spread use of the EQ-5D a number of country-specific scoring systems have been developed. The objective of this study was to identify and compare all existing EQ-5D valuation studies and country-specific scoring systems. **METHODS:** An electronic search of MEDLINE, EMBASE, NHS EED, HEED and a search through the past proceedings of the Euroqol group up to September of 2010 was conducted to identify all EQ-5D preference elicitation studies. The review included a summary and comparison of study design, model estimation, study demographics and scoring function. **RESULTS:** After screening 2940 citations identified from the literature search, 33 elicitation studies that contained a unique scoring system were included for final review. The key areas of divergence between the studies include: differences in methodology used to directly value health (i.e. SG, TTO, etc.), the number of health states that were directly valued, the transformation of the directly valued health states, the statistical methods used to derive the scoring system, and the model variables included in the scoring system. **CONCLUSIONS:** Differences in methods do exist between population studies. Knowing the extent, at which the identified methodological differences can explain the variation, will help determine whether a global preference for health exists.

PRM30
RESULTS FROM A NEW VISUAL ANALOGUE SCALE PROTOCOL FOR EQ-5D VALUATIONS

Bailey H[1], La Foucade A[1], Kind P[2]
[1]The University of the West Indies, St. Augustine, Trinidad and Tobago, [2]University of York, York, UK

OBJECTIVES: The MVH protocol uses a valuation subset of 43 EQ-5D states. Smaller valuation subsets would allow smaller samples if all respondents can value one set of states. This study tests the performance of valuation subsets based on orthogonal experiment designs. VAS values obtained for EQ-5D states are affected by the identity of adjacent states in the elicitation. This study also tests a VAS protocol developed to minimize such effects. **METHODS:** 230 respondents were each randomly assigned to one of two orthogonal EQ-5D valuation subsets. EQ-5D states were printed on 9cm x 4cm cards with rhomboid edges. Respondents first ranked a shuffled deck of 23 cards. A 1 meter VAS was then placed next to the ranked cards. Respondents transferred the cards on to the VAS with the rhomboid edges pointing to the respective values. They could see all of the cards in place at the same time. Ties were permitted and respondents were free to change the order of states in moving the cards from the ranked order to the VAS. **RESULTS:** OLS regression of VAS Scores produced internally valid models for both valuation subsets with r^2 of 0.54 and 0.56. Within sample comparisons of model versus observed VAS values produced residuals of MAD < 3% with individual values crossing 6 VAS points in only 2 cases. Rank correlation coefficients between model and observed states =0.996. Comparisons across the two subsets/subgroups produced residuals of MAD=4 VAS points and rank correlation coefficients of 0.98 and 0.96. **CONCLUSIONS:** This study suggests that orthogonal experiment designs may be suitable as the basis for the construction of EQ-5D valuation subsets, which can allow for smaller respondent samples. The statistical properties of the models arising in this study suggest that some concerns with VAS methods for EQ-5D elicitations may be reduced using this protocol.

PRM32
VALIDITY EVIDENCE AND VALIDATION PRACTICE IN PAPERS PUBLISHED IN VALUE IN HEALTH (1998-2010): A SYSTEMATIC REVIEW

Chan EKH[1], Darmawanti I[2], Mulyana OP[2], Zumbo BD[1]
[1]University of British Columbia, Vancouver, BC, Canada, [2]State University of Surabaya, Surabaya, East Java, Indonesia

OBJECTIVES: Patient-reported outcome (PRO), health-related quality of life (HRQOL), and related instruments are increasingly used in outcomes research. In developing and evaluating PRO, HRQOL, and related instruments, measurement validity is a fundamental consideration. With an eye toward investigating the methodology of validity and uses of validation, we systematically reviewed the reporting of validation practice in papers published in Value in Health, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) official journal. **METHODS:** A systematic search using the journal official website was conducted in December 2010 – including papers since its inception (January 1998) to December 2010. Two searches were independently conducted by two of the authors. Keywords used in the search included "development OR measurement OR psychometric OR psychometrics OR valid OR validation OR validity." We searched both the titles (126 citations) and abstracts (347 citations). Papers reporting empirical results on validity were included. Econometric, opinion, and conference papers were excluded. A coding sheet was developed and each article was double-coded independently by at least two of the authors. **RESULTS:** Ninety-three articles met the inclusion criteria. The percentage of reports of the broad categories of validity evidence were: construct (39.8%), discriminant (28.0%), convergent (26.9%), content (17.2%), concurrent (14.0%), predictive (5.4%), response processes (4.3%), and consequences (1.1%). A paper may report more than one type of validity evidence. In addition, 95.7% of the papers referred to the validity of the instrument, rather than the modern view of validity of the inferences from the scores. **CONCLUSIONS:** Findings reveal that validity information is not routinely presented from a modern validity perspective; that is, consider a unitary view of validity, cite recent theoretical papers in validity theory, or distinguish between validity (as a property) and validation (as a process/method). Some sources of validity evidence (e.g., consequences and response processes) are essentially ignored.

PRM33
TAKE TIME TO TRAIN: EVIDENCE TO SHOW THAT PATIENT TRAINING IN ELECTRONIC PATIENT REPORTED OUTCOMES IS BENEFICIAL

Ross J[1], Ross E[1], O'Gorman HJ[2], Peck R[1]
[1]Almac Clinical Technologies, Souderton, PA, USA, [2]Almac Group, Craigavon, UK

OBJECTIVES: To establish the concept that it is best practice to provide training to patients on how to utilize electronic patient reported outcome (ePRO) systems before patients begins study participation. **METHODS:** Methods on how to approach training patients on ePRO systems will be presented and recommendations will be discussed. To provide supporting evidence and information on best practices used in ePRO, a literature review was conducted to identify articles reporting on studies that utilized ePRO. Of that, articles were identified that reported providing training to their patients on the ePRO portion of the study prior to participation. Rates of ePRO compliance are examined and approaches to training patients are identified within the articles. **RESULTS:** Out of the 115 articles identified using ePRO, 55 (47.8%) reported providing ePRO training to their patients. ePRO compliance rates of studies with patient training were overall high (Mean=85.08, Standard Deviation=9.83). Training methods reported included: training session, providing demonstration, giving instruction, question/answer session, subject hands-on practice, written instructions/reference material, and testing of mastery of ePRO tool. Duration of training reported ranged from 1 minute to 4 hours. **CONCLUSIONS:** Best practice for training patients is to have a combination of the methods reported above. Having available a dummy system for training is most important so that demonstration can be provided, patients can practice using the system and patients can see exactly what they need to do during their actual study participation. Questions can be addressed and patients can begin the study with a clearer understanding of what to do during participation. Training patients allows for higher comfort levels in use of system, thus alleviating frustration and burden. Furthermore, when patients are trained, they require less support throughout the duration of the study. Overall, training patients will lead to better experiences for patient participation.

PRM34
EVALUATING THE SCREENING ABILITY OF PATIENT-REPORTED OUTCOME INSTRUMENTS

Coon C[1], McLeod L[1], Arnold L[2], Chandran A[3], Martin S[4]
[1]RTI Health Solutions, Research Triangle Park, NC, USA, [2]University of Cincinnati College of Medicine, Cincinnati, OH, USA, [3]Pfizer, Inc., New York, NY, USA, [4]RTI Health Solutions, Ann Arbor, MI, USA

OBJECTIVES: Assessments composed of patient-reported outcome (PRO) measures can be used in health care settings as screeners for various conditions. The objective of developing and using a PRO screening measure may be to quickly identify patients who are likely to benefit from a formal diagnostic evaluation. Alternatively, the development objective may be to avoid unnecessary diagnostic procedures, particularly when these are time or resource-intensive or invasive in nature. The PRO screener may also be administered to simply rule out the existence of a particular condition. The evaluation of a PRO screening assessment ideally occurs through analyses using a "gold standard" diagnosis of the condition of interest. **METHODS:** A number of existing statistical and psychometric methods may be used in such an evaluation, including sensitivity and specificity, positive and negative predictive value, kappa, accuracy, odds ratio, and likelihood ratio. **RESULTS:** The evaluation method selected depends on the objective of the screener itself. If the formal diagnostic procedures are particularly invasive or time- or resource-intensive, then screeners should minimize false positives; in contrast, diseases with exceptional risks when left undiagnosed call for screeners that minimize false negatives. In this research, we explore these methods using data from a study comparing various fibromyalgia screening instruments with the currently accepted gold standard diagnosis for fibromyalgia, namely the American College of Rheumatology 1990 diagnostic criteria (Wolfe et al., 1990). **CONCLUSIONS:** Using the example application, we illustrate the pros and cons of a battery of statistical methods and how they can be used to select the "best" candidate screener.

PRM35
DEVELOPMENT OF A CONCEPT LIST TO ENSURE COMPARABLE CONTENT VALIDITY BETWEEN ORIGINAL PRO QUESTIONNAIRES AND THEIR TRANSLATIONS: A REVIEW OF 15 YEARS OF LINGUISTIC VALIDATIONS

Mear I[1], Conway K[2]
[1]MAPI Institute, Lyon, France, [2]MAPI Research Trust, Lyon, France

OBJECTIVES: In its guidance on the use of PRO measures, the FDA recommends that sponsors provide evidence that the content validity and other measurement properties are adequately similar across all versions used in a clinical trial. One way to ensure the comparability of content is to ensure that all translations reflect the original item intent. As this presupposes the existence of a formal description of each item concept, it is the objective of this study to investigate if this was the case prior to the guidance publication and to make recommendations on the basis of our findings. **METHODS:** The research is based on the review of PRO translations performed between 1995 and 2009. **RESULTS:** A total of 640 questionnaires (27 generic and 613 disease- or condition-specific) were translated. All were developed before the publication of the FDA's PRO guidance. None had a formal written document listing a definition of each item concept and possible translation alternatives available before launching the translations. In all cases, the item-by-item concept list and the list of translation alternatives were developed in collaboration with the developers of the original questionnaires during the translation process. In some cases (multiple projects involving different languages for the same questionnaire), the development of the concept list was a dynamic process fed by questions raised at each new translation [e.g. Quality of Life in Inflammatory Bowel