The Ox-PAQ Initiative: Current Literature, Rationale & Approach

Report prepared for the 2nd meeting of the Ox-PAQ Steering Committee, 26th March 2013, London.

David Morley
Sarah Dummett
Laura Kelly
Jill Dawson
Mathew Humphrey
Ray Fitzpatrick
Crispin Jenkinson

Health Services Research Unit
Department of Public Health
University of Oxford

Copies of this report can be obtained from:
David Morley
Senior Project Lead
Health Services Research Unit
Tel: +44 (0)1865 289402
e-mail: david.morley@dph.ox.ac.uk

Contents

Executive Summary

The Oxford Participation and Activities Questionnaire (Ox-PAQ) initiative is a three year project with the aim of developing and validating a measure of participation and activity for use with patients experiencing a range of health conditions.

Current Literature

- There is a significant and continuing debate in the academic literature regarding the definition of 'participation' and 'activity'.
- The World Health Organisation International Classification of Functioning, Disability and Health has provided a framework within which to construct scales aimed at measuring the two constructs.
- The first stage of the Ox-PAQ initiative, to review existing measures of participation and activity, has been completed.

Rationale for the Ox-PAQ Initiative

- Evidence from the review indicates that current measures of participation and activity are largely disability and rehabilitation focused, with none exhibiting the full range of psychometric properties to a satisfactory standard.
- The development of a number of instruments is poorly reported and some have been developed with small sample sizes and without the involvement of patients in generating items.
- The Ox-PAQ initiative will address these flaws by developing a measure for generic use, with patients at the heart of the item-generation process. The measure will be assessed for the full range of psychometric properties and sample sizes will allow for detailed and legitimate psychometric analyses.

Approach

- Interviews are currently being conducted with professionals from a range of backgrounds, whose experience may inform the development of the Ox-PAQ.
- Items comprising the Ox-PAQ will subsequently be generated through interviews with patients representing a range of conditions. Recruitment of participants will be via relevant healthcare charities.
- Pre-testing of the instrument will be achieved through cognitive interviews and focus groups.
- A pilot-test survey will be conducted with data subject to factor and Rasch analysis in order to identify appropriate dimensions. Reliability will be assessed using Cronbach's Alpha, item-total correlations and intraclass correlations.
- A large scale field test will subsequently be undertaken with the Ox-PAQ being administered in conjunction with generic measures of health status (EQ-5D and SF-36) to further test validity of the measure.
- The Ox-PAQ will be re-administered at 2 weeks to assess test-retest reliability and at 3 months to assess responsiveness.

List of Abbreviations

APQ6: Activity & Participation Questionnaire

EQ-5D: EuroQol Group Questionnaire

FDA: United States Food & Drug Administration

ICF: International Classification of Functioning

IMPACT-S: ICF Measure of Activity & Participation – Screener

IPA: Impact on Participation & Autonomy Questionnaire

KAP: Keele Assessment of Participation

LIFE-H: Assessment of Life Habits Questionnaire

Ox-PAQ: Oxford Participation and Activities Questionnaire

P-Scale: Participation Scale

PAR-PRO: Participation Profile

PARTS/M: Participation Survey / Mobility

PM-PAC: Participation Measure for Post-Acute Care

POPS: Participation Objective, Participation Subjective

ROPP: Rating of Perceived Participation

SF-36: Short-Form 36 Health Survey

USER: Utrecht Scale for Evaluation of Clinical Rehabilitation

USER-P: Utrecht Scale for Evaluation of Rehabilitation-Participation

WHO: World Health Organisation (WHO)

Background

The Oxford Participation and Activities Questionnaire (Ox-PAQ) initiative is a three year project with the aim of developing and validating a measure of participation and activity for use with patients experiencing a range of health conditions. The need for such a measure is driven by population ageing and increasing demands on health and social care services. There is a growing interest in the management of long term conditions and maximising the cost effectiveness of treatment, in part by keeping people active and participating in the community. Consequently, it is essential that a well-developed and validated instrument that can meaningfully assess levels of participation and activity is widely available.

This report focuses on three particular aspects of the Ox-PAQ initiative. First, a review of current literature and measures relating to participation and activity is presented. Secondly, on the basis of the literature review, the rationale for the project is outlined. Finally, the approach to be taken in the development of the Ox-PAQ is outlined.

Current Literature

There is little agreement in the academic literature as to how we define 'participation' and 'activity'. A plethora of papers have been published with the aim of clarifying or reaching a consensus on a clear definition of both terms (e.g. Dijkers, 2010; Heinemann et al, 2010; Whiteneck et al, 2011). They do not, however, appear to have entirely achieved their goal and the debate continues. Some have called for greater clarity between the two concepts (Badley, 2008), whilst others have concluded that participation is a complex, elusive and subjective concept, influenced by one's individual perspective (Magasi et al, 2009). Herein lies the importance of investigating the meaning of concepts such as participation with those in whom we are trying to measure it. When this is undertaken, the complex nature of what participation and activity actually mean to different individuals becomes evident (Hammel et al, 2008).

Some level of agreement is reached in the World Health Organisation (WHO) International Classification of Functioning, Disability and Health (ICF; World Health Organisation, 2001), which has at least provided a framework within which to construct scales aimed at measuring the two constructs. The ICF defines participation as 'involvement in life situations' and activity as 'the execution of a task or action by an individual'. Although initially described as two separate concepts, the final version of the ICF merges activity and participation into a single taxonomy and identifies nine domains which are presented in Table 1.

1. Learning & Applying Knowledge	6. Domestic Life
2. General Tasks & Demands	7. Interpersonal Interactions & Relationships
3. Communication	8. Major Life Areas
4. Mobility	9. Community, Social & Civic Life
5. Self-care	

Table 1: Domains of the ICF classification of participation and activity

There is still significant debate over the ICF model and particularly in relation to whether activity and participation is a single entity, or whether a clear distinction should be made between the two (van der Zee et al, 2011; Whiteneck & Dijkers, 2009). Despite this continuing debate, a number of measures have now been constructed, to varying degrees, around the model.

There follows a brief review of participation and activity measures that have been developed and validated over the last 15 years. The development process, measurement characteristics and psychometric properties of each are presented in Table 2, supplemented by a short critique of each instrument.

Activity & Participation Questionnaire (APQ6; Stewart et al, 2009): The APQ6 is a six item measure with items developed from 'concepts of the Australian Bureau of Statistics (ABS) Labour Force Surveys and Census'. The developers report that the time reference and other characteristics of the questionnaire were based on previous ABS measures, but do not substantiate this further. The only psychometric analysis conducted on the APQ6 to date is assessment of test-retest reliability in a sample of 129 mental health patients. Feedback was also received from 79 participants as to the feasibility of the questionnaire, respondents being largely positive. The APQ6 clearly requires further assessment of its psychometric properties as acknowledged by the authors. Their claim, however, that 'the reported psychometric properties support the proposed use of the APQ6' is, at best, dubious.

Assessment of Life Habits (LIFE-H; Fougeyrollas et al, 1998): The LIFE-H was constructed around the Disability Creation Process model (Fougeyrollas, 1999) and originated as a 248 item (long-form) and 58 item (short-form) measure. It is not entirely clear how items were generated and the initial validation study is fundamentally flawed due to the unacceptable sample size of 49. The measures, however, have undergone significant development in recent years, although this is not well documented in the literature. There are currently 3 versions of the LIFE-H; a 16 item short-form, a 77 item 'general' version, and a 242 item long-form measure. These questionnaires have been validated in a range of age groups including older adults (Noreau et al, 2004; Poulin & Desrosiers, 2009). In their quest to ensure measurement is possible across the entire lifespan the developers have also constructed a 71 item version of the LIFE-H for infants from birth – 4 years (Lepage et al) and for children aged 5-13 (Noreau et al, 2007) in two formats; a short-form of 64 items and a long-form of 198 items. The LIFE-H suite of measures has been extensively used in Canada by the team who developed them, but not widely adopted by others. Their psychometric properties are not fully assessed and, as has been documented by others (Sakzewski et al 2007; Morris et al, 2005), the child measures in particular require further validation.

ICF Measure of Activity & Participation – Screener (IMPACT-S; Post et al, 2008): The IMPACT-S is a 32 item measure that was developed using the ICF model. Items appear to have been generated by the development team, although this is far from clear in the validation paper. A small group of patients and experts was involved in the early phases of development to assess the instrument's relevance, completeness, and acceptability. Psychometric testing based on a sample size of 275 established acceptable levels of validity and reliability, although factor analyses provided mixed results. Interestingly, the developers of IMPACT-S are one of very

few to report a response rate in their validation study; at 33% this was very low and may have biased their results. An additional study has further assessed the reliability of the measure (van der Zee et al, 2010), and an assessment of the responsiveness of the scale concludes that the measure is sufficient for evaluation studies, but questionable at the individual level (van der Zee et al, 2011).

Impact on Participation & Autonomy (IPA; Cardol et al, 1999): The IPA is a 23 item measure that was again developed using the ICF model. Items were generated by a 'multidisciplinary research group' with no patient input until the pilot phase. The subsequent validation study was conducted with a small sample of 100 participants. Results showed acceptable levels of reliability and validity, but this should be viewed with caution given the sample size. There is some evidence of responsiveness (van der Zee et al, 2013; Cardol 1999), but this requires further assessment, as noted by others (Wilkie, et al, 2011; Magasi & Post 2010). Despite its apparent limitations the IPA has been widely used, particularly in the Netherlands where it was developed.

Keele Assessment of Participation (KAP; Wilkie et al, 2005): The KAP is an 11 item measure developed for generic use in population studies. Items were generated by the developers based on the ICF model. The initial validation study incorporated an impressive sample size of 1117 and established good levels of validity and test-retest reliability. However, internal consistency and responsiveness have yet to be established. Additionally, as the authors acknowledge, the brevity of the KAP means that it 'may miss some specific details and distinctiveness of participation restriction in individuals' (Wilkie et al, 2005).

Participation Objective, Participation Subjective (POPS; Brown et al, 2004): The POPS is a 26 item scale that was developed around the ICF framework with items drawn from a variety of previously developed scales. The psychometric assessment of the instrument is at the 'lower end' of what would generally be expected, with only test-retest data being reported. The authors justify their lack of psychometric rigour through a lengthy discussion as to why traditional and modern psychometric approaches are inappropriate for the POPS. Unsurprisingly, the instrument has had little use and other reviews have documented its limitations (i.e. Wilkie et al, 2011; Noonan et al, 2010a;b).

Participation Profile (PAR-PRO; Ostir et al, 2006): The PAR-PRO is a 20 item measure that was developed around the ICF model with items generated through a process of literature review and expert panel. The measure was subsequently administered to a large sample of 594 rehabilitation inpatients. Psychometric analysis showed the instrument to have sound validity and internal consistency. However, test-retest reliability and responsiveness are yet to be assessed and this may explain why the PAR-PRO does not appear to have been incorporated into any studies to date.

Participation Survey / Mobility (PARTS/M; Gray et al, 2006): The PARTS/M is based on the ICF model, with items generated via key-informant interviews and focus groups. The instrument was administered to an impressive sample of 604 participants with a range of mobility limitations. Psychometric analyses show the PARTS/M to have good reliability and validity, but the responsiveness of the measure has not yet been reported. The major limitation of the PARTS/M is its

length. At 135 items over a 13 page booklet the measure is likely to cause significant respondent burden. This may be a contributory factor in the PARTS/M having received very limited uptake.

Participation Measure for Post-Acute Care (PM-PAC; Gandek et al, 2007): The PM-PAC is a 51 item measure that was constructed around the ICF model. Items of the PM-PAC were generated by the team of developers, with no patient input until the pre-testing phase. The instrument was subsequently administered to a sample of 395 participants from three impairment groups; neurological, musculoskeletal and 'medically complex'. The PM-PAC demonstrates reasonable levels of validity and reliability, with limited evidence of responsiveness. The development paper is an object lesson in how not to report the development of a new scale, and this may have contributed to the lack of uptake for the measure.

Participation Scale (P-Scale; van Brakel, et al, 2006): The P-Scale is an 18 item measure that is interviewer administered. Based on the ICF framework, it was developed by an international team of researchers with people with stigmatised conditions (e.g. leprosy) through a series of interviews and focus groups. The P-Scale demonstrates reasonable levels of reliability, but requires further testing of its validity as well as assessment of responsiveness. Despite this, the instrument has had some uptake in the literature and is available in 6 different languages. A shortened 13 item version is currently being validated (Stevelink et al, 2012).

Rating of Perceived Participation (ROPP; Sandström & Lundin-Olsson, 2007): The ROPP is a 22 item measure that was developed from the domains of the ICF classification with items being generated by a panel of experts. No patient input is evident until a series of interviews at the pre-testing phase. The subsequent psychometric evaluation is based on a sample size of 85, which is less than ideal. Whilst basic reliability and validity evaluation is reported, there is no assessment of responsiveness. As noted by other reviews (Wilkie et al, 2011), the ROPP requires further development if it is to be adopted by other researchers.

Utrecht Scale for Evaluation of Clinical Rehabilitation (USER; Post et al, 2009): The USER is a 30 item, clinician / therapist administered scale. It is unclear if there is any theoretical underpinning and the developers do not report how items were developed; two significant shortcomings. The measure has currently only been tested for inter-rater reliability and concurrent validity, with some evidence of responsiveness.

Utrecht Scale for Evaluation of Rehabilitation-Participation (USER-P; van der Zee et al, 2008; Post et al, 2012): The USER-P is a 31 item instrument, again based on the ICF model. Items were generated from existing scales and the measure was subsequently reviewed by experts. Psychometric analysis revealed sound validity and internal consistency, with further studies establishing test-retest reliability (van der Zee et al, 2010) and responsiveness (van der Zee et al, 2013). Although these psychometric properties are promising the absence of patient input in the item generation process is a significant limitation.

Instrument Name, Developer & Country	Intended Use	Development Process	Response Options	Items	Dimen-sions	Validity	Reliability	Respon-siveness
Activity & Participation Questionnaire (APQ6), Stewart et al, 2009), Australia	Mental illness	• Census data • Validation study n = 129	Not reported	6	0		✓ Test-retest	
Assessment of Life Habits (LIFE-H), Fougeyrollas et al, 1998, Canada	Rehabilitation	• Expert panel • Validation study n = 49	4 / 5	77	2 / 12	✓ Content ✓ Construct ✓ Concurrent ✓ Discriminant	✓ Internal consistency ✓ Test-retest ✓ Inter-rater	
ICF Measure of Activity & participation – Screener (IMPACT-S), Post et al, 2008, Netherlands	Rehabilitation	• Expert panel • Validation study n = 275	4	32	9	✓ Face ✓ Content ✓ Construct ✓ Concurrent	✓ Internal consistency ✓ Test-retest	✓ (limited)
Impact on Participation & Autonomy (IPA), Cardol et al, 1999, Netherlands	Disability	• Expert panel • Validation study n = 100	5 / 3	23	5	✓ Content ✓ Construct ✓ Concurrent ✓ Discriminant	✓ Internal consistency ✓ Test-retest	✓ (limited)
Keele Assessment of Participation (KAP), Wilkie et al, 2005, UK	Generic	• Cognitive interview • Patient interview • Validation study n = 1117	5	11	1	✓ Content ✓ Construct ✓ Concurrent ✓ Discriminant	✓ Test-retest	
Participation Objective, Participation Subjective (POPS) Brown et al, 2004, US	Brain Injury	• Generated from existing measures • Validation study n = 454	5	26	5	✓ Content	✓ Test-retest	
Participation Profile (PAR-PRO), Ostir et al, 2006, US	Disability	• Literature review • Expert panel • Validation study n = 594	5	20	4	✓ Content ✓ Construct ✓ Discriminant	✓ Internal consistency	
Participation Survey / Mobility (PARTS / M), Gray et al, 2006, US	Mobility impairment	• Key informant Interview • Focus groups • Validation study n = 604	Assorted	135	6	✓ Content ✓ Construct ✓ Concurrent	✓ Internal consistency ✓ Test-retest	

Table 2: Characteristics of current measures of activity and participation

Instrument Name, Developer & Country	Intended Use	Development	Response Options	Items	Dimen-sions	Validity	Reliability	Respon-siveness
Participation Measure for Post-Acute Care (PM-PAC), Gandek et al, 2007, US	Generic / rehabilitation	• Literature review • Focus groups • Expert opinion • Validation study n = 395	5	51	9	✓ Content ✓ Construct ✓ Known groups	✓ Internal consistency ✓ Test-retest	✓ (limited)
Participation Scale (P-Scale), van Brakel, et al, 2006, International	Rehabilitation	• Key informant Interview • Focus group • International workshop • Validation study n = 497	5	18	1	✓ Face ✓ Content ✓ Construct	✓ Internal consistency ✓ Inter-rater ✓ Intra-rater	
Rating of Perceived Participation (ROPP) Sandström & Lundin-Olsson 2007, Sweden	Neurological rehabilitation	• Expert panel • Patient interview • Validation study n = 85	5	22	9	✓ Face ✓ Construct	✓ Internal consistency ✓ Test-retest	
Utrecht Scale for Evaluation of Clinical Rehabilitation (USER), Post et al, 2009, Netherlands	Rehabilitation	• Item generation not reported • Validation study n = 319	Assorted	30	3	✓ Concurrent	✓ Inter-rater	✓ (limited)
Utrecht Scale for Evaluation of Rehabilitation-Participation (USER-P), Post et al, 2012, Netherlands	Rehabilitation	• Generated from existing measures • Expert review • Validation study n = 509	5 / 6	31	3	✓ Concurrent ✓ Discriminant ✓ Construct	✓ Internal consistency	✓ (limited)

Table 2 (continued): Characteristics of current measures of activity and participation

Rationale

The previous summary of the literature provides a sound rationale for developing the Ox-PAQ. Current measures of participation and / or activity are largely disability and rehabilitation focused, with none currently exhibiting the full range of psychometric properties that are required under current guidance. The development of a number of instruments is poorly reported and some have been developed with small sample sizes, which calls into question the adequacy of their psychometric characteristics and their generalisability. Many have failed to involve patients in generating items, something that is considered fundamental if measures are to reflect the issues of greatest concern to those we are assessing.

The Ox-PAQ initiative will address these flaws by developing a measure for generic use, with patients at the heart of the item-generation process. The measure will be assessed for the full range of psychometric properties (validity, reliability and responsiveness) and sample sizes will allow for detailed and legitimate psychometric analyses. Details of how this will be achieved are summarised below.

Approach

Development of the Ox-PAQ will proceed through a number of stages as previously incorporated in the development of widely used and highly regarded measures such as the PDQ-39 (Peto et al, 1995; Jenkinson et al, 1997a; Martínez-Martín et al, 2011) and other measures developed and validated by members of the Ox-PAQ team (e.g. Jenkinson et al, 2012; Jenkinson et al, 2008; Jenkinson et al, 2005; Jenkinson et al, 1999; Jenkinson et al, 1997b; Dawson et al, 2006; Dawson et al 1998; Dawson et al 1996a;b; Morley et al, 2013a;b; Morley et al, 2012; Morley et al, 2010; Kelly et al, 2013a;b). Such a development strategy is entirely in line with current guidelines such as those provided by the United States Food & Drug Administration (Food & Drug Administration, 2009). Each stage is listed in Table 3 and detailed further below.

Stage	Activity
1	End-user interviews
2	Patient interviews
3	Item generation and pre-testing
4	Pilot-test survey
5	Psychometric analysis of pilot-test survey
6	Large scale field test
7	Psychometric analysis of field test
8	Test-retest assessment
9	Responsiveness assessment

Table 3: Developmental stages of the Ox-PAQ Initiative

(1) Semi-structured interviews are currently being conducted with a range of professionals with experience of clinical practice, health care regulation and management, purchasing and conducting research. The purpose of these interviews is twofold. Firstly, to discuss issues with current patient-reported outcome measures and, secondly, to consider the broad topic areas a generic measure of participation and activity might include that will be meaningful and considered important. No previous measure has been designed from the outset with the views of such a variety of potential users at its core. Furthermore, the interviews will discuss practicalities of the instrument, such as preferred methods of scoring and rules for interpretation. Interviews will be transcribed and used to inform later stages of the project.

(2) Exploratory, open-ended, in-depth interviews will subsequently be undertaken with a sample of patients across a wide range of conditions, including, but not exclusively, Parkinson's disease, amyotrophic lateral sclerosis, diabetes, musculo-skeletal conditions and cancer. Recruitment of participants will be via relevant healthcare charities. The interviews will aim to identify salient dimensions of illness impact that have adversely affected participation, activities and autonomy. It is intended that the patient groups included will represent a range of conditions that together affect all key bodily systems, as well as having different symptoms, trajectories and prognoses. Approximately 25-30 patients will be interviewed in depth until 'saturation' is reached. Interviews will be recorded and whilst there will be a topic guide, patients will be free to range across these and any other relevant topics.

(3) The interviews conducted in Stage 2 will be transcribed and transcripts will be scrutinised independently by the research team for issues relating to illness impact. These issues will then be re-cast as questionnaire items. A meeting of health care researchers, and relevant interested stakeholders will then be convened to assess the list for completeness, ambiguity or repetition. The resulting candidate questionnaire will then be shown to a small group of people with long term conditions in a focus group setting for comments and to assess its face validity. Additionally the questionnaire will be pre-tested with a variety of patients in cognitive interviews using the format outlined by Willis (1994). Participants will be asked to complete the questionnaire and explain the reasons for their answers to questions on the measure, and comment upon any difficulties or ambiguities they experience. It is intended that the questionnaire will be scored with two algorithms; one based on the capabilities of respondents without assistance from aids or another person, and the second based on what is possible when given assistance (such as mechanical aids or another person).

(4) The long-form questionnaire drawn up in Stage 3 will be mailed to a sample of 600 people with a variety of long term conditions. This sample size is based on the assumption that the instrument will contain approximately 50 items. Estimates suggest that a minimum of five times as many respondents than items are required for psychometric tests to be meaningful (Norman & Streiner, 2000). As with the interviews in Stage 2, participants will be recruited via relevant healthcare charities. Assuming a typical response rate of between 60 and 70% (Asch et al, 1997) this will lead to a sample of approximately 360-420 questionnaires, which will permit rigorous testing of the instrument.

(5) Statistical analysis of data from Stage 4 will allow for identification of a shorter form questionnaire using established statistical procedures. Initial inspection of the data will select out any item with more than 5% missing responses and identify those items that display substantial 'floor' or 'ceiling' effects (i.e. responses fall predominantly at either end of the range of responses). Typically items with 70% or more of responses falling into either category are excluded (Terwee et al, 2007). Data will subsequently be factor analysed and items will be excluded from further analyses if they do not load on any given factor. Items that demonstrate a factor loading of less than 0.32 level are regarded as unacceptable and will be removed (Kline, 1994). The factor analytic procedures incorporated are likely to produce a number of clusters of items each relating to a specific aspect of activity and participation. The items relating to each factor will then be assessed for face validity and any which appear to duplicate others will be deleted. Internal reliability will then be assessed using Cronbach's alpha statistic (Cronbach, 1951). Where the internal reliability of any set of items can be improved by deleting an item or items they will be removed. Any set of items not reaching the minimum acceptable alpha coefficients of 0.7 (Scientific Advisory Committee of the Medical Outcomes Trust, 2002) will not be included in the final version. Item-total correlations, corrected for overlap, will be calculated between items and the total score to which they contribute. Items exhibiting correlations with their own scale score of 0.3 or less will be deleted in line with accepted recommendations (Nunnally & Bernstein, 1994). Consequently, items with high face validity, good correlation with the scale total to which they contribute and good results from the reliability analyses will then constitute a dimension of the resulting instrument. Rasch analysis (Hobart & Cano, 2011; Rasch, 1960) will subsequently be performed to determine which items conform to a hierarchical unidimensional structure.

(6) The short-form measure generated in Stage 5, together with existing generic health status measures, the SF-36 (Ware & Sherbourne, 1992; Ware et al, 2007; Jenkinson et al, 1999) and EQ-5D (EuroQol Group, 1990) will be mailed to 1000 patients with a wide variety of conditions, again recruited through relevant healthcare charities. Assuming a 60-70% response rate (as discussed previously) this would yield a final sample size of approximately 600-700 returned questionnaires.

(7) Confirmatory factor analyses will be undertaken to confirm the results from the pilot-test survey (Stage 5). Minor amendments may be made to the questionnaire at this time. Descriptive statistics, with confidence intervals, of results from the Ox-PAQ, EQ-5D and SF-36 will be calculated. Spearman correlations of dimensions of the SF-36 and EQ-5D measures with domains of the Ox-PAQ instrument will also be calculated. It is hypothesised that domains on the two measures which overlap will be highly correlated.

(8) Previous respondents will be asked to complete a further questionnaire. If they agree to do so they will be sent the Ox-PAQ one week after receipt of the measure from the field test in order to assess test-retest reliability by means of intra-class correlation coefficients.

(9) Previous respondents will be asked to complete a final questionnaire after three months has elapsed in order to assess the instrument's responsiveness via calculation of standardised effect sizes.

Concluding Remarks

This report has aimed to provide a brief review of current literature and measures relating to participation and activity. The review provides a sound rationale for conducting the Ox-PAQ inititative as current measures demonstrate significant limitations. Many have been poorly developed and none currently exhibit the full range of psychometric properties that would be expected. The report concludes with an outline of how the Ox-PAQ project will be conducted. The stages discussed represent best practice in the development of outcome measures, as well as including a new strategy of involving and engaging with potential users from the outset of the development process.

The Ox-PAQ team will continue to report to the steering committee on the progress of the project at regular intervals.

References

Asch DA, Jedrziewski MK, Christakis NA. (1997). Response rates to mail surveys published in medical journals. *Journal of Clinical Epidemiology*; 50(10): 1129-1136.

Badley EM. (2008). Enhancing the conceptual clarity of the activity and participation components of the International Classification of Functioning, Disability, and Health. *Social Science & Medicine*; 66(11): 2335-2345.

Brown M, Dijkers MP, Gordon WA, Ashman T, Charatz H, Cheng Z. (2004). Participation objective, participation subjective: a measure of participation combining outsider and insider perspectives. *Journal of Head Trauma Rehabilitation*; 19(6): 459-481.

Cardol M, de Haan RJ, van den Bos GA, de Jong BA, de Groot IJ. (1999). The development of a handicap assessment questionnaire: the Impact on Participation and Autonomy (IPA). *Clinical Rehabilitation*; 13(5): 411-419.

Cronbach LJ. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(2): 297-334.

Dawson J, Coffey J, Doll H, Lavis G, Cooke P, Herron M, Jenkinson C. (2006). A patient-based questionnaire to assess outcomes of foot surgery: validation in the context of surgery for hallux valgus. *Quality of Life Research*; 15(7): 1211-1222.

Dawson J, Fitzpatrick R, Murray D, Carr A. (1998). Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone & Joint Surgery (Br)*; 80(1): 63-69.

Dawson J, Fitzpatrick R, Carr A, Murray D. (1996a). Questionnaire on the perceptions of patients about total hip replacement. *Journal of Bone & Joint Surgery (Br)*; 78(2): 185-190.

Dawson J, Fitzpatrick R, Murray D, Carr A. (1996b). Questionnaire on the perceptions of patients about shoulder surgery. *Journal of Bone & Joint Surgery (Br)*; 78(4): 593-600.

Dijkers MP. (2010). Issues in the conceptualization and measurement of participation: an overview. *Archives of Physical Medicine & Rehabilitation*; 91(Suppl 1): S5-16.

EuroQol Group. (1990) EuroQol – A new facility for the measurement of health related quality of life. *Health Policy*; 16(3): 199-208.

Food & Drug Administration, Department of Health and Human Sciences. (2009). Guidance to Industry. Patient Reported Outcome Measures. Use in Medical Product Development to Support Labelling Claims. Silver Spring, MD: Food and Drug Administration.

Fougeyrollas P, Noreau L, Bergeron H, Cloutier R, Dion SA, St-Michel G. (1998). Social consequences of long term impairments and disabilities: conceptual approach and assessment of handicap. *International Journal of Rehabilitation Research*; 21(2): 127-141.

Fougeyrollas P, Cloutier R, Bergeron H, Côté J, St Michel G. (1999). The Quebec Classification: Disability Creation Process. International Network on the Disability Creation Process.

Gandek B, Sinclair SJ, Jette AM, Ware JE Jr. (2007). Development and initial psychometric evaluation of the participation measure for post-acute care (PM-PAC). *American Journal of Physical Medicine & Rehabilitation*; 86(1): 57-71.

Gray DB, Hollingsworth HH, Stark SL, Morgan KA. (2006). Participation survey/mobility: psychometric properties of a measure of participation for people with mobility impairments and limitations. *Archives of Physical Medicine & Rehabilitation*; 87(2): 189-197.

Hammel J, Magasi S, Heinemann A, Whiteneck G, Bogner J, Rodriguez E. (2008). What does participation mean? An insider perspective from people with disabilities. *Disability & Rehabilitation*; 30(19): 1445-1460.

Heinemann AW, Tulsky D, Dijkers M, Brown M, Magasi S, Gordon W, DeMark H. (2010). Issues in participation measurement in research and clinical applications. *Archives of Physical Medicine & Rehabilitation*; 91(Suppl 1): S72-76.

Hobart J, Cano S. (2011). Rasch Analysis, in Jenkinson C, Peters M, Bromberg M. (Eds.) Quality of Life Measurement in Neurodegenerative & Related Conditions. Cambridge: Cambridge University Press.

Jenkinson C, Dummett S, Kelly L, Peters M, Dawson J, Morley D, Fitzpatrick R. (2012). The development and validation of a quality of life measure for the carers of people with Parkinson's disease (the PDQ-Carer). *Parkinsonism & Related Disorders*; 18(5): 483-487.

Jenkinson C, Kennedy S, Jones G. (2008). Evaluation of the American version of the 30-item Endometriosis Health Profile (EHP-30). *Quality of Life Research*; 17(9): 1147-1152.

Jenkinson C, Burton JS, Cartwright J, Magee H, Hall I, Alcock C, Burge S. (2005). Patient attitudes to clinical trials: development of a questionnaire and results from asthma and cancer patients. *Health Expectations*; 8(3): 244-252.

Jenkinson C, Fitzpatrick R, Brennan C, Bromberg M, Swash M. (1999). Development and validation of a short measure of health status for individuals with amyotrophic lateral sclerosis / motor neurone disease: the ALSAQ-40. *Journal of Neurology*; 246(Suppl 3): 1116-1121.

Jenkinson C, Stewart-Brown S, Petersen S, Paice C. (1999). Assessment of the SF-36 Version 2 in the United Kingdom. *Journal of Epidemiology & Community Health*; 53(1): 46-50.

Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. (1997a). The Parkinson's Disease Questionnaire (PDQ-39): development and validation of a Parkinson's disease summary index score. *Age & Ageing*; 26(5): 353-357.

Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. (1997b) The PDQ-8: Development and validation of a short-form Parkinson's disease questionnaire. *Psychological Health*; 12: 805-814.

Kelly L, Jenkinson C, Ziebland S. (2013a). Measuring the effects of online health information for patients: item generation for an e-health impact questionnaire. *Patient Education & Counseling*; In Press.

Kelly L, Spencer S, Barrett G. (2013b). Using intermittent self-catheters: experiences of people with neurological damage to their spinal cord. *Disability & Rehabilitation*; In Press.

Kline P. (1994). An Easy Guide to Factor Analysis. London: Routledge.

Lepage C, Noreau, L, Fougeyrollas, P. Development, validity and reliability of a measure of participation in young children: The LIFE-H for Children 0-4. Conference Abstract / Poster.

Magasi S, Post MW. (2010). A comparative review of contemporary participation measures' psychometric properties and content coverage. *Archives of Physical Medicine & Rehabilitation*; 91(9 Suppl): S17-28.

Magasi S, Hammel J, Heinemann A, Whiteneck G, Bogner J. (2009). Participation: a comparative analysis of multiple rehabilitation stakeholders' perspectives. *Journal of Rehabilitation Medicine*; 41(11): 936-944.

Martinez-Martin P, Jeukens-Visser M, Lyons KE, Rodriguez-Blazquez C, Selai C, Siderowf A, Welsh M, Poewe W, Rascol O, Sampaio C, Stebbins GT, Goetz CG, Schrag A. (2011). Health-related quality-of-life scales in Parkinson's disease: Critique & Recommendations. *Movement Disorders*; 26(13): 2371-2380.

Morley D, Jenkinson C, Doll H, Lavis G, Sharp R, Cooke P, Dawson J. (2013a). The Manchester Oxford Foot Questionnaire (MOXFQ): Development & Validation of a Summary Index Score. *Bone & Joint Research*; In Press.

Morley D, Dummett S, Kelly L, Peters M, Dawson J, Fitzpatrick R, Jenkinson C. (2013b). The PDQ-Carer: Development and validation of a summary index score. *Parkinsonism & Related Disorders*; 19(4): 448-449.

Morley D, Selai C, Thompson A. (2012). The self-report Barthel Index: preliminary validation in people with Parkinson's disease. *European Journal of Neurology*; 19(6): 927-929.

Morley D, Selai C, Schrag A, Thompson A, Jahanshahi M. (2010). Refinement and validation of the Parental Illness Impact Scale. *Parkinsonism & Related Disorders*; 16(3): 181-185.

Morris C, Kurinczuk JJ, Fitzpatrick R. (2005). Child or family assessed measures of activity performance and participation for children with cerebral palsy: a structured review. *Child: Care, Health & Development*; 31(4): 397-407.

Noonan VK, Kopec JA, Noreau L, Singer J, Mâsse LC, Zhang H, Dvorak MF. (2010a). Comparing the validity of five participation instruments in persons with spinal conditions. *Journal of Rehabilitation Medicine*; 42(8): 724-734.

Noonan VK, Kopec JA, Noreau L, Singer J, Mâsse LC, Dvorak MF. (2010b). Comparing the reliability of five participation instruments in persons with spinal conditions. *Journal of Rehabilitation Medicine*; 42(8): 735-743.

Noreau L, Desrosiers J, Robichaud L, Fougeyrollas P, Rochette A, Viscogliosi C. (2004). Measuring social participation: reliability of the LIFE-H in older adults with disabilities. *Disability & Rehabilitation*; 26(6): 346-352.

Noreau L, Lepage C, Boissiere L, Picard R, Fougeyrollas P, Mathieu J, Desmarais G, Nadeau L. (2007). Measuring participation in children with disabilities using the Assessment of Life Habits. *Developmental Medicine & Child Neurology*; 49(9): 666-671.

Norman GR, Streiner DL. (2000). Principal components and factor analysis. In GR Norman & DL Streiner (eds) *Biostatistics: The Bare Essentials*. Hamilton: B.C. Decker Inc.

Nunnally JC, Bernstein IH. (1994). Psychometric Theory (Third Edition). McGraw-Hill Series in Psychology, McGraw-Hill, Inc.

Ostir GV, Granger CV, Black T, Roberts P, Burgos L, Martinkewiz P, Ottenbacher KJ. (2006). Preliminary results for the PAR-PRO: a measure of home and community participation. *Archives of Physical Medicine & Rehabilitation*; 87(8): 1043-1051.

Peto V, Jenkinson C, Fitzpatrick R, Greenhall R. (1995). The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. *Quality of Life Research;* 4(3): 241-248.

Post MW, de Witte LP, Reichrath E, Verdonschot MM, Wijlhuizen GJ, Perenboom RJ. (2008). Development and validation of IMPACT-S, an ICF-based questionnaire to measure activities and participation. *Journal of Rehabilitation Medicine*; 40(8): 620-627.

Post MW, van de Port IG, Kap B, Berdenis van Berlekom SH. (2009). Development and validation of the Utrecht Scale for Evaluation of Clinical Rehabilitation (USER). *Clinical Rehabilitation*; 23(10): 909-917.

Post MW, van der Zee CH, Hennink J, Schafrat CG, Visser-Meily JM, van Berlekom SB. (2012). Validity of the Utrecht scale for evaluation of rehabilitation-participation. *Disability & Rehabilitation*; 34(6): 478-485.

Poulin V, Desrosiers J. (2009). Reliability of the LIFE-H satisfaction scale and relationship between participation and satisfaction of older adults with disabilities. *Disability & Rehabilitation*; 31(16): 1311-1317.

Rasch G. (1960). Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press.

Sakzewski L, Boyd R, Ziviani J. (2007). Clinimetric properties of participation measures for 5- to 13-year-old children with cerebral palsy: a systematic review. *Developmental Medicine & Child Neurology*; 49(3): 232-240.

Sandström M, Lundin-Olsson L. (2007). Development and evaluation of a new questionnaire for rating perceived participation. *Clinical Rehabilitation*; 21(9): 833-845.

Scientific Advisory Committee of the Medical Outcomes Trust (2002). Assessing health status and quality of life instruments: Attributes and review criteria. *Quality of Life Research* 11(3), 193-205.

Stevelink SA, Hoekstra T, Nardi SM, van der Zee CH, Banstola N, Premkumar R, Nicholls PG, van Brakel WH. (2012). Development and structural validation of a shortened version of the Participation Scale. *Disability & Rehabilitation*; 34(19): 1596-1607

Stewart G, Sara G, Harris M, Waghorn G, Hall A, Sivarajasingam S, Gladman B, Mowry B. (2010). A brief measure of vocational activity and community participation: development and reliability of the Activity and Participation Questionnaire. *Australian & New Zealand Journal of Psychiatry*; 44(3): 258-266.

Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology;* 60(1): 34–42.

van Brakel WH, Anderson AM, Mutatkar RK, Bakirtzief Z, Nicholls PG, Raju MS, Das-Pattanayak RK. (2006). The Participation Scale: measuring a key concept in public health. *Disability & Rehabilitation*; 28(4): 193-203.

van der Zee CH, Baars-Elsinga A, Visser-Meily JM, Post MW. (2013). Responsiveness of two participation measures in an outpatient rehabilitation setting. *Scandinavian Journal of Occupational Therapy*, In Press.

van der Zee CH, Kap A, Rambaran Mishre R, Schouten EJ, Post MW. (2011). Responsiveness of four participation measures to changes during and after outpatient rehabilitation. *Journal of Rehabilitation Medicine*; 43(11): 1003-1009.

van der Zee CH, Priesterbach AR, van der Dussen L, Kap A, Schepers VP, Visser-Meily JM, Post MW. (2010). Reproducibility of three self-report participation measures: The ICF Measure of Participation and Activities Screener, the Participation Scale, and the Utrecht Scale for Evaluation of Rehabilitation-Participation. *Journal of Rehabilitation Medicine*; 42(8): 752-757.

van der Zee CH, Post MW, Kap A, Schouten EJ, van Royen R, Rambaran Mishre R, Berdenis van Berlekom SH. (2008). Measurement of participation as an outcome of outpatient rehabilitation: a prospective multi-centre study. *Revalidata*; 146: 12-16.

Ware JE Jr, Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. (2007). User's Manual for the SF-36 Version 2, Second Edition. Lincoln RI: Qualitymetric.

Ware JE Jr, Sherbourne CD. (1992). The MOS 36-item short-form health survey (SF-36) I: Conceptual framework and item selection. *Medical Care*, 30, 473-483.

Whiteneck GG, Bogner JA, Heinemann AW. (2011). Advancing the measurement of participation. *Archives of Physical Medicine & Rehabilitation*; 92(4): 540-541.

Whiteneck CG, Dijkers MP. (2009). Difficult to measure constructs: conceptual and methodological issues concerning participation and environmental factors. *Archives of Physical Medicine & Rehabilitation*; 90(11 Suppl): S22-35.

Wilkie R, Jordan JL, Muller S, Nicholls E, Healey EL, van der Windt DA. (2011). Measures of social function and participation in musculoskeletal populations: Impact on Participation and Autonomy (IPA), Keele Assessment of Participation (KAP), Participation Measure for Post-Acute Care (PM-PAC),Participation Objective, Participation Subjective (POPS), Rating of Perceived Participation (ROPP), and The Participation Scale. *Arthritis Care & Research*; 63(S11): 325-336.

Wilkie R, Peat G, Thomas E, Hooper H, Croft PR. (2005). The Keele Assessment of Participation: a new instrument to measure participation restriction in population studies. Combined qualitative and quantitative examination of its psychometric properties. *Quality of Life research*; 14(8): 1889-1899.

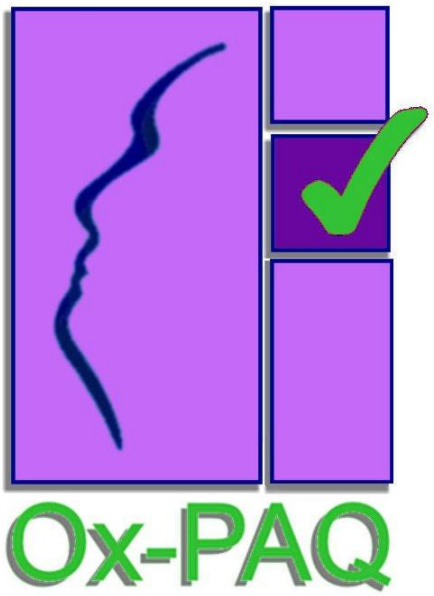
Willis G. (1994). Cognitive interviewing and questionnaire design: A training manual. *Cognitive Methods Staff working Paper Series, No.7*, Hyattsville, MD: Office of Research and Methodology, National Center for Health Statistics.

World Health Organisation. (2001). *International Classification of Functioning, Disability & Health.* Geneva: World Health Organisation.

The Ox-PAQ Steering Committee

Asim Humayun	Genus
Corinne Savill	Novartis
Deanna Dubose	Abbott
Diana Rofail	Roche
Frank De Felice	MSD
Helen Smith	GSK
Kristina Dziekan/Luisa Alvares	Medtronic
Lars Joensson	Bayer
Mary Uhlenhopp	Amgen
Neil Mulcock	Lundbeck
Tara Symonds	Pfizer
Vanessa Stevens	Vifor
Veronica Guijarro	Britannia

Alistair Newton	European Brain Council
Crispin Jenkinson	University of Oxford
David Morley	University of Oxford
Jill Dawson	University of Oxford
Ken Paterson	NHS Scotland
Laura Kelly	University of Oxford
Mary Baker	European Brain Council
Mathew Humphrey	University of Nottingham
Ray Chaudhuri	Kings College
Ray Fitzpatrick	University of Oxford
Sarah Dummett	University of Oxford
Teus van Laar	University of Groningen
Thomas Lonngren	Strategic Advisor


Ox-PAQ