and easy to complete. The measure will be further tested in additional qualitative and quantitative studies to evaluate its measurement properties.

**PRM174**

**COMPARING EQ-5D VALUATION STUDIES: A SYSTEMATIC REVIEW AND METHODOLOGICAL REPORTING CHECKLIST**

Roy F1, Guebel K2, Doble BM2, Pullenayagum E2

1McMaster University, Hamilton, ON, Canada, 2Monash University, Clayton, Australia

**OBJECTIVES:** There has been a growing interest in developing country-specific scoring algorithms for the EQ-5D around the world. This study systematically reviewed existing EQ-5D valuation studies to highlight their strengths and limitations, explores heterogeneity in observed utilities using meta-regression, and proposes a methodological checklist, the Checklist for Valuation Studies of the EQ-5D (CREATE) for reporting EQ-5D valuation studies. METHODS: A comprehensive literature search was conducted in Medline, EMBASE, the National Health Service Economic Evaluation Database (NH EED) via Wiley’s Cochrane Library, and Wiley’s Health Economic Evaluation Database from inception through November 2012, as well as bibliographies of key papers and the EuroQol General Meeting Proceedings from 2001 to 2012 for English-language reports of EQ-5D valuation studies. Two reviewers independently screened the titles and abstracts for relevance. Three reviewers performed data extraction and compared the characteristics and scoring algorithms developed in the included valuation studies. RESULTS: Of the 31 studies included in the review, 19 used the time-tradeoff (TTO) technique, 10 used the visual analogue scale (VAS) technique, and 2 used both TTO and VAS. More than 70% of studies used respondents from the general population selected by random or quota sampling, and face-to-face interviews or postal surveys. Studies valued between 7 and 198 total states, with 1 to 23 states valued per respondent. Different model fit indices and validation techniques have been used, which vary by demographic factors, including gender, education, per cent urban population, and national health care expenditure, with differences in observed utilities for moderate or severe health states. Conclusions: EQ-5D values in studies published in the last 10 years have varied widely in their design and in the resulting scoring algorithms. Therefore, we propose the CREATE for those conducting valuation studies.

**PRM175**

**CLINICAL OUTCOME ASSESSMENTS AND RATER TRAINING MATERIALS - MAXIMIZING CONSISTENCY**

Nastri A1, A2

1TransPerfect, New York, NY, USA, 2TransPerfect, Research Triangle Park, NC, USA

**OBJECTIVES:** Consistency between translated clinical outcome assessments and corresponding documentation is critical for global clinical trials. If the translation of rater training materials is completed independently or without reference to the linguistically validated clinical outcomes assessments, inconsistencies between the scale and the rater training materials are created. Steps must be taken to ensure consistency between rater training materials and their corresponding scales. METHODS: A review of prior projects involving rater training materials and corresponding scales, as well as discussion with Sponsors who have conducted translations with other organizations was conducted to provide insight into key areas where reference between the translation of the clinical outcomes assessment and the rater training material proved critical. RESULTS: Reviews of prior projects revealed that early reference to previously translated scales yielded improved linguistic consistencies between overlapping text found in rater training documentation, as well as yielding time efficiencies. If the rater training documentation was translated independently of the scale, inconsistencies existed between the scale and the rater training materials. This required additional reconciliation efforts, adding both timeline and cost to the project scope to rectify the inconsistencies. By centralizing the translation process around the community based report, the inconsistent versioning of multiple patient populations were reviewed along with the process for determining their respective respondent criteria. RESULTS: For each project, the goal of consistent training was to determine patient understanding of concepts in the intended language populations. Respondents were developed to represent the intended study population as closely as possible. Criteria ranged from multiple cancer types, to multiple age groups, as well as care sites and patient mixes. In each case, more than one respondent of each sub-category was required in order to ensure sufficient feedback. Conclusions: Cognitive interviewing of a single patient population typically includes around 5 respondents. With multiple populations, an increased total number of respondents is recommended with this requirement for each sub-category. Additional subject restrictions presents a challenge. While exact representation of multiple patient populations may not be possible, development of a more broadly representative selection process that provides population proves beneficial in ensuring that a questionnaire is fit for purpose.

**PRM176**

**DETERMINING THE CURRENT STATE-OF-PLAY FOR VARIANTS OF THE ‘SHORT FORM’ HEALTH-RELATED QUALITY OF LIFE INSTRUMENT IN SPINAL CORD INJURY: A SYSTEMATIC LITERATURE REVIEW**

Vial D1, Thompson M1, Cockbain C2, Hogan S2, Johnson M1, Bourke A1

1University of Oxford, Oxford, UK, 2PharmaQuest Ltd, Banbury, UK, 3i4 Innovation Ltd., Oxford, Oxfordshire, UK

**OBJECTIVES:** Researchers at Oxford are developing a new self-reported assessment tool called the E-Health Impact Questionnaire (e-HIQ), which will be used to assess the impact of using health-related websites. A translatability assessment was performed on the draft items to establish whether they will be cross-culturally and linguistically valid should the questionnaire be translated in the future. The phrase ‘taking steps’ to ‘taking actions’, retaining the positive tone yet avoiding a literal translation. There were also three questions that, although distinct in English, have overlapping meaning when translated. The questions included the terms ‘solidarity’, ‘have a lot in common’ and ‘identify with people’. Although clear terms in English, many languages only have one way of saying all three terms and therefore a change was advised. Conclusions: This study highlights the importance of including a translatability assessment during the development of a measure as it enables developers to incorporate an international element into its development. In turn, this will enable an easier transition into various translations.

**PRM177**

**LINGUISTIC VALIDATION IN MULTIPLE PATIENT POPULATIONS: COGNITIVE INTERVIEWING APPROACHES**

Dobie BM1, Zarrar K1

1TransPerfect, Atlanta, GA, USA, 2TransPerfect, Research Triangle Park, NC, USA

**OBJECTIVES:** While requirements for identifying cognitive interviewing respondents are not specifically outlined in the FDA Guidance for Industry Patient-Reported Outcome Measures, best practices dictate that respondents should be as closely representative of the study patient population as possible. For a questionnaire intended for a single patient population or a study with a single patient population, this concept is sufficient for determining criteria. On the other hand, some cases lend themselves to the use of multiple patient populations and more varied respondent criteria. Methods: Past translation and validation studies included cognitive interviewing of multiple patient populations were reviewed along with the process for determining their respective respondent criteria. Results: For each project, the goal of cognitive interviewing was to determine patient understanding of concepts in the intended language populations. Respondents were developed to represent the intended study population as closely as possible. Criteria ranged from multiple cancer types, to multiple age groups, as well as care sites and patient mixes. In each case, more than one respondent of each sub-category was required in order to ensure sufficient feedback. Conclusions: Cognitive interviewing of a single patient population typically includes around 5 respondents. With multiple populations, an increased total number of respondents is recommended with this requirement for each sub-category. Additional subject restrictions presents a challenge. While exact representation of multiple patient populations may not be possible, development of a more broadly representative selection process that provides population proves beneficial in ensuring that a questionnaire is fit for purpose.

**PRM178**

**DETERMINING THE CURRENT STATE-OF-PLAY FOR VARIANTS OF THE ‘SHORT FORM’ HEALTH-RELATED QUALITY OF LIFE INSTRUMENT IN SPINAL CORD INJURY: A SYSTEMATIC LITERATURE REVIEW**

Vial D1, Bryan S2, Noonan V3, Whitehurst DGT4

1Maastricht University, Maastricht, The Netherlands, 2University of British Columbia, Vancouver, BC, Canada, 3Rick Hansen Institute, Vancouver, BC, Canada, 4Simon Fraser University, Burnaby, BC, Canada

**OBJECTIVES:** The SF-36 and SF-12 are generic health-related quality of life (HRQoL) instruments used in many areas of research, including spinal cord injury (SCI). Repeated concerns have been raised within the broader disability literature about the appropriateness of such measures for subjects with significant physical impairment, yet continued use is widespread. This systematic review provides a comprehensive descriptive overview of how the SCI literature has used variants of multiple HRQoL instruments. Methods: A systematic database and bibliographic search was conducted to identify SCI-related publications reporting outcome data using a HRQoL instrument related to the 36-item and 12-item short form measures (e.g. SF-36, SF-12, RAND-36, RAND-12, SF-6D). Three descriptive considerations formed the basis of the analysis; 1) the frequency of use across identified variants; 2) the degree to which instruments have been administered as partial measures (i.e. selective use of items or domains); and 3) a summary of data collection methodology. Results: One hundred seventy-four articles met the selection criteria. In total, 9 variants were identified, the SF-36 was the most widely used complete measure (n=76), followed by the SF-12 (n=24), SF-6D (n=4), RAND-12 (n=3), and VR-36 (n=2). Partial use of instruments was common (n=54); some studies identified limitations regarding the Physical Functioning scale and tested modified measures (n=13), such as the SF-36L and the SF-36WW. Data collection methodology was often ambiguous or missing (n=65). Conclusions: The SF-36 is a common HRQoL measure within SCI research, despite oft-cited concerns regarding its measurement properties. Attempts to add, delete or modify items have resulted in a large number of variants, often with minimal support. Developing established measures with conceptual and psychometric properties is appealing because it enables comparisons to be made across clinical specialties. However, the trade-off between comparability and context-specific validity requires further explicit consideration within the SCI research.

**PRM179**

**RAPID IDENTIFICATION AND RECRUITMENT OF PATIENTS FROM THE HEALTH IMPACT MEASUREMENT NETWORK [THIN] PRIMARY CARE PATIENT DATA FOR A HEALTH-RELATED QUALITY OF LIFE (HRQoL) STUDY OF PATIENTS WITH ATRIAL FIBRILLATION (AF) IN THE UNITED KINGDOM**

Vial D1, Thompson M1, Cockbain C2, Hogan S2, Johnson M3, Bourke A1

1Cegedim Strategic Data Medical Research Ltd, London, UK, 2Timberwolf Consulting, London, UK, 3Monash University, Clayton, Australia

**OBJECTIVES:** There has been a growing interest in developing country-specific scoring algorithms for the EQ-5D around the world. This study systematically reviewed existing EQ-5D valuation studies to highlight their strengths and limitations, explores heterogeneity in observed utilities using meta-regression, and proposes a methodological checklist, the Checklist for Valuation Studies of the EQ-5D (CREATE) for reporting EQ-5D valuation studies. Methods: A comprehensive literature search was conducted in Medline, EMBASE, the National Health Service Economic Evaluation Database (NH EED) via Wiley’s Cochrane Library, and Wiley’s Health Economic Evaluation Database from inception through November 2012, as well as bibliographies of key papers and the EuroQol General Meeting Proceedings from 2001 to 2012 for English-language reports of EQ-5D valuation studies. Two reviewers independently screened the titles and abstracts for relevance. Three reviewers performed data extraction and compared the characteristics and scoring algorithms developed in the included valuation studies. Results: Of the 31 studies included in the review, 19 used thetime-tradeoff (TTO) technique, 10 used the visual analogue scale (VAS) technique, and 2 used both TTO and VAS. More than 70% of studies used respondents from the general population selected by random or quota sampling, and face-to-face interviews or postal surveys. Studies valued between 7 and 198 total states, with 1 to 23 states valued per respondent. Different model fit indices and validation techniques have been used, which vary by demographic factors, including gender, education, per cent urban population, and national health care expenditure, with differences in observed utilities for moderate or severe health states. Conclusions: EQ-5D values in studies published in the last 10 years have varied widely in their design and in the resulting scoring algorithms. Therefore, we propose the CREATE for those conducting valuation studies.