Respondent burden and patient-perceived validity of the PDQ-39.

Kim, M-Y; Dahlberg, A; Hagell, Peter

Published in:
Acta Neurologica Scandinavica

DOI:
10.1111/j.1600-0404.2005.00549.x

2006

Respondent burden and patient perceived validity of the PDQ-39

Mi-Young Kim, RN BSc[1*], Anneli Dahlberg, RN BSc[1*], Peter Hagell, RN PhD[1,2]

[1] Department of Health Sciences, Lund University, Lund, Sweden
[2] The Vårdal Institute, the Swedish Institute for Health Science, Lund University, Lund, Sweden

*) Equal contributions by both authors.

Number of words (main text): 2558

Short title: Respondent burden & perceived validity of the PDQ-39

Corresponding author:
Peter Hagell
Department of Health Sciences
Lund University
P.O. Box 157
SE-221 00 Lund
Sweden
Tel: +46 46 222 1937
E-mail: Peter.Hagell@med.lu.se

Respondent burden & perceived validity of the PDQ-39

Abstract

Objectives: To evaluate the respondent burden and patient perceived content validity of the Parkinson's disease (PD) specific health status questionnaire PDQ-39, and the linguistic validity of its revised Swedish version.

Material & methods: Eighteen PD patients completed the revised Swedish version of the PDQ-39. Respondent burden was assessed by the time taken to complete the questionnaire. Content and linguistic validity was evaluated qualitatively.

Results: Patients with mild, moderate and advanced PD needed a mean of 9.5, 11.3 and 20.1 minutes, respectively, to complete the PDQ-39. One third of patients identified irrelevant items and 50% identified important health-related areas that were missing. Revisions had eliminated previous linguistic problems with the Swedish PDQ-39.

Conclusions: Undue respondent burden challenges the appropriateness of the PDQ-39 among patients with more advanced disease. Overall content validity was acceptable but compromised by lack of important content areas. Observations support the linguistic validity of the revised Swedish PDQ-39.

Key words: Health status, Parkinson’s disease, Respondent burden, Validity

In order for patient-reported outcome measures (PRO) to yield data that can be interpreted with confidence it is essential that they are scientifically sound. In addition to quantitative measurement properties, there are also qualitative properties that need to be considered. Content validity relates to how well the content of an instrument covers the construct it intends to tap, and is evaluated through critical review of its items in relation to the intended purpose (1). In the case of PRO intended to reflect health from the perspective of the patient it is central that patients representing the instrument's target population are involved in establishing content validity (2, 3). Another important aspect of the usefulness of PRO is the burden they pose upon patients requested to complete them (4). The Scientific Advisory Committee of the Medical Outcomes Trust (5) has thus highlighted respondent burden as a main attribute when reviewing the quality and suitability of PRO. Respondent burden is perhaps particularly important when a PRO is to be used as one out of several outcomes in clinical trials and research among people with chronic, disabling disorders such as Parkinson's disease (PD). In order to be a feasible component of clinical research and practice, PRO should thus be easy, unambiguous and relatively quick for patients to complete (2). In the case of the 39-item PD Questionnaire (PDQ-39) (6), the most widely used disease-specific health status questionnaire for PD (7), respondent burden does not appear to have been documented and direct patient evaluations of its content validity have been few and shown mixed results (8, 9).

Increasing international conduct of clinical research emphasizes another important aspect of the usefulness of self-administered PRO, i.e., to ensure that questionnaires are valid across cultures and languages (2-5, 10). A fundamental aspect of the cross-cultural adaptation of PRO is to ensure that the new translation is appropriate (11, 12). Linguistic shortcomings may, e.g., alter the conceptual meaning of items and/or cause ambiguities that influence responses and render inequalities across language versions that necessarily are not captured by traditional psychometric indices (13). Evaluations of the original Swedish version of the PDQ-39 thus identified linguistic shortcomings that influenced patient responses and appeared to compromise measurement validity (8, 14).

The objectives of this study were to evaluate the respondent burden and patient perceived content validity of the PDQ-39, and the linguistic validity of its revised Swedish version.

Material and methods

The study was approved by the Research Ethics Committee, Faculty of Medicine, Lund University, Sweden.

Questionnaire

The PDQ-39 consists of 39 items covering eight subscales (15). Each item relates to a frame question, or question stem (*Due to having Parkinson's disease, how often during the last month have you…?*), appearing at the top of each questionnaire page. For each item, respondents are requested to affirm one of five response categories ("never" – "occasionally" – "sometimes" – "often" - "always, or cannot do at all"). In a previous study on the original Swedish version of the PDQ-39 (8), three aspects of the questionnaire were identified as particularly problematic: (i) anonymity of the frame question, (ii) ambiguous distinction between response categories, in particular between "occasionally" and "sometimes", and (iii) a double negative in the wording of 2 out of 3 items in the Social Support subscale (items 28 and 29) when read together with the response categories. Although the first was not linguistic but rather related to the appearance of the questionnaire, a majority of patients pointed out that anonymity of the frame question made them fail to consider it, which may have affected their responses. Based on these experiences the questionnaire was revised regarding wording and appearance (but not content). The frame question, along with the response category definitions, was repeated halfway through each questionnaire page. The translation of the response category "occasionally" (original Swedish version: "vid enstaka tillfällen") was changed to "seldom" ("sällan"). The problematic wording of items 28 and 29 (identical wording in both items) in the Social Support subscale was changed from "…not received support…" ("…inte fått stöd…") to "…lacked support…" ("…saknat stöd…"). A minor revision of the wording of the frame question was also made. Revisions were made in communication with the developers of the original British version of the PDQ-39.

Patients

Patients were recruited consecutively from the movement disorder daycare unit at the Department of Neurology, Lund University Hospital, Sweden, during 4 months. In keeping with recommendations for this type of evaluations (16, 17), a gender mixed sample of about 20 patients representing various ages and all five stages of disease severity according to Hoehn & Yahr (HY) (18) was aimed for. Selection criteria were native Swedish patients with

clinically diagnosed idiopathic PD (19), and not previously familiar with the PDQ-39. Patients with significant co-morbidities or other medical conditions were excluded.

Procedures

Patients completed the revised Swedish version of the PDQ-39 according to instructions included in the questionnaire (8, 15). The time taken to complete the questionnaire was recorded for each patient. After completing the PDQ-39, patients were asked whether they experienced any problems with the wording of the questionnaire and whether they found any items difficult to respond to. They were also asked if they thought any items were of no or negligible relevance regarding health, functioning and well-being in PD, and if any important aspects thereof were lacking. Patients were also asked to comment on the ease of using the response categories and to rate (on a 1-10 scale anchored by 1 = "worst possible" and 10 = "best possible") and comment on their overall impression of the questionnaire as a measure of health, functioning and well-being in PD. All ratings, comments and responses were recorded and reviewed for accuracy by the patients at the end of each interview.

Analyses

Due to the restricted sample size patients were categorized into mild (HY stages I-II), moderate (HY stage III) and advanced (HY stages IV-V) PD, and analyzed accordingly. Quantitative variables were checked regarding assumptions underlying the use of parametric and non-parametric statistics and described and analyzed accordingly, using SPSS for Windows, version 12.0.1 (SPSS, Inc., Chicago, IL). Qualitative interview responses were reviewed, summarized and grouped according to content.

Results

Eighteen out of 22 eligible patients gave informed consent (Table 1). All participating patients were treated with levodopa (+ a dopadecarboxylase inhibitor) alone or in combination with a dopamine agonist, COMT-inhibitor, and/or amantadine. Three patients were treated with bilateral deep brain stimulation of the subthalamic nuclei. Participants' mean PD duration was 13 (SD, 8.8; range, 2-30) years.

Patients needed a mean of 14.8 (SD, 7.7; range, 4-30) minutes to complete the PDQ-39. Patients with mild, moderate and advanced PD needed a mean (SD) of 9.5 (1.3), 11.3 (7.1) and 20.1 (6.9) minutes, respectively, to complete the PDQ-39 (Fig. 1). The

difference across grouped HY stages was significant (one-way ANOVA: F=5.235; df=2; P=0.019).

No linguistic problems regarding the wording or intelligibility of any aspect of the revised Swedish PDQ-39 were identified. However, 3 patients (17%) with moderate to advanced PD tended not to relate to the frame question when responding. While the linguistic distinction between response categories was not considered ambiguous, other concerns regarding the response scale were expressed by 8 patients (44%) with moderate to advanced PD (. Of these, 4 found frequency statements difficult to relate to due to unpredictable motor fluctuations. Others expressed general difficulties assessing frequency and/or making retrospective ratings.

Six patients (33%) representing mild, moderate and advanced PD, identified items of no or negligible relevance regarding health, functioning and well-being in PD. One of these, a patient with longstanding advanced PD, perceived only 3 items as relevant. Item 3 ("carrying shopping bags") was the item most frequently considered irrelevant (by 3 patients). Other items were mentioned by one patient each. Nine patients (50%), also representing mild through advanced PD, identified areas that they considered important but missing from the questionnaire. Of these, 5 (with moderate to advanced disease) missed items related to medication and motor fluctuations, and another 2 (mild and moderate PD) raised issues related to exercise. Other areas mentioned by one patient each included fatigue, social issues, assistive devices, and support groups. Patients' overall impression of the PDQ-39 received a median rating of 8 (range, 2-10).

Patients' comments provided additional information related to the acceptability of the questionnaire. Two patients (11%) with advanced PD found the questionnaire too long, which caused one of them problems concentrating throughout. Six patients (33%; representing all 3 severity groups) expressed difficulties in responding to items due to perceived ambiguity regarding what situation they refer to, whether use of assistive devices should be taken into account, or due to expression of more than one idea (e.g., item 29, "family or close friends"). Two patients (11%; moderate and advanced PD) expressed concerns regarding the acceptability of items 26-29 ("worried by others' reactions", "problems with close relationships", "support from spouse/partner" and "support from friends/family", respectively), which they found too personal. One patient with mild PD found some items redundant, e.g., items 27/28 ("problems with close relationships"/"support from spouse or partner").

Discussion

This study provides general support for the patient-perceived content validity of the PDQ-39, and supports the linguistic validity of its revised Swedish version. However, it also indicates that important health-related areas are lacking from the PDQ-39, and that its respondent burden should not be underestimated, particularly not among patients with more advanced disease.

Despite the recognized importance of the burden that PRO pose upon respondents (2-5), we have not been able to identify any previous documentation of this aspect of the PDQ-39. About 10 or up to 15 minutes have been suggested as a desirable time for self-completion in order for PRO to be acceptable for use in clinical trials and research (4, 20). Dunbar et al. (21) compared the required self-completion time for 4 generic and 3 condition-specific PRO in patients with knee arthoplasties. In this study respondents needed an average of 23 minutes to complete the Sickness Impact profile (SIP), whereas other PRO required about 8-12 minutes each (21). In agreement with the suggestions of Andresen (4) and McKenna (20), the authors considered the SIP to be the least suitable of the evaluated PRO (21). The present findings thus appear to suggest that the PDQ-39 has an acceptable level of respondent burden among patients with mild to moderate PD, whereas this is challenged in more advanced PD. Respondent burden is obviously related to the number of items. However, aspects such as questionnaire layout, complexity of item wording and the response system also contribute (2, 3, 22). Thus, while there may be reason to consider item reduction of the PDQ-39, the observed problems related to the response categories may also contribute to its respondent burden. Although not systematically recorded, a significant proportion of patients in this study were thus observed to take relatively long deciding on their responses. An alternative to circumvent undue respondent burden may be to use the 8-item PDQ-8 (23), a short form of the PDQ-39, instead. Items for the PDQ-8 were selected based on item-to-total correlations where the items with the strongest correlations to their respective subscales of the original UK PDQ-39 were selected (15, 23). However, firm evaluation of the PDQ-8 and its item selection is needed before confidence can be put into its use in various countries (24). For example, in its American (25) and Singapore Chinese (9) versions, 4 and 7 of the proposed PDQ-8 items, respectively, failed to meet the selection criterion used for the original UK PDQ-8.

No linguistic problems were identified in the revised Swedish PDQ-39, thus indicating that the attempts to rectify those encountered in the original Swedish version (8)

had been successful. However, although the wording of the response categories was not considered problematic by this group of patients, 44% experienced some problems regarding the use of the response scale. While it is recognized that further data are needed to allow firm conclusions, this may suggests an inherent problem with the use of a 5-grade, retrospective frequency related response scale in PD. Apart from the Swedish PDQ-39, we are only aware of two other adaptations, the American (25) and English Singapore (26) versions, that have involved PD patients to assess and/or produce linguistic validity. Interestingly, the problems identified when adapting the PDQ-39 for use in the US were very similar to those found with the original Swedish version. Thus, in both instances there was a problem with the response categories “occasionally” and “sometimes” as well as with items 28 and 29, and for both adaptations these problems were considered severe enough to call for revisions. Due to the apparent general lack of documented evaluations of linguistic validity, it is unknown to what extent these or other shortcomings exist also in other adaptations of the PDQ-39.

The attempt to overcome patients’ tendency to overlook the frame question when responding to items was not completely successful as a few patients still tended to forget to relate to the frame question. A simple solution to this could be to omit the frame question altogether. However, whereas some observations indicate that specification of the origin of experienced symptoms does not influence responses (27), potential attribution effects should not be ignored. Recent studies have thus reported influences on health questionnaire responses depending on whether items are asked with specific attributions or not, although the direction of this influence has differed (28, 29).

Patients’ ratings of their overall impression of the PDQ-39 as a measure of health, functioning and well-being in PD indicates acceptable content validity, which is in keeping with previous experiences (8, 9). However, a third of the patients expressed reservations regarding the relevance of various aspects of the questionnaire, and 50% identified important health-related aspects that were lacking. In accordance with previous findings (8), these related primarily to aspects of motor complications. Taken together, these observations point to room for improvement of the content validity of the PDQ-39. While other PD-specific health status questionnaires do include items that tap aspects of, e.g., motor complications, these largely lack other important areas, such as self care and daily activities (7). From this perspective, it thus seems unlikely that these would be superior to the PDQ-39, although empirical head-to-head comparisons will be needed to firmly evaluate this.

In conclusion, this study offers important implications for the development towards evidence based health outcome measurement in PD and illustrates that available PRO

should not be applied uncritically. The fact that an instrument has become widely used is not, in itself, a guarantee for its quality or appropriateness. First, the perceived overall content validity of the PDQ-39 as a PD-specific health status questionnaire is acceptable but compromised by lack of important content areas. This implies that additional items or modules may need to be considered to optimize content validity of the PDQ-39, particularly among patients experiencing motor complications. Second, item reduction and/or revision of the response scale may be necessary to improve its feasibility in clinical research and practice. Such developments need to be coupled by thorough qualitative and quantitative evaluations. Finally, this study illustrates that systematic small-scale qualitative evaluations with representative respondents can detect and successfully guide revision of linguistic problems. Such evaluations, coupled with psychometric assessments, are particularly important in the case of disease-specific PRO and should be employed and documented before large-scale implementation of new language adaptations. A quantitative evaluation of the psychometric properties of the revised Swedish PDQ-39 is underway.

Acknowledgements

This study was supported by Rådet för Hälso- och Sjukvårdsforskning (HSF), the Swedish Research Council, the Skane County Research and Development Foundation, and the Department of Nursing, Lund University.

We thank the patients for their cooperation, and Christine Nilsson RN, Department of Neurology, Lund University Hospital, for assistance with patient recruitment.

References

1. Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. 2nd ed. Oxford: Oxford University Press, 1995.
2. Doward LC, Meads DM, Thorsen H. Requirements for Quality of Life Instruments in Clinical Research. Value Health 2004;**7**(Suppl. 1):S13-S16.
3. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998;**2**:1-74.
4. Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;**81**(Suppl. 2):S15-S20.
5. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002;**11**:193-205.
6. Peto V, Jenkinson C, Fitzpatrick R, Greenhall R. The development and validation of a short measure of functioning and well-being for individuals with Parkinson´s disease. Qual Life Res 1995;**4**:241-248.
7. Marinus J, Ramaker C, van Hilten JJ, Stiggelbout AM. Health related quality of life in Parkinson's disease: A systematic review of disease specific instruments. J Neurol Neurosurg Psychiatry 2002;**72**:241-248.
8. Hagell P, McKenna SP. International use of health status questionnaires in Parkinson's disease: Translation is not enough. Parkinsonism Relat Disord 2003;**10**:89-92.
9. Luo N, Tan LCS, Li SC, Soh LK, Thumboo J. Validity and reliability of the Chinese (Singapore) version of the Parkinson's Disease Questionnaire (PDQ-39). Qual Life Res 2005;**14**:273-279.
10. Cella DF, Lloyd SR, Wright BD. Cross-cultural instrument equating: Current research and future directions. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials, 2nd ed. Philadelphia: Lippincott-Raven Publishers, 1996;707-715.
11. McKenna SP, Whalley D. Coping with language barriers. Good Clinical Practice Journal 1997;**4**:14-17.
12. Leplège A, Verdier A. The adaptation of health status measures: Methodological aspects of the translation procedure. In: Shumaker S, Berzon R, editors. International use and performance of health-related quality of life instruments. Oxford: Oxford Rapid Communication, 1995;93-101.

13. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 health survey for use in Japan. J Clin Epidemiol 1998;**51**:1037-1044.
14. Hagell P, Whalley D, McKenna SP, Lindvall O. Health status measurement in Parkinson's disease: Validity of the PDQ-39 and Nottingham Health Profile. Mov Disord 2003;**18**:773-783.
15. Jenkinson C, Fitzpatrick R, Peto V. The Parkinson's disease questionnaire. User manual for the PDQ-39, PDQ-8 and PDQ Summary Index. Oxford: Health Services Research Unit, Department of Public Health, University of Oxford, 1998.
16. Fayers PM, Machin D. Quality of life: Assessment, analysis and interpretation. West Sussex: John Wiley & Sons, Ltd., 2000.
17. Sprangers MAG, Cull A, Bjordal K, Groenvold M, Aaronson NK, for the EORTC Study Group on Quality of Life. The European Organization for Research and Treatment of Cancer approach to quality of life assessment: Guidelines for developing questionnaire modules. Qual Life Res 1993;**2**:287-295.
18. Hoehn MM, Yahr M. Parkinsonism: Onset, progression and mortality. Neurology 1967;**17**:427-442.
19. Gibb WRG, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988;**51**:745-752.
20. McKenna SP. Measuring quality of life in schizophrenia. Eur Psychiatry 1997;**12**(Suppl 3):267s-274s.
21. Dunbar MJ, Robertsson O, Ryd L. Non-psychometric properties of seven outcome questionnaires as applied to 3600 patients from the national Swedish knee arthroplasty registry. Transactions of the Annual Meeting of the Orthopaedic Research Society Annual Meeting 1999;**24**:150.
22. Saris WE, Gallhofer IN, van der Veld W. A scientific method for questionnaire design: SQP. Amsterdam: University of Amsterdam, 2003.
23. Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The PDQ-8: Development and validation of a short-form Parkinson's disease questionnaire. Psychol Health 1997;**12**:805-814.
24. Gandek B, Ware JE, Aaronson NK, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. J Clin Epidemiol 1998;**51**:1171-1178.

25. Bushnell DM, Martin ML. Quality of life and Parkinson's disease: Translation and validation of the US Parkinson's Disease Questionnaire (PDQ-39). Qual Life Res 1999;**8**:345-350.
26. Tan LCS, Luo N, Nazri M, Li SC, Thumboo J. Validity and reliability of the PDQ-39 and the PDQ-8 in English-speaking Parkinson's disease patients in Singapore. Parkinsonism Relat Disord 2004;**10**:493-499.
27. Markum RA. Assessment of the reliability of and the effect of neutral instructions on the symptom ratings on the Moos Menstrual Distress Questionnaire. Psychosom Med 1976;**38**:163-172.
28. Dubuc N, Haley SM, Kooyoomjian JT, Jette AM. Assessing disability in older adults: the effects of asking questions with and without health attribution. J Rehabil Med 2004;**36**:226-231.
29. Marx RG, Hogg-Johnson S, Hudak P, et al. A comparison of patients' responses about their disability with and without attribution to their affected area. J Clin Epidemiol 2001;**54**:580-586.

Table 1 Patient characteristics

	Participating patients (n=18)	Drop-outs (n=4)
Gender (M/F)	11/7 [a]	2/2 [a,b]
Age (years)	67.9 (10.9; 36-81) [c]	69.5 (10.9; 56-81) [b]
HY stage	III (II, IV) [d]	II (II, III) [b,d]
I	2 [a]	0 [a]
II	2 [a]	3 [a]
III	6 [a]	1 [a]
IV	4 [a]	0 [a]
V	4 [a]	0 [a]

[a] Number of patients.

[b] No significant differences between participating patients and drop-outs (Fischer's exact test, t-test and Mann-Whitney U-test, respectively).

[c] Mean (SD; min – max)

[d] Median (q1, q3)

M, male; F, female; HY, Hoehn & Yahr stage of Parkinson's disease.

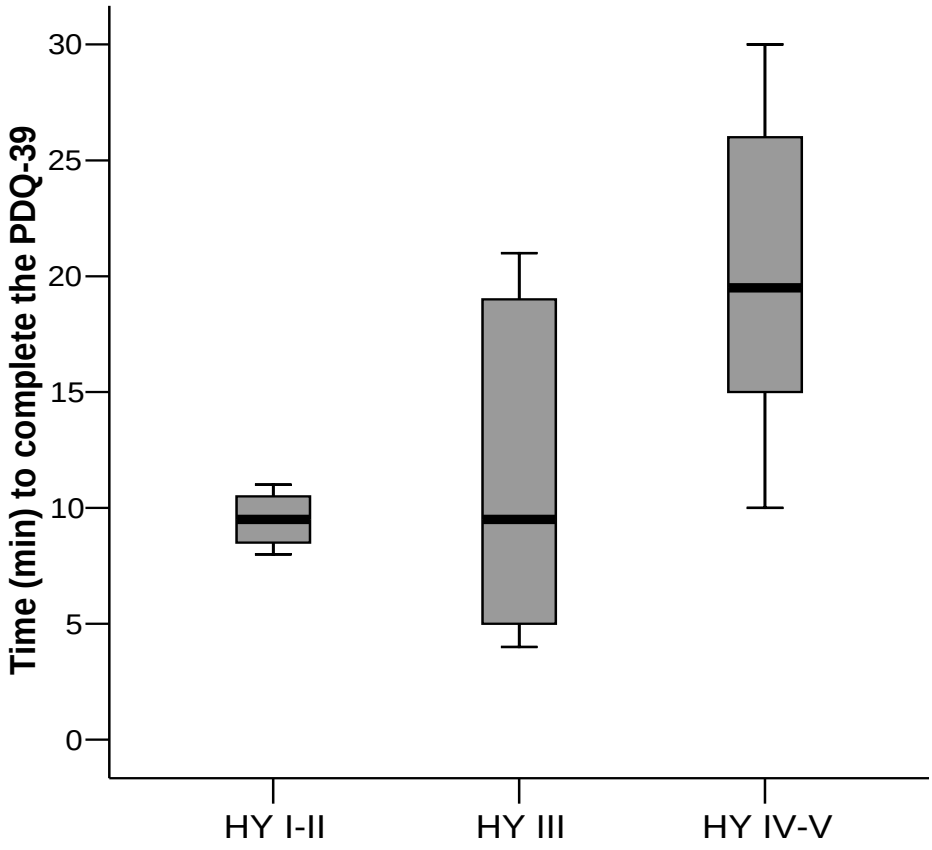
Legend to Figure

Fig. 1:

Distribution of the time needed to complete the PDQ-39 by HY stages grouped as mild (HY I-II; n=4), moderate (HY III; n=6) and advanced (HY IV-V; n=8) PD. Solid horizontal lines are median values, boxes are inter-quartile ranges, error bars are ranges.

HY, Hoehn & Yahr stage of Parkinson's disease; PDQ-39, the 39-item Parkinson's Disease Questionnaire.

Fig. 1


Time (min) to complete the PDQ-39
30
25
20
15
10
5
0
HY I-II
HY III
HY IV-V