Parkinsonism and Related Disorders 21 (2015) 1312–1316

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

journal homepage: www.elsevier.com/locate/parkreldis

Comparing results from long and short form versions of the Parkinson's disease questionnaire in a longitudinal study

Crispin Jenkinson [a, *], Carl Clarke [b], Richard Gray [c], Paul Hewitson [a], Natalie Ives [d], David Morley [a], Caroline Rick [d], Keith Wheatley [e], Adrian Williams [f]

[a] Health Services Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK
[b] Department of Neurology, City Hospital, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK
[c] Clinical Trial Service Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK
[d] Birmingham Clinical Trials Unit, University of Birmingham, Birmingham, UK
[e] Cancer Research UK Clinical Trials Unit, School of Cancer Studies, University of Birmingham, UK
[f] Department of Neurology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

ARTICLE INFO

Article history:
Received 18 April 2015
Received in revised form
20 August 2015
Accepted 1 September 2015

Keywords:
Parkinson's disease questionnaire
PDQ-39
PDQ-8
PDQ single index
Patient reported outcomes

ABSTRACT

Background: The purpose of this study was to determine the extent to which summary index scores from the short form Parkinson's Disease Questionnaire (PDQ-8) replicate those from the parent form (PDQ-39) in a longitudinal study.

Methods: Longitudinal data gained from the PD-MED trial were examined (n = 1867), to determine the extent the PDQ-8 replicates results from the PDQ-39 at baseline and follow up. The sensitivity to change of the PDQ-8 was also compared with that of the PDQ-39. Finally, results on the two measures were compared with those from the Hoehn and Yahr (HY) clinical staging scale.

Results: Results of the Single Index summary score gained from the PDQ-8 were found to closely replicate those gained from the PDQ-39 at each of the three time points. Furthermore at each time point the intraclass correlation coefficient between the two measures was very high (ICC range 0.93–0.96). Similarly, the two measures gave very similar accounts of change (e.g. from baseline to follow up at one year effect sizes were 0.18 for the single index calculated using the PDQ-39, and 0.09 when calculated using the PDQ-8). Similar levels of correlation were found between the two indices when correlated with the HY scale.

Conclusions: The PDQ-8 closely replicates results gained from the PDQ-39 when calculating single indices. In instances where a single summary score of the impact of PD on self-reported quality of life is needed, it is likely the PDQ-8 will provide reliable and accurate information.

1. Introduction

Patient reported outcome measures (PROs) play an increasingly important role in the evaluation of medical care [1] and have been advocated as potentially important end-points in clinical trials [2]. Traditionally, neurologists have chosen to develop rating scales for Parkinson's disease (PD) based on clinical assessment [3] and which classically focus on neurological symptoms and physical impairment [4–7]. However, typically such instruments fail to address the full impact of the illness upon subjectively assessed quality of life (QoL) of patients [8]. Consequently, a number of PD specific PROs have been developed [9] to capture the overall impact of PD on health-related quality of life with the most widely used and validated being the Parkinson's Disease Questionnaire (PDQ-39) [10–12]. Use of the instrument has been recommended in a number of critical reviews of competing PROs in PD [13–15].

The PDQ-39 is a 39 item self-report questionnaire which measures eight dimensions of health. The instrument was developed on the basis of interviews with people with Parkinson's (PwP) and consequently measures areas of concern which are of particular salience to this patient group. Furthermore, scores from the eight dimensions can be aggregated onto the same metric to provide a single index of the overall impact of PD on self-reported health

* Corresponding author. Health Services Research Unit, Nuffield Department of Population Health, Rosemary Rue Building, University of Oxford, Old Road Campus, Headington, Oxford, OX3 7LF, UK.
E-mail address: crispin.jenkinson@dph.ox.ac.uk (C. Jenkinson).

http://dx.doi.org/10.1016/j.parkreldis.2015.09.008
1353-8020/

status [16]. Such summary indices are useful in giving a global score of the impact of disease, and are useful in trials by reducing the risk of chance findings due to multiple comparisons across dimensions [17].

Further research developed a shorter form PDQ which can be used to create the single index. The PDQ-8 [18] was developed by selecting the item from each dimension most highly correlated with the corrected dimension total. The resulting PDQ-8 summary index (PDQ-8-SI) has been shown to produce, in cross sectional and test-retest studies, results that are encouragingly similar to the PDQ-39 summary index (PDQ-39-SI) [19]. However, to date limited information has been available concerning the sensitivity to change of the PDQ-8 in relation to the PDQ-39 over time. This is an important issue when selecting and using instruments in evaluative studies [2]. Consequently, the aim of this study was to compare data generated from the PDQ-8-SI and the PDQ-39-SI over time, in a longitudinal study.

2. Methods

Data reported here are from PD-MED, a randomised clinical trial evaluating the comparative clinical and cost-effectiveness of different classes of drugs in PwP. Patients were categorised as receiving treatment for 'either' early or 'late' PD. Those classified as 'late' were those whose symptoms were no longer controlled by their first class of treatment. The primary outcome measure for the trial was health-related QoL as measured by the PDQ-39. In this paper data is not broken down by treatment arm but is broken down by 'early'/'late' categories. Full details of the trial design and results are published elsewhere [20].

The trial was awarded Multi-Centre Research Ethics Committee (MREC) approval and Clinical Trials Authorisation from the Medicines and Healthcare Regulatory Authority (MHRA). All respondents gave informed written consent to participate in the study.

2.1. Participants

PwP from over 80 neurology and care of the elderly units across the United Kingdom participated in clinic-based tests and postal evaluations via questionnaires. PD MED recruited 2120 patients – 1620 early and 500 later patients. 1366 (64.4%) of PwP in PD MED were male and 754 (36.6%) female. In this study only participants who had complete data to enable calculation on the PDQ-8-SI and PDQ-39-SI are included in the analyses. Consequently, 1434 (88.51%) PD MED early and 433 (86.6%) late respondents are reported. The mean age at recruitment into the study was 70.46 years (range 27–94), and mean disease duration was 11.22 years (range 4.9–38.6 years).

2.2. Materials

Three validated measures form the basis of the analyses reported here:

- The PDQ-39 [11]: As previously introduced, a 39 item self-report questionnaire which measures eight dimensions of health, namely mobility, activities of daily living, emotional well-being, stigma, social support, cognition, communication, and bodily discomfort. Dimension scores are coded on a scale of 0 (perfect health as assessed by the measure) to 100 (worst health as assessed by the measure). A number of studies indicate that the instrument possesses sound levels of reliability, validity and responsiveness [10–12,21–23]. The PDQ-39-SI is calculated by summing the eight dimensions of the instrument and standardising the score on a scale of 0–100.
- The PDQ-8 [18]: As previously introduced, an 8 item self-report questionnaire derived from its parent questionnaire, the PDQ-39 [11]. The PDQ-8 has been shown to exhibit appropriate levels of reliability, validity and responsiveness [12,18,19]. The PDQ-8-SI is calculated by summing the eight items of the instrument and standardising the score on a scale of 0–100. It should be stressed that the PDQ-8 was not administered as a separate instrument in this study. Rather, it was calculated from PDQ-39 data which may have influenced the manner in which items were completed. It is, however, standard practice to assess short form instruments in this way [18,24–26]. All PDQ data was collected by paper and pen completion via postal surveys.
- The modified Hoehn and Yahr (HY) staging scale [4,27]: A widely used clinical measure of disability in PwP, the HY scale classifies seven stages of disease which are rated by a clinician. The scale is regarded as fulfilling reasonable criteria for reliability and validity [26]. All HY data was obtained in clinic visits.

2.3. Statistical analysis

Trial data from baseline and three follow up points (one, two and three years) were subject to analysis. The data are analysed broken down by 'early'/'late' category, but not analysed by treatment arm. Descriptive statistics (mean, standard deviation (SD), median, minimum, maximum) were calculated for the PDQ-39 and PDQ-8 single indices. Concordance between the two indices was evaluated by the intraclass correlation coefficient (ICC; two-way mixed average, absolute agreement) in conjunction with the calculation of 95% confidence intervals (CIs). Mean change scores were calculated for the summary index of the PDQ-39 and the PDQ-8. Effect sizes, i.e. change in score in relation to its SD [28] were also calculated for the summary index on both the PDQ-39 and PDQ-8. Scores on the two PDQ indices were correlated with the HY scale cross-sectionally, using Spearman's rho. Data was analysed using SPSS Version 19.

3. Results

Tables 1 and 2 report scores on the PDQ-39-SI and the PDQ-8-SI, broken down by 'early' and 'late' PD respectively, for those respondents who completed all items on the PDQ-39 which enables calculation of the summary scores. No meaningful differences were found between scores on the PDQ-39-SI and the PDQ-8-SI at any of the time points. Indeed, mean differences between the two scores were very small, ranging from 0.6 to 1.1 points. ICCs suggested that the results for both 'early' and 'late' respondents from both measures were remarkably similar at each time point with ICCs ranging from 0.93 (95% CI 0.92–0.94) to 0.96 (95% CI 0.96–0.97). Change scores on the two versions of the PDQ were calculated and ICCs calculated between them and ranged from 0.89 (95% CIs 0.88–0.90) to 0.90 (95% CIs 0.89–0.91).

The sensitivity to change of the PDQ-8-SI was compared with that of the PDQ-39-SI. Mean change scores over time were found to be similar and found to be highly correlated (ICCs ranged from 0.89 to 0.90). Table 3 reports mean changes between baseline and follow up at one, two and three years. Effect sizes were also calculated, and indicate the PDQ-8-SI replicates the results of the parent form.

Scores from respondents assessed by a clinician on the Hoehn and Yahr scale (H&Y) are presented in Table 4. PDQ-39-SI and PDQ-8-SI scores were correlated with this score at the three follow up points. Both versions of the PDQ correlated moderately, and, importantly, reflected similar levels of magnitude, with the H&Y