dard valuation method. The aim of this study is to review and critique temporary health state valuation methods and identify areas for future research. METHODS: We reviewed the literature and evaluated preference-based temporary health state valuation methods according to five criteria: (1) Consistency with quality-adjusted life year theory; (2) Ease of use; (3) Relevance to temporary health state-specific domains; (4) Sensitivity to health state duration; and (5) Extent of bias. Our goal was to provide a critical assessment of methods that could be used to obtain values for use in cost-utility analyses. RESULTS: We identified six temporary health state valuation methods. Methods modified standard approaches by prorating utilities, using a chained approach, or trading-off waiting time or sleep instead of death in a time trade-off. These modifications capture the effect of duration better than standard methods. The strength of methods varied. No method was well tested for validity and reliability with respect to temporary health states. CONCLUSION: The literature on temporary health state valuation methods is sparse and inadequate. Our critique did not identify a method that is appropriate for valuation of all temporary health states. Selection of the most appropriate method should depend on the duration of and type of temporary health state being considered. Further research should focus on the validity, reliability and feasibility of valuation under different circumstances. Utility values obtained using temporary health state methods should be compared to those using standard methods to quantify biases.

**Abstracts**

**PMC31**

**CONTROLLING MEASUREMENT ERROR OF PATIENT-REPORTED-OUTCOMES DURING THE IMPLEMENTATION STAGE OF CLINICAL TRIALS**

Gnanasakthy A

Novartis Pharmaceuticals, East Hanover, NJ, USA

**OBJECTIVE:** Measurement errors may be introduced in the development, cultural adaptation, implementation, and analysis of PRO assessments. Recent publications provide guidance to minimize measurement errors during the development and cultural adaptation stages. Very little guidance is available to control errors, especially in multinational studies, introduced during the implementation of PRO assessments. The objective of this abstract is to highlight errors that may be introduced during the implementation stage, specifically during the production of data capture modules (e.g. Case Report Forms) for multi-national studies.

**METHODS:** A rigorous process was put in place to monitor errors introduced during the CRF development process with the aim of having a library of PRO instruments readily available for use in clinical trials. After typesetting, CRF pages were proof read by three independent reviewers including a native speaker. Suspected errors including poor grammar and typographical errors found in original PRO instruments were reconciled with author’s permission and documented.

**RESULTS:** A total of 40 PRO instruments were used in 39 Phase III multinational studies involving 69 languages in 2006–2007. Three instruments had multiple versions for the same language and the author of another instrument did not have a list of available translations. Two types of errors were found at the final stage of proof reading by native speakers. The first, ambiguous or outdated terminology. The second, typesetting errors which may have altered the meaning of the phrase or question. **CONCLUSION:** An adequate process must be in place to monitor, document and minimize errors that may be introduced during the implementation stage of PRO assessments. Failure to do so, especially in multi-national studies, may invalidate the resources spent during the development and translation stages and increase the Company Risk.

**PMC32**

**PREDICTING SF-6D PREFERENCE-BASED UTILITIES USING MEAN SF-36 HEALTH DIMENSION SCORES WHEN PATIENT LEVEL DATA ARE NOT AVAILABLE**

Ara R1, Brazier JE2

1University of Sheffield, Sheffield, South Yorkshire, UK, 2The University of Sheffield, Sheffield, South Yorkshire, UK

**OBJECTIVES:** The objective of the study is to derive an algorithm to predict a cohort preference-based SF-6D index using the eight mean health dimension scores when patient level data is not available. **METHODS:** Health related quality of life data (n = 6890) collected from patients with a wide range of health conditions was used to explore the relationship between the SF-6D and the eight dimension scores. Ordinary least square regressions were derived using the eight dimension scores and first order interactions. Models were assessed for goodness of fit and predictive abilities using standard statistics such as variance explained; residuals and the proportion of predicted values within the minimal important difference. The models were also compared on their abilities to predict mean cohort SF-6D scores using mean dimension scores using both within-sample and out-of-sample published datasets. **RESULTS:** The OLS equations obtained explained over 83% of the variance in the individual SF-6D scores. While the models over-predict the lower health states and under-predict the higher SF-6D scores on the individual level, the mean absolute errors are in the region of 0.040. When using mean dimension scores from within-sample subgroups and out-of-sample published datasets, the majority of predicted scores were well within the minimal important difference (0.041) for the SF-6D. The models are reasonably accurate at predicting incremental values between study arms (mean error 0.012; mean absolute error 0.017) and when predicting incremental changes over time (mean error 0.004; mean absolute error 0.024). **CONCLUSION:** This paper presents a mechanism to estimate a mean cohort preference-based SF-6D score from published mean dimension scores. This study is unique in that it uses published mean statistics to validate the results. The out-of-sample validation demonstrates the algorithms can be used to inform both clinical and economic research. Further research is required in different health conditions.

**PMC33**

**PREDICTING A MEAN EQ-5D PREFERENCE-BASED SCORE FROM THE 8 MEAN SF-36 DIMENSION SCORES WHEN INDIVIDUAL DATA IS NOT AVAILABLE**

Ara R1, Brazier JE2

1University of Sheffield, Sheffield, South Yorkshire, UK, 2The University of Sheffield, Sheffield, South Yorkshire, UK

**OBJECTIVES:** The objective of the study is to derive a method to predict a cohort EQ-5D preference-based index score using published statistics of the eight dimension scores describing the SF-36 health profile. **METHODS:** Ordinary least square regressions are used to obtain models from patient level data covering a wide range of health conditions. The eight dimension scores, the squares age and gender are used to derive a relationship with the EQ-5D index. Models obtained are compared for goodness of fit using standard techniques such as descriptive statistics, variance explained, the residuals and the proportion of values within the minimal important difference. Predictive abilities are also compared when using summary statistics from both within-sample subgroups and datasets published studies. **RESULTS:** The models obtained explain more than 56% of the variance in the EQ-5D scores. For the individual predicted values, the mean predicted EQ-5D score is correct to two decimal places and the mean absolute error is approximately 0.13. Using summary statistics to