

The TSM was reliable ($\alpha > 0.70$) for all scales. In addition, patients who experienced post-menopausal bleeding were less satisfied than those who did not experience post-menopausal bleeding, indicating the preliminary validity of the TSM. **CONCLUSION:** The TSM is currently being used in clinical trials and further research will be conducted on the predictive validity of the questionnaire in predicting dropouts from clinical trials. Although the TSM was developed with women, the questions were worded in a generic and gender-neutral manner. Thus, it is intended that the TSM be used and validated in a variety of trials and conditions.

GS3

THE VALUATION OF DISEASE-SPECIFIC HEALTH STATES TO FACILITATE ECONOMIC EVALUATION OF PHARMACEUTICALS

Kok E, Stolk E, van Busschbach J

Institute for Medical Technology Assessment (iMTA), Erasmus University Rotterdam, Rotterdam, The Netherlands

OBJECTIVE: Disease-specific questionnaires have often been used to reflect the gradation of specific disorders. However, disease-specific questionnaires do not allow appraisal of clinical outcomes within economic evaluations. Here, a solution could be to determine retrospective quality of life-weights for health states of disease specific questionnaires that previously have been used in trials. The main problem will be to reduce the large number of health states. We applied this method in the study of the quality of life effects of Benign Prostatic Hyperplasia (BPH). **METHODS:** Outcomes of BPH are measured using the IPSS, which health states can be converted into quality of life values using time trade-off. For that, a factor analysis and a decrease of answer levels are needed to reduce the number of health states to be valued. **RESULTS:** After factor analysis two main factors remain, namely “obstructive” ($\alpha = 0.8018$) and “irritative” ($\alpha = 0.7165$). Each of these factors identified by factor analysis was described at three levels in order to restrict the number of possible health states. The resulting 9 health states were valued by the general public. The quality of life values range from 0.8330 for the worst BPH-state to 0.9599 for the best BPH-state. The domain irritative appeared to be experienced as more “severe” than the domain obstructive. **CONCLUSION:** The study shows that by using a factor analysis the large number of health states can be reduced, which offers the possibility to convert disease-specific outcomes into utilities. The study also shows that the two different domains of BPH (obstructive and irritative) have a different impact on the quality of life values. Because the IPSS is now validated for the use in economic appraisal, new and already published research can be used to determine the cost per QALY of different interventions in BPH.

RESPIRATORY DISORDERS

RS1

COMPARISON OF IMPUTATION METHODS FOR MISSING ASTHMA QUALITY OF LIFE DATA

Runken MC¹, Bramley TJ², Eaddy MT², Schulz RM²

¹Glaxo Wellcome, Research Triangle Park, NC, USA;

²University of South Carolina, Columbia, SC, USA

OBJECTIVE: To explore the differences among missing data imputation methods when comparing treatment options. **METHODS:** Data from a randomized clinical trial evaluating the effects of treatment on asthma quality of life (QOL) were obtained. The instrument employed to measure QOL was the Asthma Quality of Life Questionnaire. Six common statistical techniques for handling missing data were evaluated: Complete case analysis (CCA), Markov modeling, hot deck, general imputation (GI), last value carried forward (LVCF), and regression. Scores in the tenth week were modeled using ANCOVA with the baseline score as the covariate. **RESULTS:** A total of 303 subjects were available for analysis, however only 134 subjects had data available for CCA. Two of the three treatment groups had 100 subjects while the remaining treatment group had 103 subjects. Overall LVCF had the lowest average score (5.37, $\sigma = 1.03$), indicating lower quality of life, whereas CCA produced the highest average score (5.83, $\sigma = 0.84$). The effect of treatment group was found to be statistically insignificant across all six methods; however, the LVCF method ($P = 0.1667$) and the GI method ($P = 0.1227$) were substantial closer to achieving statistical significance than other methods. **CONCLUSIONS:** Differing imputation methods did not affect the decision of the analysis. CCA produced the highest average imputation scores while LVCF produced the lowest average imputation scores.

Table 1. AQLQ Scores at the Tenth Week Utilizing Different Imputation Methods

Imputation Method	Treatment 1	Treatment 2	Treatment 3	Overall	Significance of Treatment Group
CCA (n = 134)	5.84 (0.91)	5.68 (0.93)	5.95 (0.63)	5.83 (0.84)	0.228
Markov Model	5.71 (0.92)	5.82 (0.87)	5.85 (0.82)	5.80 (0.87)	0.367
Hot Deck	5.74 (0.91)	5.82 (0.91)	5.88 (0.71)	5.82 (0.84)	0.449
GI	5.62 (1.00)	5.53 (0.98)	5.76 (0.74)	5.64 (0.92)	0.123
LVCF	5.42 (1.06)	5.23 (1.10)	5.47 (0.93)	5.37 (1.03)	0.168
Regression	5.69 (0.83)	5.76 (0.77)	5.82 (0.63)	5.76 (0.74)	0.362