Health-related quality of life as an outcome variable in Parkinson's disease

Pablo Martinez-Martin and Mónica M. Kurtis

Ther Adv Neurol Disord

[2012] 5(2) 105–117

DOI: 10.1177/

1756285611431974

© The Author(s), 2012.

Reprints and permissions:
[http://www.sagepub.co.uk/
journalsPermissions.nav](http://www.sagepub.co.uk/journalsPermissions.nav)

Abstract: In the past three decades, health-related quality of life (HRQoL) has become an outcome variable in Parkinson's disease clinical trials. This review considers the measuring tools that have been developed, suitability of data reporting, complexity of outcome interpretation, and clinical application to provide evidence regarding available therapeutic interventions to date. In the introduction, different terms regarding quality of life are clearly defined. The methodology section offers an overview of generic, disease specific, and recommended HRQoL scales in Parkinson's disease and the most important psychometric attributes a scale should meet. The interpretation of HRQoL outcomes is complex and not intuitive. Thus, appropriate reporting of data is crucial in order to calculate relative change, a result that facilitates understanding to what extent an intervention is beneficial. The concept of minimally important change/difference is explained as well as the different approaches to its calculation (anchor-based and distribution-based methods). In the results section, a brief overview of the impact on HRQoL of currently available treatments in Parkinson's disease is provided. Special emphasis is given to data assessment, highlighting reports that helped understanding of the clinical significance of the intervention and therefore aided in making therapeutic decisions. The discussion section emphasizes the need for more clinical trials with HRQoL as a primary outcome and standardized reporting in order to further our understanding of the complexity of treatment effects and make evidence-based clinical decisions regarding HRQoL in patients with Parkinson's disease.

Keywords: anchor-based methods, clinical trials, distribution-based methods, health-related quality of life, interpretability, minimally important difference, number needed to treat, Parkinson's disease

Introduction

Currently, the expression 'quality of life' is frequently used in conversation and by mass media to evoke feelings of well being and satisfaction in important aspects for the individual that are related to personal judgment and social standards. More than three decades ago, the term was introduced in medicine as the objective of care, in contrast to the indiscriminate 'quantity of life' ('to live better rather than to live more'). Since then, terminology regarding quality of life in medical settings and literature has expanded and can be rather confusing since there are no universally accepted definitions and differing terms are erroneously used interchangeably. In our understanding, quality of life encompasses 'individuals' perception of their position in life in the context of the culture and value systems in which they

live and in relation to their goals, expectations, standards and concerns' [The WHOQOL Group, 1995], whereas health-related quality of life (HRQoL) is 'the perception and evaluation, by patients themselves, of the impact caused on their lives by the disease and its consequences' [Martinez-Martin, 1998]. Health status influences and predicts HRQoL, but should not be considered equivalent to HRQoL [Den Ouden *et al.* 2007] since it assesses physical and mental symptoms, disability, and social dysfunction related to a health situation in a mere descriptive way, lacking judgments about their impact on the individual's well being and expectations.

The arrival of the current biopsychosocial model to medicine has led to participation of patients in their own care, thus their judgments

Correspondence to:
P. Martinez-Martin MD, PhD

Alzheimer Disease Research Unit, CIEN Foundation and CIBERNED, Carlos III Institute of Health, Alzheimer Center Reina Sofia Foundation, C/. Valderrebollo 5, 28031 Madrid, Spain

pmartinez@fundacioncien.es

Mónica M. Kurtis MD
Movement Disorders Unit, Department of Neurology, Hospital Ruber Internacional, Madrid, Spain

about perceived health status, preferences and satisfaction with treatment are now considered essential part of health care. Subsequently, instruments for measurement of these judgments have been developed and are integrated in assessment procedures as ‘patient-reported outcomes’ (PROs), defined as ‘information provided by patients on their health status and treatment’. PROs are scientific measures for evaluation of health status and change and they should be developed, validated, and analyzed as any other variable in a study [Acquadro *et al.* 2003]. Together with clinician-reported evaluation, ancillary tests (biological, physiological, imaging), and caregiver-reported assessment, PROs are useful indicators of disease activity and treatment efficacy and thus should be included in clinical trials and taken into account for evidence-based practice [European Medicines Agency, 2005; Food and Drug Administration, 2009]. HRQoL is a kind of PRO which refers to quality of life aspects affected by health status and health care.

Methods

HRQoL scales

The most frequently included domains in HRQoL measures are symptoms, physical function, mental status (emotion, cognition), social (well being, activities), role function, and global judgment of health [Berzon *et al.* 1993; Schipper *et al.* 1996; Fitzpatrick *et al.* 1998; Fitzpatrick and Alonso, 1999]. Therefore, HRQoL measures combine personal judgments and reflections about well being and satisfaction with health. HRQoL scales may be classified as generic or specific. Generic scales, in turn, may be health profiles or utility measures. Profiles adopt the format of questionnaires and seek to cover the most important domains of health. Typical examples of this category are the Sickness Impact Profile (SIP) [Bergner *et al.* 1976], the Nottingham Health Profile (NHP) [Hunt *et al.* 1981] and the Medical Outcomes Study-Short form (SF-36) [Ware and Sherbourne, 1992]. Utility measures provide a unique number (‘utility’) that represents quality of life status in relation to best health status (1.0) and death (0.0). They are used for econometric analysis and decision making in health policy. The EQ-5D [EuroQol Group, 1990] is a prototypic example of these kind of measures. Their possible disadvantages are related with potential inability to assess areas of specific interest and lack of sensitivity to detect changes.

Specific instruments address areas of particular interest for a disease, population, function, or dimension. Their advantages are the ability to evaluate health problems with high prevalence in the target population for which they were developed and their responsiveness. Therefore, specific instruments are more appropriate than generic measures for clinical trials, although ideally they should be combined to obtain relevant data on general health.

Table 1 shows a list of the generic and specific HRQoL scales used in Parkinson’s disease (PD) with their respective acronym and some of their basic characteristics. These scales have been recently reviewed by an *ad hoc* Movement Disorder Society Task Force concerning their psychometric attributes and validity evidence. Four generic (SIP, NHP, SF-36, EQ-5D) and five specific (PDQ-39, PDQ-8, PIMS, PDQL, SCOPA-PS) scales gained the qualification of ‘recommended’ for application in patients with PD. That designation means a scale fulfills ‘criteria of successful clinimetric testing, established validation, reliability, and sensitivity assessments, and prior application in PD’ [Goetz *et al.* 2007; Martinez-Martin *et al.* 2011].

Scale attributes that influence outcomes

The outcome analysis of a clinical trial including HRQoL as an endpoint of interest will be affected by the characteristics of the scale applied [Sprangers *et al.* 2002; Marquis *et al.* 2004; European Medicines Agency, 2005; Food and Drug Administration, 2009]. The most relevant psychometric properties when testing a scale are summarized in Table 2. A study can be unsuccessful if any of these problems occur: missing data in a considerable proportion of the sample; accumulation of patients scoring in the high or low end of the scale; indexes reflecting high random error due to poor item cohesion or measure stability; lack of association with other measures applied for assessing the same or related constructs or, in contrast, nonnegligible correlation with measures evaluating constructs nonrelated with that of interest; large distance between correlative scores; and lack of sensitivity to capture a real change.

Interpretation of outcomes

Validation procedures in the realm of HRQoL instruments are quite uniform, but interpretation of outcomes is not, and this fact leads to

Table 1. Health-related quality of life scales applied in Parkinson's disease studies.

Scale	Acronym	Number of		Time frame	Reference
		Items	Domains		
Generic Scales					
<i>Sickness Impact Profile</i>	SIP	136	12	Today	Bergner <i>et al.</i> [1976]
<i>Nottingham Health Profile</i>	NHP	38	6	At present time	Hunt <i>et al.</i> [1981]
<i>EuroQoL 5 Dimensions</i>	EQ-5D	5	5	Today	EuroQol Group [1990]
<i>Medical Outcomes Study-Short Form 36 items</i>	SF-36	36	8	Last 4 weeks	Ware and Sherbourne [1992]
<i>Schedule for the Evaluation of Individual Quality of Life – Direct Weighting</i>	SEIQOL-DW	–	5	At present time	Hickey <i>et al.</i> [1996]
<i>Quality of life questionnaire 15 Dimensions</i>	15-D	15	15	At present time	Sintonen [2001]
WHO Quality of Life Assessment Short Version	WHOQOL-BREF	26	4	Last 4 weeks	Saxena <i>et al.</i> [2001]
<i>Questions on Life Satisfaction</i>					
– <i>Movement Disorders</i>	QLS -MD	12	1	Last 4 weeks	Kuehler <i>et al.</i> [2003]
– <i>Deep Brain Stimulation</i>	QLS-DBS	5	1		
Specific Scales					
<i>Belastungsfragebogen Parkinson Kurzversion</i>	BELA-p-k	19	4	At present time	Ellgring <i>et al.</i> [1993]
<i>Parkinson’s Disease Questionnaire</i>					
– <i>39 items</i>	PDQ-39	39	8	Last month	Peto <i>et al.</i> [1995],
– <i>8 items</i>	PDQ-8	8	8		Jenkinson <i>et al.</i> [1997]
<i>Parkinson’s Impact Scale</i>	PIMS	10	4	At present time	Calne <i>et al.</i> [1996]
<i>Parkinson’s Disease Quality of Life Questionnaire</i>	PDQL	37	4	Last 3 months	de Boer <i>et al.</i> [1996]
<i>Fragebogen Parkinson Lebensqualität</i>	Fragebogens PLQ	44	9	Past week	Van den Berg [1998]
<i>Parkinson’s Disease Quality of Life Scale</i>	PDQUALIF	32	7	Flexible	Welsh <i>et al.</i> [2003]
<i>Parkinson’s Problem Schedule</i>	PPS	39	3	(unknown)	Brod <i>et al.</i> [1998]
<i>Scales for Outcomes in Parkinson’s Disease-Psychosocial</i>	SCOPA-PS	11	1	Last month	Marinus <i>et al.</i> [2003]

uncertainty and confusion. Difficulty in establishing the meaningfulness of change is related to the complexity of the outcome itself (difference by severity level; multiple items and domains; influence of nonrelated aspects to the disease/treatment) [Crosby *et al.* 2003; European Medicines Agency, 2005] and to the lack of intuitive association between the magnitude of change and significance.

Reporting of HRQoL is often incomplete since most studies simply report the mean and standard deviation at baseline and at the end of follow up, with a *p* value that determines statistical significance (calculated, for example, by a *t*-test or repeated measures analysis of variance [ANOVA]). However, it is important to keep in mind that *p* value 'is the probability of having observed our data when the null hypothesis is true' [Altman,

1991]. The usual cut-off, *p* = 0.05, is an arbitrary level and nothing will be very different if a *p* value is 0.045 or 0.055. Really, a '*p*' value depends on the standard error of the observed value (related, in turn, to the sample size) and informs about the null hypothesis, but says nothing about the real meaningfulness of the observed difference [Osoba *et al.* 1998; Hurley, 2010]. Confidence interval and power of the study add relevant information for interpreting the certainty of results [Altman, 1991; Sprangers *et al.* 2002; Charter and Feldt, 2009], but give no information on the importance of change. A better way of approaching the significance of change is to provide the percentage that the change represents of the baseline value (relative change = (mean₂ – mean₁)/ mean₁) [Deyo and Centor, 1986]. The higher the magnitude of relative change, the more likely that it will be meaningful.

Table 2. Psychometric properties especially relevant for outcome analysis.

Attribute	Meaning
Data quality	The completeness of item- and scale-level data
Acceptability	Extent to which the scores are well distributed in the sample
Reliability	Extent to which the scale is free from random error
A - Internal consistency	– intercorrelations of the items at one point in time
B - Reproducibility	– stability over time (test–retest) and interrater agreement
Construct validity	Degree of coherence in the behavior of scores according to theoretical assumptions related with the concept measured
A - Convergent and divergent validity	– degree of association with other measures used for evaluating the same construct or a nonrelated construct
B - Known-groups validity	– ability of the scale to detect differences between (extreme) groups at a point in time
Precision (sensitivity)	Ability of the instrument to detect small differences
Responsiveness	Ability of the scale to detect change over time
Interpretability	Degree to which an easily understandable meaning can be assigned to the quantitative scores

For interpretation of clinical significance when analyzing HRQoL measures, the minimally important change (MIC) or difference (MID) [Jaeschke *et al.* 1989, 1991; Barrett *et al.* 2005; Sloan *et al.* 2005] is a relevant criterion [European Medicines Agency, 2005]. A definition of MID is ‘the smallest difference in score in the outcome of interest that informed patients or informed proxies perceive as important, either beneficial or harmful, and that would lead the patient or clinician to consider a change in the management’ [Brozek *et al.* 2006]. The MID works as a threshold value, marking the border between a negligible and a significant benefit (or harm).

Two strategies have been used most frequently to determine a MID:

1. The anchor-based approach considers the relationship between change in the HRQoL measure and an independent clinical or patient-based measure (the anchor). The aim is to assign patients to groups showing no change, mild, moderate, or large (positive or negative) change in health status. The anchor should be easily interpretable and closely related with the target [Guyatt *et al.* 2002]. It may be an external measure of symptoms, severity level, functional status, etc. or a within-patient rating, typically in the format of a transition question on, for example, a scale from 0 (no change) to 7 (a very large change) (Table 3). The mean change in score observed for the group of patients referring small change is

considered the MID [Jaeschke *et al.* 1989; Juniper *et al.* 1994; Guyatt *et al.* 2002].

2. Distribution-based methods seek to determine MID on the basis of the observed scores and, therefore, they simply express the change in a standardized way. Different formulas have been proposed for this purpose, such as the effect size $[(\text{mean}_2 - \text{mean}_1) / \text{SD}_1]$ [Cohen, 1977; Terwee *et al.* 2003], the standardized response mean $[(\text{mean}_2 - \text{mean}_1) / \text{SD}_{\text{difference}}]$ [Crosby *et al.* 2003; Terwee *et al.* 2003], half a standard deviation at baseline [Norman *et al.* 2003; Sloan *et al.* 2005], the standard error of measurement $[\text{SEM} = \text{SD} \times \sqrt{(1 - r_{xx})}]$ [Wyrwich *et al.* 1999, 2005], and the reliable change index $[1.96 \times S_{\text{diff}}]$ [Jacobson and Truax, 1991; Fitzpatrick *et al.* 2004].

There is no consensus about which of the proposed methods is most suitable. Anchor-based methods are conceptually more sound for determining MID, but they have some problems: transition scales are single-item measures, whereas HRQoL is a complex construct; the anchor choice may be inappropriate; minimal, moderate and large change levels are arbitrary, dependent on the patients’ subjectivity and lacking clinical control; and the assessment is based on a retrospective judgment, which may be defective. On the other hand, distribution-based methods do not directly explore the importance of change and the relationship between values obtained by distribution-based methods and clinical significance is not intuitive; some methods depend heavily on

Table 3. Transition questions usable as within-patient anchor measures.

Rating	Change
0 – No change	No
1 – Almost the same; hardly any better (worse) at all	Small
2 – A little bit better (worse)	Moderate
3 – Somewhat better (worse)	
4 – Moderately better (worse)	
5 – A good deal better (worse)	Large
6 – A great deal better (worse)	
7 – A very great deal better (worse)	
Based on Guyatt <i>et al.</i> [2002].	

the sample's characteristics (although this may be appropriate); and some (e.g. SEM) really determine the minimal detectable or real change rather than an important change [Crosby *et al.* 2003; de Vet *et al.* 2006; Martinez-Martin *et al.* 2008; Revicki *et al.* 2008; Turner *et al.* 2010].

Finally, the MID may be challenged in some aspects: is the same magnitude of effect valid for another sample? Is it valid for mild and severe patients? Is the change equally significant across the range of scores of a scale? Determining the MID for each sample and context through estimations from several methods (triangulation) to offer a range of values presumably containing the real MID value [Wyrwich *et al.* 2007; Revicki *et al.* 2008] has been proposed. Another approach is based on dichotomization of patients according to whether they reach or do not reach, for example, a certain score threshold (e.g. 65/100 points), a change by a certain score (e.g. increase of 3 points or more) or a percentage of score of the baseline (e.g. improve by 10% or more), or an improvement over the established MID. Using proportions, this data may be transformed into the number needed to treat (i.e. $NNT = 1/\text{proportion of patients achieving benefit}$) [Guyatt *et al.* 1998, 2002; Norman *et al.* 2001; Brozek *et al.* 2006; Schünemann *et al.* 2006], a well-understood concept that helps the interpretation of outcomes at a glance.

Results

Application of HRQoL measures to PD treatments

The discovery of the beneficial effects of treatment with levodopa [Barbeau, 1962; Cotzias

et al. 1967] as a precursor and source of lacking dopamine in PD, started the era of rational medical treatment in the late 1960s. Since then, other strategies have been developed and currently PD is the only neurodegenerative disease with efficacious treatment since both medical and surgical therapies have demonstrated successful symptomatic control. However, treatment evaluation is complex as the spectrum of PD symptoms has broadened to include many nonmotor symptoms. Furthermore, other considerations such as social and professional limitations and treatment adverse events must be taken into account. When evaluating intricate treatment effects in clinical trials, a multidimensional construct such as HRQoL is useful, since it is a comprehensive outcome that informs about the final effect from the patient's perspective. In this results section, the impact of currently available treatments in PD on HRQoL is reviewed. Special emphasis is given to appropriate reporting of data, highlighting reports that helped to understand the clinical significance of the intervention and therefore aided in making therapeutic decisions.

L-dopa

To the best of the authors' knowledge there are no clinical trials evaluating the effect on HRQoL of standard levodopa (L-dopa) therapy *versus* placebo. This is probably because the beneficial effect of dopamine therapy was well established before the concept of HRQoL was developed in the late 1970s. There are data from four clinical trials testing controlled release *versus* standard L-dopa or transfer from standard to controlled-release formulations. According to the most recent review on the effect of medical and surgical interventions on HRQoL in PD [Martinez-Martin and Deuschl, 2007], these trials have shown significant improvement in HRQoL in favor of controlled release formulation, in total scores [Grandas *et al.* 1998; Martinez-Martin *et al.* 1998, 1999; Martinez-Martin and Koller, 1999], in all domains [Martinez-Martin, 1998; Martinez-Martin and Koller, 1999], or for some dimensions (home management, emotions, pain, and physical mobility) [Pahwa *et al.* 1997; Grandas *et al.* 1998; Martinez-Martin *et al.* 1999]. The studies in which the magnitude of effect on HRQoL was reported showed 25% improvement with a generic measure and 30% with a specific HRQoL scale. The authors conclude that there is level I and III evidence showing that conversion from standard L-dopa to controlled

release-L-dopa is efficacious in improving HRQoL, even at long term.

Duodenal L-dopa infusion

The first clinical trial with HRQoL as an outcome testing the effect of a duodenal L-dopa infusion suggested that this intervention may be more efficacious than oral L-dopa to improve patient HRQoL in the short term [Nyholm *et al.* 2005]. The trial was randomized, controlled and used a disease-specific scale (PDQ-39) and a generic scale (15D) to measure HRQoL but only included 24 patients with 1.5 months follow up and is therefore considered level II evidence. PDQ-39 summary index (SI) scores at the end of the trial were lower for the infusion group and median scores of seven of the eight dimensions of the PDQ-39 were significantly decreased with infusion. No data on baseline PDQ SI scores were reported, therefore RC and ES could not be calculated.

A more recent prospective, open-label clinical trial evaluating the effect of L-dopa infusions provided more data regarding HRQoL which was established as a secondary outcome (nonmotor symptoms measured by the NMSS were the primary outcome) [Honig *et al.* 2009]. Based on their sample of 22 patients, this group reported mean baseline (44.2 ± 18.4) and follow up (20.7 ± 12.0) PDQ-8 scores, and calculated mean change in score of 23.4 ($p = 0.0003$), relative change = -53%, and ES = 1.3. This article is an example of good data reporting, reporting outcomes based on magnitude of change and distribution-based methods.

Dopamine agonists

The effect of dopamine agonists (DAs) on management of PD symptoms has been tested by a wide number of clinical trials. To the best of the authors' knowledge, the first article reporting the effect of a DA on HRQoL applied generic measures (FSQ and EQ-5D) and was published in 1997 [Guttman, 1997]. Since then, other reported data have included HRQoL results, but according to the most recent review [Martinez-Martin and Kurtis, 2009] no study has included this construct as a primary outcome. Only one clinical trial [Noyes *et al.* 2006] reported a *post hoc* study specifically designed to analyze HRQoL. However, it was based on the extension of a study comparing the development of motor complications with pramipexole *versus* L-dopa as initial treatment

[Parkinson Study Group, 2000; Holloway *et al.* 2004]. The 4-year extension study provided ambiguous [Holloway *et al.* 2004] and conflicting *post hoc* results [Noyes *et al.* 2006].

Using a procedure similar to that used by the Movement Disorders task force [Task Force of the Movement Disorder Society, 2002], this review classified DAs according to their efficacy in improving HRQoL. The authors conclude that there is level I evidence demonstrating that pramipexole, extended release ropinirole, rotigotine, cabergoline and pergolide are efficacious in improving HRQoL as add-on therapy to L-dopa in patients with wearing-off and motor fluctuations. The most widely studied agonist was pramipexole where there was level I evidence demonstrating that it is efficacious in improving HRQoL (SF-36), at 3 months follow up in PD patients with depression on stable L-dopa treatment, without motor fluctuations or dyskinesia [Barone *et al.* 2006]. Recently, a study on mild to moderate PD patients confirmed a modestly beneficial effect of pramipexole on depression [Barone *et al.* 2010].

However, as initial treatment, there is level I evidence showing that pramipexole is nonefficacious in improving HRQoL (EQ-5D, PDQUALIF) at 2 years follow up [Parkinson Study Group, 2000]. This is also the case for rotigotine, where a class I study demonstrated that this agonist is nonefficacious in improving HRQoL (as measured by the EQ-5D) of early PD patients not receiving levodopa or other DA, at 6 months follow up [Jankovic *et al.* 2007]. To date, there is insufficient evidence regarding subcutaneous apomorphine injections and infusions as to their effect on HRQoL.

One of the highest quality agonist trials was an evidence level I study which favored tolcapone as an add on to L-dopa therapy over pergolide in patients with fluctuations. The authors reported that at 3 months follow up, tolcapone produced significantly greater improvement than pergolide in HRQoL as measured by SIP ($p = 0.005$ RC = 22.6% *versus* 20.3%; ES = 0.32 *versus* 0.28) and PDQ-39 scores ($p = 0.005$; RC = 15.7 *versus* 10.3; ES = 0.59 *versus* 0.35) [Koller *et al.* 2001].

Other medical therapies

There are two randomized, prospective, double-blind, placebo-controlled clinical trials which show improvement of HRQoL with the addition of entacapone to L-dopa in nonfluctuating PD

patients [Olanow *et al.* 2004; Fung *et al.* 2009]. The first study included 750 patients with 26 weeks of follow up and considered quality of life as a secondary outcome, measured by a disease-specific scale (PDQ-39) and two recommended generic scales (SF 36 and SIP). The second more recent study included 184 patients and investigated HRQoL with the PDQ-8. Both of them report change between baseline and final scores for each group, showing significant differences in favor of the patients on entacapone. However, this way of reporting data does not allow for calculations of relative change and effect size. A recent review that included medical treatment with catechol-O-methyltransferase (COMT) and monoamine transferase (MAO) inhibitors [Gallagher and Schrag, 2008] found that the addition of entacapone to patients with motor fluctuations is unclear [Fenelon *et al.* 2003; Reichmann *et al.* 2005]. Tolcapone (only tested in patients with motor fluctuations) resulted in significant improvement in HRQoL measured by generic scales (SIP) in two of four studies [Baas *et al.* 1997; Welsh *et al.* 2000]. These inconsistent results lead to uncertainty about the effect of tolcapone and entacapone as adjuncts to L-dopa therapy to improve patients HRQoL in PD patients with motor fluctuations. Rasagiline has improved HRQoL as monotherapy in early PD [Parkinson Study Group, 2002; Biglan *et al.* 2006] but not clearly in more advanced disease [Parkinson Study Group, 2005].

Surgery

Historically, the first study to use HRQoL measures in PD considered a surgical intervention [Wilson and Goetz, 1990]. Traditionally more surgery studies than medical studies have approached the issue of HRQoL. With respect to lesional surgery, most clinical data on quality of life considered unilateral pallidotomy. Published studies through December 2010, included two, high-quality, class Ib comparative studies, one contrasting this surgery with best medical therapy [de Bie *et al.* 1999] and another comparing it with bilateral subthalamic nucleus (STN) deep brain stimulation (DBS) [Esselink *et al.* 2004]. Considering these studies, including a total of 71 patients, followed for an average of 6 months, the average improvement in HRQoL using the PDQL for unilateral pallidotomy was around 16%.

As far as deep brain stimulation, most clinical trials considering HRQoL to date, regard bilateral

STN DBS. Considering four class I studies on bilateral STN published to date [Esselink *et al.* 2004; Deuschl *et al.* 2006; Follet *et al.* 2010; Williams *et al.* 2010] with a total of 855 patients (mean: 214; 34–366) and mean follow up of 12 months, the average improvement in HRQoL using the PDQ-39 was 15.4%. These results are lower than those reported in reviews dating 4 or 5 years ago, where published trials on bilateral subthalamic DBS showed a HRQoL improvement of 60.5% in generic measures [Martinez Martin and Deuschl, 2007] and 33.8–34.5% in disease-specific scales [Kleiner-Fisman *et al.* 2006; Martinez Martin and Deuschl, 2007]. These reviews concluded that the intervention obtaining maximal improvement in global HRQoL was bilateral subthalamic DBS.

However, recent data challenge this traditionally established view, in favor of unilateral globus pallidum interna (GPI) DBS. In a recent large class I clinical trial considering bilateral GPI DBS, this intervention provided an average improvement in HRQoL that was similar to most recently reported results with bilateral STN DBS (11.2%, as measured by the PDQ-39, at 24 months follow up) [Follet *et al.* 2010]. However, in a class I study considering unilateral GPI DBS in contrast to unilateral STN (included 42 patients with a 6-month follow up), the average improvement in HRQoL as measured by the PDQ-39 after unilateral STN was 14.6%, with ES of 0.38, while results for unilateral pallidal DBS showed a magnitude of change of 38% and an ES of 0.88 [Zahodne *et al.* 2009]. This study is a good example of how to report data since baseline and follow-up scores were reported so that RC and ES could be easily calculated for both interventions.

Exercise interventions

In a recent systematic review and meta-analysis on the effectiveness of exercise for people with PD [Goodwin *et al.* 2008], four studies were found that included HRQoL as an outcome variable [Ellis *et al.* 2005; Schmitz-Hubsch *et al.* 2005; Burini *et al.* 2006; Ashburn *et al.* 2007]. Two of these studies were considered to be of high quality [Ellis *et al.* 2005; Burini *et al.* 2006]. Reported findings included 292 participants in total, across three quality of life outcomes, including the Sickness Impact Profile (SIP-68), the Parkinson's disease questionnaire (PDQ-39), and the EuroQOL (EQ-5D). Only one reported a statistically significant

benefit in favor of the exercise intervention group [Ashburn *et al.* 2007]. The authors extracted relevant data from the four papers and calculated the standardized mean difference on the ES (SMD = 0.27, 95% CI 0.04–0.51). Their results suggest that exercise interventions are likely to result in improvements in HRQoL, although more good quality studies are needed.

Discussion

HRQoL measures are currently essential for the evaluation of patients in clinical research. A very recent meta-analysis on HRQoL in later PD found that only three studies included patient-rated quality of life measures [Stowe *et al.* 2011]. These results highlight the need for including (generic and specific) HRQoL assessments in all clinical trials. For trials not focusing specifically on this construct, short versions or brief scales (e.g. EQ-5D and PDQ-8) may be enough, although more precise information will be needed when HRQoL is the main endpoint (e.g. SF-36 and PDQ-39). We would like to insist on the need to use HRQoL measurement as a primary outcome variable in adequately powered clinical trials [Martinez-Martin and Deuschl, 2007]. The lack of studies focusing mainly on HRQoL is paradoxical since this construct has been proposed as a comprehensive and rational outcome for evaluating the complex effects potentially related to PD therapy [Clarke and Guttman, 2002; Martinez-Martin and Kurtis, 2009].

This review emphasizes the need to standardize reporting of clinical trial outcomes regarding HRQoL, providing enough data (at least, mean $\pm$ SD at baseline and follow up) for the calculation of informative parameters of change, such as RC and ES. Many reports only give *p* values that merely provide statistical probability for rejection of the null hypothesis; they do not give information about the magnitude and clinical importance of change. A small sample size can impede reaching statistical significance whereas other outcome parameters are indicative of a potentially important change. For example, in a study with duodenal L-dopa infusion (*n* = 12) there was a –10 point difference from follow up to baseline in the PDQ-39 summary index (*p* = 0.076) [Isacson *et al.* 2008]. The difference was not ‘statistically significant’, but our calculations show that RC was –27.4% and the ES = 0.88, indicative of a potentially important effect. Inclusion in the protocol of a transition question

allowing to assign the MID directly, details of scoring at baseline at follow up, and/or *a priori* settled thresholds for calculation of NNT are necessary for an appropriate interpretation of outcomes. Obviously, outcomes should be reported for all groups in the study (e.g. placebo and active treatment, surgical and best medical therapy, etc.), both within and between groups. A consensus on a common minimal body of explicit and standardized elements for reporting HRQoL outcomes, similar to the FDA recommendations, would be welcome [Food and Drug Administration, 2009].

Another relevant point is that some therapies have been tested in diverse situations (for example, as add-on therapy in advanced phases of disease and as initial treatment in early stages). This variability requires specific evaluation of intervention impact in each scenario, and obviously results have to be interpreted accordingly. In this sense, it seems necessary to determine the MID for each study, adjusted to the characteristics of the sample and circumstances of the study, rather than considering a fixed value. We propose obtaining MID values from several anchor- and distribution-based methods applied simultaneously, a strategy producing a range of values that should contain the true MID [Wyrwich *et al.* 2007; Revicki *et al.* 2008; Honig *et al.* 2009]. Determining the importance of change is a fundamental step for making decisions in clinical research, pharmacoeconomics, assignment of resources, and health policy.

The application of HRQoL scales in the investigation of PD treatment has demonstrated efficacy of certain interventions. Traditionally, surgical trials have included HRQoL as an outcome more often than clinical drug trials, a trend that is rapidly changing as HRQoL progressively gains attention from the scientific community. Summarizing the literature on medical therapy and HRQoL in PD: controlled release L-dopa has consistently shown improvement; DAs are efficacious as add-on therapy to L-dopa in patients with wearing-off and motor fluctuations; addition of entacapone to L-dopa in stable PD patients has demonstrated improvement; and rasagiline has demonstrated beneficial effects in monotherapy in early PD patients. There is insufficient evidence to conclude on the effects of COMT inhibitors on fluctuating PD patients or on exercise, although they are likely efficacious. In patients with advanced disease, there is level II evidence that intraduodenal infusion of L-dopa improves

HRQoL, while there is insufficient evidence to conclude on the effect of apomorphine infusions. Regarding PD surgery, recent results looking specifically at HRQoL, challenge the traditional supremacy of bilateral STN target in favor of unilateral pallidal DBS [Zahodne *et al.* 2009], although both interventions are efficacious.

More clinical trials with HRQoL as a primary outcome and standardized reporting are necessary to further our understanding of the complexity of treatment effects in order to make evidence based clinical decisions regarding HRQoL in patients with PD.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest statement

The authors declare no conflicts of interest in preparing this article.

References

- Acquadro, C., Berzon, R., Dubois, D., Leidy, N.K., Marquis, P., Revicki, D., *et al.* (2003) Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the patient-reported outcomes (PRO) harmonization group meeting at the Food and Drug Administration, February 16, 2001. *Value Health* 6: 522–531.
- Altman, D.G. (1991) *Practical Statistics for Medical Research*. London: Chapman & Hall, pp. 165–171.
- Ashburn, A., Fazakarley, L., Ballinger, C., Pickering, R., McLellan, L.D. and Fitton, C. (2007) A randomised controlled trial of a home-based exercise programme to reduce risk of falling among people with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 78: 678–684.
- Baas, H., Beiske, A.G., Ghika, J., Jackson, M., Oertel, W.H., Poewe, W., *et al.* (1997) Catechol-O-methyltransferase inhibition with tolcapone reduces the "wearing off" phenomenon and levodopa requirements in fluctuating parkinsonian patients. *J Neurol Neurosurg Psychiatry* 63: 421–428.
- Barbeau, A. (1962) The pathogenesis of Parkinson's disease: a new hypothesis. *Can Med Assoc J* 87: 802–807.
- Barone, P., Poewe, W., Albrecht, S., Debieve, C., Massey, D., Rascol, O., *et al.* (2010) Pramipexole for the treatment of depressive symptoms in patients with Parkinson's disease: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol* 9: 573–580.
- Barone, P., Scarzella, L., Marconi, R., Antonini, A., Morgante, L., Bracco, F., *et al.* (2006) Pramipexole versus sertraline in the treatment of depression in Parkinson's disease: a national multicenter parallel-group randomized study. Depression/Parkinson Italian Study Group. *J Neurol* 253: 601–607.
- Barrett, B., Brown, D., Mundt, M. and Brown, R. (2005) Sufficiently important difference: expanding the framework of clinical significance. *Med Decis Making* 25: 250–261.
- Bergner, M., Bobbitt, R.A., Pollard, W.E., Martin, D.P. and Gilson, B.S. (1976) The Sickness Impact Profile: validation of a health status measure. *Med Care* 14: 57–67.
- Berzon, R., Hays, R.D. and Shumaker, S.A. (1993) International use, application and performance of health-related quality of life instruments. *Qual Life Res* 2: 367–368.
- Biglan, K.M., Schwid, S., Eberly, S., Blindauer, K., Fahn, S., Goren, T., *et al.* (2006) Rasagiline improves quality of life in patients with early Parkinson's disease. *Mov Disord* 21: 616–623.
- Brod, M., Mendelsohn, G.A. and Roberts, B. (1998) Patients' experiences of Parkinson's disease. *J Gerontol Psych Sci* 53B: 213–222.
- Brozek, J.L., Guyatt, G.H. and Schünemann, H.J. (2006) How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. *Health Qual Life Outcomes* 4: 69.
- Burini, D., Farabollini, B., Iacucci, S., Rimatori, C., Riccardi, G., Capecci, M., *et al.* (2006) A randomised controlled cross-over trial of aerobic training versus Qigong in advanced Parkinson's disease. *Eur Med J* 42: 231–238.
- Calne, S., Schulzer, M., Mak, E., Guyette, C., Rohs, G., Hatchard, S., *et al.* (1996) Validating a quality of life rating scale for idiopathic Parkinsonism: Parkinson's Impact Scale (PIMS). *Parkinsonism Relat Disord* 2: 55–61.
- Charter, R.A. and Feldt, L.S. (2009) A comprehensive approach to the interpretation of difference scores. *App Neuropsychol* 16: 23–30.
- Clarke, C.E. and Guttman, M. (2002) Dopamine agonist monotherapy in Parkinson's disease. *Lancet* 360: 1767–1769.
- Cohen, J. (1977) *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Cotzias, G.C., Van Woert, M.H. and Schiffer, L.M. (1967) Aromatic amino acids and modification of parkinsonism. *N Engl J Med* 276: 374–379.

- Crosby, R.D., Kolotkin, R.L. and Williams, G.R. (2003) Defining clinical meaningful change in health-related quality of life. *J Clin Epidemiol* 56: 395–407.
- de Bie, R.M., de Haan, R.J., Nijssen, P.C., Rutgers, A.W., Beute, G.N., Bosch, D.A., *et al.* (1999) Unilateral pallidotomy in Parkinson's disease: a randomised, single-blind, multicentre trial. *Lancet* 354: 1665–1669.
- de Boer, A.G., Wijker, W., Speelman, J.D. and de Haes, J.C. (1996) Quality of life in patients with Parkinson's disease: development of a questionnaire. *J Neurol Neurosurg Psychiatry* 61: 70–74.
- de Vet, H.C., Terwee, C.B., Ostelo, R.W., Beckerman, H., Knol, D.L. and Bouter, L.M. (2006) Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes* 4: 54.
- Den Ouden, B.L., Van Heck, G.L. and De Vries, J. (2007) Quality of life and related concepts in Parkinson's disease: a systematic review. *Mov Disord* 22: 1528–1537.
- Deuschl, G., Schade-Brittinger, C., Krack, P., Volkmann, J., Schäfer, H., Bötzel, K., *et al.* (2006) A randomized trial of deep-brain stimulation for Parkinson's disease. *N Engl J Med* 355: 896–908.
- Deyo, R.A. and Centor, R.M. (1986) Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 39: 897–906.
- Ellgring, H., Seiler, S., Perleth, B., Frings, W., Gasser, T. and Oertel, W. (1993) Psychosocial aspects of Parkinson's disease. *Neurology* 43(Suppl.): 41–44.
- Ellis, T., de Goede, C.J., Feldman, R.G., Wolters, E.C., Kwakkel, G. and Wagenaar, R.C. (2005) Efficacy of a physical therapy program in patients with Parkinson's disease: a randomised controlled trial. *Arch Phys Med Rehabil* 86: 626–632.
- Esselink, R.A., de Bie, R.M., de Haan, R.J., Lenders, M.W., Nijssen, P.C., Staal, M.J., *et al.* (2004) Unilateral pallidotomy versus bilateral subthalamic nucleus stimulation in PD: a randomized trial. *Neurology* 62: 201–207.
- European Medicines Agency (2005) Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003637.pdf
- EuroQol Group (1990) EuroQol—a new facility for the measurement of health related quality of life. *Health Policy* 16: 199–208.
- Fenelon, G., Gimenez-Roldan, S., Montastruc, J.L., Bermejo, J.F., Durif, F., Bourdeix, I., *et al.* (2003) Efficacy and tolerability of entacapone in patients with Parkinson's disease treated with levodopa plus a dopamine agonist and wearing-off motor fluctuations: a randomized, double-blind, multicentre study. *J Neural Transm* 110: 239–251.
- Fitzpatrick, R. and Alonso, J. (1999) Quality of life in health care: concepts and components. In: Martinez-Martin, P. and Koller, W.C. (eds), *Quality of Life in Parkinson's Disease*. Barcelona: Masson SA, pp. 1–15.
- Fitzpatrick, R., Davey, C., Buxton, M.J. and Jones, D.R. (1998) Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assessm* 2(14): 7–8.
- Fitzpatrick, R., Norquist, J.M. and Jenkinson, C. (2004) Distribution-based criteria for change in health-related quality of life in Parkinson's disease. *J Clin Epidemiol* 57: 40–44.
- Follett, K.A., Weaver, F.M., Stern, M., Hur, K., Harris, C.L., Luo, P., *et al.* (2010) Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. *N Engl J Med* 362: 2077–2091.
- Food and Drug Administration (2009) Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. U.S. Department of Health and Human Services. Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>
- Fung, V.S.C., Herawati, L. and Wan, Y. (2009) Quality of life in early Parkinson's disease treated with levodopa/carbidopa/entacapone. *Mov Disord* 24: 25–31.
- Gallagher, D.A. and Schrag, A. (2008) Impact of newer pharmacological treatments on quality of life in patients with Parkinson's disease. *CNS Drugs* 22: 563–586.
- Goetz, C.G., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stebbins, G.T., *et al.* (2007) Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format and clinimetric testing plan. *Mov Disord* 22: 41–47.
- Goodwin, V.A., Richards, S.H., Taylor, R.S., Taylor, A.H. and Campbell, J.L. (2008) The effectiveness of exercise interventions for people with Parkinson's disease: a systematic review and meta-analysis. *Mov Disord* 23: 631–640.
- Grandas, F., Martínez-Martín, P., Linazasoro, G., for The STAR Multicenter Study Group (1998) Quality of life in patients with Parkinson's disease who transfer from standard levodopa to Sinemet CR: the STAR Study. *J Neurol* 245(Suppl.1): 31–33.

- Guttman, M. (1997) Double-blind comparison of pramipexole and bromocriptine treatment with placebo in advanced Parkinson's disease. International Pramipexole-Bromocriptine Study Group. *Neurology* 49: 1060–1065.
- Guyatt, G.H., Juniper, E.F., Walter, S.D., Griffith, L.E. and Goldstein, R.S. (1998) Interpreting treatment effects in randomised trials. *Br Med J* 316: 690–693.
- Guyatt, G.H., Osoba, D., Wu, A.W., Wyrwich, K.W., Norman, G.R., for The Clinical Significance Consensus Meeting Group (2002) Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 77: 371–383.
- Hickey, A.M., Bury, G., O'Boyle, C.A., Bradley, F., O'Kelly, F.D. and Shannon, W. (1996) A new short form individual quality of life measure (SEIQoL-DW): application in a cohort of individuals with HIV/AIDS. *Br Med J* 313: 29–33.
- Holloway, R.G., Shoulson, I., Fahn, S., Kieburtz, K., Lang, A., Marek, K., *et al.* (2004) Pramipexole vs levodopa as initial treatment for Parkinson disease: a 4-year randomized controlled trial. *Arch Neurol* 61: 1044–1053.
- Honig, H., Antonini, A., Martinez-Martin, P., Forgacs, I., Faye G.C., Fox, T., *et al.* (2009) Intrajejunal levodopa infusion in Parkinson's disease: A pilot multicenter study of effects on nonmotor symptoms and quality of life. *Mov Disord* 24: 1468–1474.
- Hunt, S.M., McKenna, S.P., McEwen, J., Williams, J. and Papp, E. (1981) The Nottingham Health Profile: subjective health status and medical consultations. *Soc Sci Med* 15A: 221–229.
- Hurley, D. (2010) Does statistical significance translate into clinical efficacy? Not always, especially when it comes to PD trials. *Neurol Today* 10: 4–5.
- Isacson, D., Bingefors, K., Kristiansen, I.S. and Nyholm, D. (2008) Fluctuating functions related to quality of life in advanced Parkinson disease: effects of duodenal levodopa infusion. *Acta Neurol Scand* 118: 379–386.
- Jacobson, N.S. and Truax, P. (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 59: 12–19.
- Jaeschke, R., Guyatt, G.H., Keller, J. and Singer, J. (1991) Interpreting changes in quality-of-life score in N of 1 randomized trials. *Control Clin Trials* 12 (Suppl. 4): 226–233.
- Jaeschke, R., Singer, J. and Guyatt, G.H. (1989) Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 10: 407–415.
- Jankovic, J., Watts, R.L., Martin, W. and Boroojerdi, B. (2007) Transdermal rotigotine: double-blind, placebo-controlled trial in Parkinson disease. *Arch Neurol* 64: 676–82.
- Jenkinson, C., Fitzpatrick, R., Peto, V., Greenhall, R. and Hyman, N. (1997) The PDQ-8: Development and validation of a short-form Parkinson's Disease Questionnaire. *Psychol Health* 12: 805–814.
- Juniper, E.F., Guyatt, G.H., Willan, A. and Griffith, L.E. (1994) Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 47: 81–87.
- Kleiner-Fisman, G., Herzog, J., Fisman, D.N., Tamma, F., Lyons, K.E., Pahwa, R., *et al.* (2006) Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. *Mov Disord* 21(Suppl. 14): 290–304.
- Koller, W., Lees, A., Doder, M., Hely, M., for the Tolcapone/Pergolide Study Group. (2001) Randomized trial of tolcapone *versus* pergolide as add-on to levodopa therapy in Parkinson's disease patients with motor fluctuations. *Mov Disord* 16: 858–866.
- Kuehler, A., Henrich, G., Schroeder, U., Conrad, B., Herschbach, P., Ceballos-Baumann, A. (2003) A novel quality of life instrument for deep brain stimulation in movement disorders. *J Neurol Neurosurg Psychiatry* 74: 1023–1030.
- Marinus, J., Visser, M., Martinez-Martin, P., van Hilten, J.J. and Stiggelbout, A.M. (2003) A short psychosocial questionnaire for patients with Parkinson's disease: the SCOPA-PS. *J Clin Epidemiol* 56: 61–67.
- Marquis, P., Chassany, O. and Abeto, L. (2004). A comprehensive strategy for the interpretation of quality-of-life data based on existing methods. *Value Health* 7: 93–104.
- Martinez-Martin, P. (1998) An introduction to the concept of "quality of life in Parkinson's disease". *J Neurol* 245(Suppl. 1): 2–5.
- Martinez-Martin, P., Carod-Artal, F.J., da Silveira Ribeiro, L., Ziolkowski, S., Vargas, A.P., Kummer, W., *et al.* (2008) Longitudinal psychometric attributes, responsiveness, and importance of change: An approach using the SCOPA-Psychosocial questionnaire. *Mov Disord* 23: 1516–1523.
- Martinez-Martin, P., del Ser, T., Aguilar, M. and Peto, V. (1998) The QoL Study Group. Controlled-release levodopa-carbidopa (CRLC) treatment and quality of life of Parkinson's disease patients as measured by the PDQ-39. *Eur J Neurol* 5(Suppl. 3): 171.
- Martinez-Martin, P. and Deuschl, G. (2007) Effect of medical and surgical interventions on health-related

- quality of life in Parkinson's disease. *Mov Disord* 22: 757–765.
- Martinez-Martin, P., Grandas, F., Linazasoro, G., Bravo, J.L., The STAR Study Group (1999) Conversion to controlled-release levodopa/carbidopa treatment and quality of life as measured by the Nottingham Health Profile. *Neurologia* 14: 338–343.
- Martinez-Martin, P., Jeukens-Visser, M., Lyons, K.E., Rodriguez-Blazquez, C., Selai, C., Siderowf, A., et al. (2011) Health-related quality-of-life scales in Parkinson's disease: Critique and recommendations. *Mov Disord* 26: 2371–2380
- Martinez-Martin, P. and Koller, W.C. (1999) Quality of life benefit induced by controlled-release carbidopa/levodopa formulation. In: Martinez-Martin, P. and Koller, W.C. (eds), *Quality of Life in Parkinson's Disease*. Barcelona: Masson SA, pp. 93–105.
- Martinez-Martin, P. and Kurtis, M.M. (2009) Systematic review of the effect of dopamine receptor agonists on patient health-related quality of life. *Parkinsonism Relat Disord* 15(Suppl. 4): 58–64.
- Norman, G.R., Sloan, J.A. and Wywich, K.W. (2003) Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 41: 582–592.
- Norman, G.R., Srdihar, F.G., Guyatt, G.H. and Walter, S.D. (2001) Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 39: 1039–1047.
- Noyes, K., Dick, A.W., Holloway, R., for the Parkinson Study Group (2006) Pramipexole versus levodopa in patients with early Parkinson's disease: effect on generic and disease-specific quality of life. *Value Health* 9: 28–38.
- Nyholm, D., Nilsson Remahl, A.I., Dizdar, N., Constantinescu, R., Holmberg, B., Jansson, R., et al. (2005) Duodenal levodopa infusion monotherapy vs oral polypharmacy in advanced Parkinson's disease. *Neurology* 64: 216–223.
- Olanow, C.W., Kieburtz, K., Stern, M., Watts, R., Langston, J.W., Guarnieri, M., et al. (2004) Double-blind, placebo-controlled study of entacapone in levodopa-treated patients with stable Parkinson disease. *Arch Neurol* 61: 1563–1568.
- Osoba, D., Rodrigues, G., Myles, J., Zee, B. and Pater, J. (1998). Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 16: 139–144.
- Pahwa, R., Lyons, K., McGuire, D., Silverstein, P., Zwiebel, F., Robischon, M., et al. (1997) Comparison of standard carbidopa-levodopa and sustained-release carbidopa-levodopa in Parkinson's disease: pharmacokinetic and quality-of-life measures. *Mov Disord* 12: 677–681.
- Parkinson Study Group (2000) Pramipexole vs levodopa as initial treatment for Parkinson disease: A randomized controlled trial. Parkinson Study Group. *JAMA* 284: 1931–1938.
- Parkinson Study Group (2002) A controlled trial of rasagiline in early Parkinson disease: the TEMPO Study. *Arch Neurol* 59: 1937–1943.
- Parkinson Study Group (2005) A randomized placebo-controlled trial with of rasagiline in levodopa-treated patients with Parkinson disease and motor fluctuations: the PRESTO study. *Arch Neurol* 62: 241–248.
- Peto, V., Jenkinson, C., Fitzpatrick, R. and Greenhall, R. (1995) The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. *Qual Life Res* 4: 241–248.
- Reichmann, H., Boas, J., Macmahon, D., Myllyla, V., Hakala, A., Reinikainen, K., et al. (2005) Efficacy of combining levodopa with entacapone on quality of life and activities of daily living in patients experiencing wearing-off type fluctuations. *Acta Neurol Scand* 111: 21–28.
- Revicki, D., Hays, R.D., Cella, D. and Sloan, J. (2008) Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 61: 102–109.
- Saxena, S., Carlson, D. and Billington, R. (2001) The WHO quality of life assessment instrument (WHOQOL-Bref): the importance of its items for cross-cultural research. *Qual Life Res* 10: 711–721.
- Schipper, H., Clinch, J.J. and Olweny, C.L.M (1996) Quality of life studies: definitions and conceptual issues. In: Spilker, B. (ed.), *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd edn. Philadelphia, PA: Lippincot-Raven Publishers, pp. 11–23.
- Schmitz-Hubsch, T., Pyfer, D., Kielwein, K., Fimmers, R., Klockgether, T. and Wüllner, U. (2005) Qigong exercise for the symptoms of Parkinson's disease: a randomised controlled pilot study. *Mov Disord* 21: 543–548.
- Schünemann, H.J., Akl, E.A. and Guyatt, G.H. (2006) Interpreting the results of patient reported outcome measures in clinical trials: The clinician's perspective. *Health Qual Life Outc* 4: 62.
- Sintonen, H. (2001) The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 33: 328–336.

- Sloan, J.A., Cella, D. and Hays, R.D. (2005) Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol* 58: 1217–1219.
- Sprangers, M.A.G., Moinpour, C.M., Moynihan, T.J., Patrick, D.L., Revicki D.A., for the Clinical Significance Consensus Meeting Group (2002) Assessing meaningful change in quality of life over time: A users' guide for clinicians. *Mayo Clin Proc* 77: 561–571.
- Stowe, R., Ives, N., Clarke, C.E., Handley, K., Furnston, A., Deane, K., *et al.* (2011). Meta-analysis of the comparative efficacy and safety of adjuvant treatment to levodopa in later Parkinson's disease. *Mov Disord* 26: 587–598.
- Task Force of the Movement Disorder Society (2002) Management of Parkinson's disease: an evidence-based review. *Mov Disord* 17(Suppl. 4): 1–6.
- Terwee, C.B., Dekker, F.W., Wiersinga, W.M., Prummel, M.F. and Bossuyt, P.M. (2003) On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res* 12: 349–362.
- The WHOQOL Group (1995) The World Health Organization Quality of Life Assessment (WHOQOL): position paper from the world health organization. *Soc Sci Med* 41: 1403–1409.
- Turner, D., Schünemann, H.J., Griffith, L.E., Beaton, D.E., Griffiths, A.M., Critch, J.N., *et al.* (2010) The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol* 63: 28–36.
- Van den Berg, M. (1998) Leben mit Parkinson: Entwicklung und psychometrische Testung des Fragenbogens PLQ. *Neurol Rehabil* 4: 221–226.
- Ware, J.E. and Sherbourne, C.D. (1992) The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30: 473–483.
- Welsh, M., McDermott, M.P., Holloway, R.G., Plumb, S., Pfeiffer, R., Hubble, J., *et al.* (2003) Development and testing of the Parkinson's disease quality of life scale (PDQUALIF). *Mov Disord* 18: 637–645.
- Welsh, M.D., Dorflinger, E., Chernik, D. and Waters, C. (2000) Illness impact and adjustment to Parkinson's disease: before and after treatment with tolcapone. *Mov Disord* 15: 497–502.
- Williams, A., Gill, S., Varma, T., Jenkinson, C., Quinn, N., Mitchell, R. *et al.* (2010) Deep brain stimulation plus best medical therapy *versus* best medical therapy alone for advanced Parkinson's disease (PD SURG trial): a randomised, open-label trial. *Lancet Neurol* 9: 581–591.
- Wilson, R.S. and Goetz, C.G. (1990) Neurologic illness. In: Spilker, B. (ed.), *Quality of Life Assessment in Clinical Trials*. New York: Raven Press Ltd, pp. 347–356.
- Wyrwich, K.W., Bullinger, M., Aaronson, N., Hays, R.D., Patrick, D.L., Symonds, T., *et al.* (2005) Estimating clinically significant differences in quality of life outcomes. *Qual Life Res* 14: 285–295.
- Wyrwich, K.W., Metz, S.M., Kroenke, K., Tierney, W.M., Babu, A.N. and Wolinsky, F.D. (2007) Triangulating patient and clinician perspectives on clinically important differences in health-related quality of life among patients with heart disease. *Health Serv Res* 42: 2257–2274.
- Wyrwich, K., Nienaber, N.A., Tierney, W.M. and Wolinsky, F.D. (1999) Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 37: 469–478.
- Zahodne, L.B., Okun, M.S., Foote, K.D., Fernandez, H.H., Rodriguez, R.L., Wu, S.S., *et al.* (2009) Greater improvement in quality of life following unilateral deep brain stimulation surgery in the globus pallidus as compared to the subthalamic nucleus. *J Neurol* 256: 1321–1329.

Visit SAGE journals online
<http://tan.sagepub.com>

SAGE JOURNALS
 Online