUS domains had a Cronbach’s alpha of at least 0.84 individually. The reliability of the overall CHAQ score in MPSII patients was comparable to its original sample (high 0.80s). The reliability of the CHAQ in conjunction with the HS-FOCUS was superior to each in isolation. The compound instrument also showed high face, external and statistical validity.

CONCLUSIONS: This example illustrates the potential success possible using this FOCUS methodology for meeting the instrumentation needs for rare diseases.

METHODS

PMD21
A DESCRIPTIVE ANALYSIS TO INVESTIGATE THE DIFFERENCES BETWEEN TRADERS AND NON-TRADERS IN TIME TRADE OFF
Miller LAN, Mody RR
West Virginia University, Morgantown, WV, USA

OBJECTIVES: The purpose of this study was to examine differences between traders and non-traders and characteristics that are associated with being a non-trader. When eliciting utilities using the Time Trade-Off (TTO) method, studies have shown that we can usually anticipate a high level (15–60%) of non-traders—those people who are not willing to trade any of their life for improved health. METHODS: A cross-sectional supervised self-administered survey was used to assess the perception of health-related quality of life (HRQOL) and utilities using the TTO method in the general population of 3 groups: Caucasian