THE EFFECT OF RECALL PERIOD ON CANCER PATIENTS’ RATINGS OF THE SEVERITY OF MULTIPLE SYMPTOMS

Shi Q1, Trask PC2, Wang S, Mendola T, Cleeland C

1University of Texas M. D. Anderson Cancer Center, Houston, TX, USA; 2University of New London, CT, USA

In response to the US Food and Drug Administration’s concern on choice of suitable recall period for patient-reported outcomes (PRO), we examined the effects of recall period on PRO ratings by comparing ratings made using 24-hour recall and 7-day recall periods of the MD Anderson Symptom Inventory (MDASI). METHODS: Forty-two patients, two at their 3rd to 8th week of chemotherapy at the Radiation Treatment Center at M.D. Anderson Cancer Center were asked to rate their symptoms using the MDASI on two separate occasions, one week apart. By the initial visit, patients were randomly assigned to rate their symptoms using either a 24-hour recall or a 7-day recall. On their next visit, patients were asked to rate their symptoms using the recall period not used at their first visit. RESULTS: Correlation coefficients of global symptom severity between 24-hour and 7-day recall periods were 0.89. Examining individual items, all correlation coefficients were over 0.7 except for distress (r = 0.67). The percentages of moderate to severe symptoms (5 or greater) were consistent in the 24- and 7-day recall periods, with no significant difference in the prevalence of moderate to severe symptoms being found between the two recall periods. Cronbach’s α coefficients in both 24-hour and 7-day recalls were all over 0.8. Symptoms from both recall periods were more severe for patients with poorer performance status. Among 20 patients who underwent cognitive debriefing, 70% thought the 7-day recall was “more appropriate” for answering the MDASI, but 83% did not think that recall period would influence their answers. CONCLUSIONS: This study demonstrated that a 7-day recall version of the MDASI has psychometric properties consistent with the 24-hour recall version, which may allow its use in future clinical trials. In addition, this study may help ease the choice of recall period when symptoms are outcome measures.

RELATIONSHIP BETWEEN QUALITY OF LIFE AND HEALTH-RELATED MEASURES INCLUDING SYMPTOMS, BIOCHEMICAL MARKERS AND TUMOR BURDEN

Vink JJ, Vink AI, Silva MP

Eastern Virginia Medical School, Norfolk, VA, USA

OBJECTIVES: Examine the relationship of quality of life in neuroendocrine tumor patients using the Norfolk QOL-NET by correlating the total questionnaire score with each of the Norfolk QOL-NET domains, with tumor burden, biochemical status and the Norfolk Carcinoid Symptom Score tool. METHODS: During their visits to the Neuroendocrine Unit at Eastern Virginia Medical School, 29 adult patients diagnosed with neuroendocrine tumor(s) signed the consent form and completed the Norfolk QOL-NET. Data related to current tumor burden, biochemical status and the validated Carcinoid Symptom Score was obtained from their files matching the date they completed the questionnaires. RESULTS: The Norfolk QOL-NET total score correlated positively with all of its domains – 0.8, physical functioning (r = 0.96, p < 0.0001), depression/anxiety (r = -0.74, p < 0.001), gastrointestinal (r = 0.78, p < 0.0011), flushing (r = 0.62, p < 0.0001), respiratory (r = 0.65, p < 0.0002), positive attitude (r = 0.52, p < 0.0004), and cardiovascular (r = 0.46, p < 0.0012); with the Norfolk Carcinoid Symptom Score (r = 0.6, p < 0.0001); with tumor burden (r = 0.004), and serotonin (r = 0.62, p < 0.013). Serotonin was the only biochemical marker that correlated positively with a poor quality of life in patients with neuroendocrine tumors. CONCLUSIONS: We demonstrated a strong correlation between Norfolk QOL-NET and symptoms, biochemical markers and tumor burden. Norfolk QOL-NET seems sensitive to symptom change, physical functioning, respiratory and cardiovascular disease progression or remission. Norfolk QOL-NET should be an important tool for measuring patients’ perception of the burden of their disease, relating to the tumor burden and the biochemical abnormality as well as the impact of treatment modalities. The Norfolk quality of life tool may also be a useful guide in deciding changes in therapy to alter apparent health status as well as an endpoint in clinical studies.

TEST-RETEST RELIABILITY OF THE EQ-SD VISUAL ANALOG SCALE ACROSS POPULATIONS AND CONDITIONS

Wills CT, Pickard AS

1University of Illinois at Chicago, Chicago, IL, USA; 2College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA

OBJECTIVES: As electronic versions of HRQoL measures such as the EQ-SD become available, it is important to understand the reliability of different modes of technology. The aim of this study was to summarize the evidence of test-retest reliability for the EQ-SD visual analog scale (VAS), a scale ranging from 0 (worst imaginable health) to 100 (best imaginable health). METHODS: A structured literature search was conducted in MEDLINE using keywords relevant to EQ-SD, visual analog scales, and test-retest reliability. Original research studies that reported information on the test-retest reliability of the EQ-SD VAS were included. Demographic characteristics, interval between observations, and intra-class correlation coefficients (ICCs) were abstracted. RESULTS: Of the 25 studies that examined test-retest reliability of EQ-SD, 14 reported evidence of test-retest reliability for EQ-SD VAS. Most of the papers were studies that assessed the validity of EQ-SD for certain countries or languages (n = 5/14, 36%) or for use in patient groups / certain medical conditions (n = 7/14, 50%). The most common interval between observations was 2 weeks (n = 4/14, 29%), with analyses conducted on a subgroup of self-reported stable patients, based on self-report, in 4 studies. TRT ICs ranged from ICC = 0.38 for Alzheimer’s patients to ICC = 0.90 (95% CI: 0.88–0.92) for a methodological study conducted in Spain, with a median ICC = 0.8 across the 14 studies. Almost 80% of studies (11/14) reported ICs above 0.7, a reliability threshold considered acceptable at the group level. CONCLUSIONS: EQ-SD VAS demonstrated acceptable TRT reliability in most studies of populations and medical conditions except in Alzheimer’s disease, where proxies but not patients provided reproducible assessments.

DETERMINING THE MINIMALLY IMPORTANT DIFFERENCES OF FOUR PREFERENCE-BASED HEALTH INDICES: A SIMULATION APPROACH

Loo N1, Johnson JA2, Coons SJ3

1National University of Singapore, Singapore, Singapore; 2University of Alberta, Edmonton, AB, Canada; 3University of Arizona, Tucson, AZ, USA

OBJECTIVES: To estimate the minimally important differences (MIDs) for the EQ-5D, HUI2, HUI3, and SF-6D health index scores using health-state transitions described by each instrument’s health classification systems as anchors. METHODS: We estimate that the smallest difference in health states defined by each instrument’s multi-attribute health classification (MAHC) systems are associated with important differences in health preferences. Based on this assumption, the MID was defined as the difference in index score between two health states defined by each MAHC system differing in only one health dimension or attribute and by only one functional level. Thus, for each instrument, we enumerated all the theoretically possible pairs of minimally different health states and calculated the differences in index scores for those pairs of health states. RESULTS: Based on our definitions, the total number of pairs of minimally different health states is 405 for the EQ-5D, 127,600 for the HUI2, 6,382,810 for the HUI3, and 86,700 for the SF-6D. The mean (standard deviation) MID estimate was 0.040 (0.026) for the EQ-5D (US algorithm), 0.082 (0.032) for the SF-6D (UK algorithm), 0.045 (0.039) for the HUI2, 0.032 (0.028) for the HUI3, and 0.027 (0.028) for the SF-6D. The effect sizes corresponding to these MID estimates range from 0.19 to 0.28. In general, these MID estimates are quite comparable to those estimated using other anchor-based methods. CONCLUSIONS: This new approach to estimating the MIDs of four commonly used preference-based HRQoL index scores provides new and useful information for identifying and interpreting meaningful change (or differences) in scores.


Ferrari V, Kind P

1University of York, York, North Yorkshire, UK; 2University of York, York, UK

OBJECTIVES: The development of national health policies requires a clear understanding about how objective and subjective measures of health status vary over time. This task is only possible when generic self-reported instruments are considered part of a population surveys alongside traditional health indicators. This study examines variations in self-reported health status in England as measured by the EuroQol EQ-SD
questionnaire. METHODS: Data from the Health Survey for England (HSE) of 1996 (n = 16,443) and 2006 (n = 14,142) were selected in order to assess variation in population health status over a 10 year period. Both surveys covered population aged 16 years and over living in private households. The sample is regularly drawn using a multistage stratified random design in phases including postcode sectors as the primary sampling units. Given that only the EQ-SD descriptive system is included to describe self-reported health in the HSE, a predicted EQ-SDVAS was estimated for each respondent based on a regression model developed from data of the 1993 York Measurement and Valuation of Health Project. RESULTS: Despite being older (2.59 years on average, p-value <0.01) and having a slightly higher proportions of women (0.8 percent, p-value 0.156), the 2006 HSE reflects that English population has significantly (p-value <0.01) reduced its prevalence of self-reported health problems in the last 10 years in three of the five EQ-5D dimensions: usual activities, pain/discomfort and anxiety/depression. Mobility and self-care dimensions, although higher in prevalence, did not reach statistical significance at 5% level when both years were compared. Health improvements over time were also identified in the self-rated EQ-SDINDEX and predicted EQ-SVAS (p-value <0.001), having the 16−44 age-group and women the highest health gains. CONCLUSIONS: EQ-SD is a useful tool for monitoring population health. Our findings will assist local policymakers and public health authorities by improving their knowledge about trends in self-perceived health.

ASSESSING THE QUALITY OF CONJOINT ANALYSIS APPLICATIONS IN HEALTH: A PILOT EVALUATION OF THE ISPOR CHECKLIST FOR CONTENT VALIDITY AND PRACTICE IN CONJOINT ANALYSIS.

Marshall DA, Hauber AE, Bridges JP, Cameron R, Weaver L, Dionne J, Johnson FK.

University of Calgary, Calgary AB, Canada, RTI Health Solutions, Research Triangle Park, NC, USA, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD, USA.

OBJECTIVES: Increasingly, conjoint analysis, a stated-preference method, is applied in health outcomes research. Variation in method type and quality make it difficult to assess substantive findings. The ISPOR Conjoint Analysis Database Project was established to identify and evaluate conceptual and linguistic equivalence in conjoint applications in the literature using the 10-point ISPOR Checklist for Good Research Practice in Conjoint Analysis (the Checklist). METHODS: Multiple electronic databases published between 1980 and 2008 were searched to identify conjoint-analyses in human health studies. Only English-language publications were incorporated. Included studies were subject to detailed data extraction including descriptive information, methodological details on survey type, experimental design, survey format, attributes and levels, sample size, number of conjoint tasks per respondent, and analysis methods. Review articles and conference abstracts where there is a relative lack of qualitative data in many applications of conjoint analysis were piloted to identify key elements to be included in the database using a standardized taxonomy and test the Checklist as an evaluative framework for the methodological assessment of these studies. RESULTS: The search identified 2,362 citations – 264 met inclusion criteria. The number of applied studies increased substantially over time (1980 – 5 and 2007 – 42) in a broad range of applications, cancer being the most frequent. Based on the pilot results, discrete-choice experiments using fractional factorial designs were most common. Attribute number ranged from 3–6, choice tasks per respondent ranged from 8–16 and sample size ranged from 30–335. Studies generally repeated key estimates more than required by the 10-point Checklist, especially in the first 15 items. CONCLUSIONS: Conjoint analysis in health has expanded to include a broad range of applications and methodological approaches. The Checklist provides a framework to assess their quality. The conjoint analysis Database project will complete the assessment of the quality and variability of these studies based on the pilot findings.

THE TRANSLATION AND LINGUISTIC VALIDATION OF THE NEUROPATHY TOTAL SYMPTOM SCORE-6: SELF-ASSESSED VERSION (NTSS-6 5A)

Portado T, Gordon-Stables R, Wild D

Oxford Translations Ltd, Oxford, UK

The NTSS-6 5A has been translated into many different languages. It is designed to assess the severity of peripheral neuropathy symptoms. The objective of this study was to produce translations that are conceptually equivalent to the original and to language versions, ensuring the relevance of the translations within the target cultures. A standard methodology was employed: 2 forward translations, a reconciliation of the forward translations, 2 back translations, back translation review; or an in-country review; linguistic validation interviews with 5 patients with diabetic peripheral neuropathy in each country, and 2 proofreadings. Numerous cultural and linguistic issues became apparent throughout the translation process, including the following: – Many different pain types are described (e.g. stabbing, shooting, electric-shock like, boring, aching) which were particularly difficult as this vocabulary was unavailable in some languages. A decision was made to assign the pain types into two groups; firstly dull aching pains, and secondly sharper, stabbing pains. These could then be more easily conveyed and translated. – For many countries, there was no direct translation for ‘pins and needles’. If the country had no idiomatic description of this, ‘feeling as if ant bites are present’ was used. – Some items asked about ‘feet’; many of the countries involved have no specific word for ‘feet’, so ‘from ankle to toes’ was translated. – Some languages were unable to convey ‘asleep feeling’ in a limb; this wording was therefore converted to ‘numbness’. The NTSS-6 5A has been translated and linguistically validated using a rigorous translation process. A number of cultural and linguistic issues became apparent and were resolved. The measure is now appropriate for use in multilingual trials.

A NOVEL COMPARISON OF QUALITATIVE DATA SOURCES: CONTENT ANALYSIS OF SEMI-PERSONAL STRUCTURED INTERVIEWS VERSUS WEBLOGS (BLOGS)

Assael S1, Wild D1


OBJECTIVES: Blogs have become a fruitful source of qualitative data in recent years as participants post their thoughts and experiences. In this study, the potential use of this new data source could provide valuable insight in the early stages of research development. As such, the objective of this study was to assess the potential use of blogs in research development by comparing the data available in blogs with that gained from conducting semi-structured interviews with patients. METHODS: The subject of menopausal hot flushes was used to demonstrate the comparison. Twenty semi – structured interviews were conducted with women reporting to suffer hot flashes. The interviews focused on a description of the symptoms and their impact on HRQoL. The themes emerging from the content analysis of these interviews was then compared to the themes found in twenty blog entries. Four researchers conducted the analysis, two in each data source group. RESULTS: Both the semi – structured interview data and the blog data provided numerous descriptions of the symptoms of hot flushes, with no discrepancies in thematic content. The interviews did however allow for explicit discussion of the relationship between hot flashes and night sweats, which could only be inferred in the blog analysis. Similarly, the effect of hot flashes on physical and social functioning, and psychological wellbeing, produced similar themes in both data sources. However, while the interviews permitted clarification of the impact of symptoms on HRQoL, blogs analysis often relied on inference. CONCLUSIONS: The broad themes elicted from both data sources were comparable. However, the interactive nature of the interviews produced richer, more reliable data than that contained within the blogs. As such the role of blog analysis could be that of a cost effective adjunct to literature searches when developing research protocols.

USE OF THE MINI INTERNATIONAL NEUROPSYCHIATRIC INTERVIEW (M.I.N.I.) – VERSION 6 – IN AN INTERNATIONAL STUDY

Soubrie P, D’Arcey P, Sheehan D, Lecrubier Y, Even C

Map Research Institute, Lyon, France, Map Research Institute, LYON, France, University of South Florida, Tampa, FL, USA, Centre Hospitalier Sainte-Anne, Paris, France.

The M.I.N.I. is a semi-structured interview designed to explore and diagnose psychiatric disorders in research and clinical settings. Various versions of the US English instrument exist (M.I.N.I. Kid, M.I.N.I. Plus, etc.). Since its development in 1990 some or all versions have been translated into more than 40 languages. OBJECTIVES: Before using version 6 in 12 countries, it had to be translated according to a rigorous methodology to meet 3 requirements: (1) concordance with existing translations, (2) conceptual equivalence across and (3) linguistic consistency within language groups. METHODS: The translation process was conducted as follows: forward translations on the basis of translations of version 5 produced by psychiatrists appointed by the authors, backward translation, clinician review and whenever possible, review by the psychiatrists who had coordinated the initial translations. The authors contributed to the process by identifying the original concepts and reviewing the backward