Validated of 19-items wearing-off (WOQ-19) questionnaire to Portuguese

Carlos Eduardo MANTÈSE1,2, Marcio MEDEIROS1, Artur SCHUMACHER-SCHUH1,3, Carlos Roberto de Melo RIEDER1,4

ABSTRACT
Background: The treatment of Parkinson disease with dopaminergic therapy improves functionality and quality of life. However, as the disease progresses, the wearing-off phenomenon develops. To improve the recognition of this phenomenon, the 19-item wearing-off questionnaire (WOQ-19) was developed. Objective: To translate and validate the WOQ-19 into Portuguese. Methods: The questionnaire was translated into Portuguese and, subsequently, back-translated into English and analyzed. The final version was tested in Parkinson disease patients for reliability through the test-retest paradigm and internal consistency. Also, sensitivity and specificity were obtained in different cut-off positive items. Results: The WOQ-19 showed good test stability, with an intraclass correlation coefficient of 0.877 (95%CI 0.690–0.951; p<0.001), and good internal consistency, with Cronbach alpha of 0.815. Two items of positive cut-off showed the best accuracy: 0.873 (95%CI 0.791–0.954). Sensitivity was 0.975 (95%CI 0.892–1) and specificity was 0.714 (95%CI 0.565–0.863). Conclusion: The Portuguese version of the WOQ-19 showed excellent diagnostic properties and can be used to diagnose wearing-off phenomena.

Keywords: Parkinson Disease; Antiparkinson Agents.

INTRODUCTION

Parkinson disease (PD) is the second most prevalent neurodegenerative disorder, with no curative treatment. In the beginning, it can have a fair medical treatment with dopaminergic therapy, which improves functionality and quality of life1. As the disease progresses, some complications of the treatment can be very disturbing, such as wearing-off (WO) phenomenon. It is the shortening of the drug effects that leads to complex medication posology or adjuvant...
therapies. It is associated with the worsening of quality of life and functionality. Sometimes, it is not easily recognized, if the symptoms are mild or non-motor.

To improve the recognition of this phenomenon, a 32-item questionnaire was developed. It describes some symptoms and patients must check whether they have them and if they improve with medication. For practical reasons, using the same research, it was adapted to 19 items (19-item wearing-off questionnaire — WOQ-19), which should have the same properties. It has been used in clinical studies and has been validated in several languages with some different clinimetric properties. A recent review of the Movement Disorder Society set WOQ-19 as the recommended tool for wearing-off screening. Our group performed a systematic review which showed a sensitivity of 0.8–0.88 and specificity of 0.68–0.8, when 2 cut-off items were used, compared to the clinical evaluation, the gold-standard. This questionnaire was used in a sample of Brazilian patients, with similar properties. However, it was a plain translation, with no proper validation.

Thus, we aimed to validate the WOQ-19 in Portuguese and to access its clinimetric properties.

**METHODS**

The work has two parts, language, cultural, and social translation, following Beaton's guidelines, and clinimetric properties testing.

**Patient selection**

One-hundred-fifty-six PD patients were selected from the movement disorder unit in Hospital de Clínicas de Porto Alegre, a public teaching hospital, in Porto Alegre, Brazil. Inclusion criteria were: Patients had to be diagnosed with PD by UK Brain Bank criteria and had to be on dopaminergic therapy. Also, the patient or caregivers had to be able to fill it out. Epidemiological and clinical data were collected, with a neurologic evaluation and the Mini-Mental State Examination and the Movement Disorder Society Unified Parkinson’s Disease rating scale (MDS-UPDRS). Levodopa equivalent dose was calculated as described by Tomlinson. All patients were examined by one of the two authors, CEM or MM, both neurologists with a special interest in Movement Disorders, in ON state, both blinded to the WOQ-19 results.

Twenty patients were evaluated by test-retest paradigm 2 weeks apart, through intraclass correlation coefficient (ICC) and kappa, to assess reliability. We used the Cronbach’s alpha for internal consistency in those patients, expecting a value greater than 0.75.

For sensibility and specificity the questionnaire results, were compared to the clinical evaluation as a gold-standard.

**Statistical analysis**

A descriptive analysis of all variables was conducted, expressed as mean and standard deviation (SD) or median and interquartile range (IQR). WOQ-19 results were compared to the WO clinical assessment of positive and negative patients. Also, the number of positive cut-off items for better sensibility and specificity was analyzed. Moreover, we compared the difference between patients with positive and negative questionnaires. We calculated the sensitivity, specificity, positive predictive value, negative predictive value and a receiver operating characteristic (ROC) curve with area under the curve (AUC). A post-hoc analysis was conducted, comparing these variables in a group of 20 patients with shorter disease duration. Internal consistency was tested with Cronbach’s alpha. The Shapiro-Wilk normality test was applied to determine the homogeneity of variance. Student's t-test was used for parametric and Mann-Whitney U for non-parametric variables, chi-square for proportion. For statistical analysis, the SPSS 18 package was used. This study followed the STARD guideline.

**Ethical aspects**

The project was approved by the Hospital de Clínicas Ethics Committee. The project was approved by Duke University and Dr. Stacy, developer of the questionnaire.
Informed consent was obtained from all patients or their legal guardians.

RESULTS

Language, cultural, and social translation
After initial translation and back-translation, the item “Numbness” presented with discordance. The initial form *Dormência,* came back wrong. To clarify this issue, we consulted the Italian (*Sensazione di addormentamento ad una parte del corpo*) and Spanish (*entumecimiento, hormigueos*) versions, languages more similar to Portuguese, and the item was settled as *Dormência em parte do corpo.* All 20 initial patients seem to understand the questionnaire.

Clinimetric properties
One hundred fifty-six PD patients were included. Demographics and clinical data are described in Table 1. Mean age was 64.7 (±10.2) years, 51% were male, and mean disease duration was 12.4 (±5.3) years. All patients were on levodopa therapy. Total MDS-UPDRS was 78.22 (±35.3). The median of the positive answers was 5 (3–8) for the entire questionnaire, 4 (2–6) for motor questions and 1 (0–2) for non-motor ones.

The test-retest paradigm was applied to 20 patients. The ICC, when considering the number of positive items, was 0.877 (95%CI 0.791–0.954) (Figure 1).

The questionnaire turned out in 128 WO-positive patients and 35 WO-negative ones. Clinical evaluation showed 121 WO-positive patients and 35 WO-negative ones. The median positive answer in clinical WO-positive patients was 6 (4–9) and clinical WO-negative patients, 1 (0–3); p<0.001.

Table 2 shows the sensitivity and specificity by number of positive items; 2 positive items have the best accuracy. Sensitivity was 0.975 (95%CI 0.947–1), specificity was 0.714 (0.565–0.863), positive predictive value (PPV) was 0.921 (95%CI 0.875–0.968), and negative predictive value was 0.892 (95%CI 0.778–1). A ROC curve was plotted with an AUC of 0.875 (95%CI 0.873–0.968) (Figure 1).

A sample of 20 patients with shorter disease duration (4,93±1.36 years) analyzed with 2 positive cutoff points. Sensitivity was 1, specificity was 0.714.

### Table 1. Epidemiological data.

|                          | Total n=156 (100%) | WO by clinical assessment n=121 (77.6%) | No WO by clinical assessment n=35 (22.4%) |
|--------------------------|--------------------|----------------------------------------|-----------------------------------------|
| Age: mean (±SD), years   | 64.75 (10.21)      | 63.5 (10.34)                           | 68.78 (8.73)                            |
| Male: n (%)              | 80 (51.3%)         | 66 (54.5%)                             | 14 (40%)                                |
| Age at disease onset: mean (±SD), years | 52.28 (10.36) | 50.98 (10.42)                          | 56.75 (8.91)                            |
| Disease duration: mean (±SD), years | 12.47 (5.39) | 12.60 (5.40)                           | 12.02 (5.4)                             |
| LVD equivalent: mean (±SD), mg | 1,188.02 (505.46) | 1,236.31 (484.75) | 1,021.09 (546.26)                      |
| MDS-UPDRS                |                    |                                        |                                         |
| Part 1: mean (±SD)       | 13.03 (7.88)       | 12.75 (7.71)                           | 13.97 (8.45)                            |
| Part 2: mean (±SD)       | 19.49 (11.19)      | 19.33 (10.26)                          | 20.08 (14.14)                           |
| Part 3: mean (±SD)       | 39.06 (20.82)      | 37.99 (20.07)                          | 42.74 (23.17)                           |
| Part 4: mean (±SD)       | 6.64 (4.12)        | 7.86 (3.38)                            | 2.43 (3.76)                             |
| Total: mean (±SD)        | 78.22 (35.30)      | 77.93 (33.18)                          | 79.2 (42.37)                            |
| Questionnaire 2 items: n (%) | 128 (82%)         | 118 (97.5%)                            | 10 (28.5%)                              |
| Items: median (IQR)      | 5 (3–8)            | 6 (4–9)                                | 1 (0–3)                                 |
| Motor: median (IQR)      | 4 (2–6)            | 5 (3–7)                                | 1 (0–2)                                 |
| Non motor: median (IQR)  | 1 (0–2)            | 1 (0–3)                                | 0 (0–1)                                 |

LVD: levodopa; MDS-UPDRS: Movement Disorder Society Unified Parkinson’s Disease rating scale; SD: standard deviation; IQR: interquartile range. Student’s t-test was used for parametric and Mann-Whitney U for non-parametric variables.

### Table 2. Items cut-off.

| Number positive items | Sn       | Sp       |
|-----------------------|----------|----------|
| 0                     | 1        | 0        |
| 1                     | 0.983    | 0.457    |
| 2                     | 0.975    | 0.714    |
| 3                     | 0.901    | 0.743    |
| 4                     | 0.810    | 0.80     |
| 5                     | 0.678    | 0.857    |
| 6                     | 0.545    | 0.857    |

Sn: sensitivity; Sp: specificity.
The Cronbach’s alpha showed good internal consistency. To our knowledge, this is the first study to apply kappa for categorical correlation, resulting in a moderate correlation. When applying the ROC curve, two positive items seemed to have the best accuracy. The same result was demonstrated by other authors. In addition, the 9-item questionnaire, which is a shortened version of this one, showed that the questionnaire loses specificity when a positive item is used as a cutoff point.

This study has some limitations, as the specificity interval confidence was wide, partly because we have few patients without WO. This might happen because WO is associated with longer disease duration, and our patients had a long disease compared to others who validated this questionnaire in other languages. Also, our sample did not have many patients at the onset of the disease, we had only 2 patients with Hoehn and Yahr less than 2. It might be because our sample comes from a tertiary teaching hospital. This could make the symptoms milder and harder to identify. To mitigate this bias, we analyzed a sample of patients with 20 patients with 6 or fewer years of disease. It does not have the statistic power, however it showed similar results to those of the entire group. This approach must lessen this bias. However, WO was once thought to be a later complication, though it can be present since the onset of the disease. Stocchi, using the Italian version of the same questionnaire, reported as early as 2.5 years of disease duration for 41.8% of patients presented WO. As having few patients early in the course of the disease, our validation loses power for this subset of patients. We do not report the time with levodopa use, because there was an important recall bias. Since WO is associated with levodopa use, the amount of time use could indicate a risk to develop WO. However, since all patients were on levodopa, we do not believe it could influence the diagnostic properties.

Another potential bias was that the questionnaire was explained to patients before they filled it out. There might be a difference among other groups, as it is not reported how the questionnaire was applied. Bares, when applying the shorter version of the questionnaire, stated it can have some misunderstanding about the questionnaire, so it was recommended it was explained to patients.

In conclusion, we validated the Portuguese version of the WOQ-19. The Brazilian version of the questionnaire demonstrated to be reliable and valid and can be considered a tool to diagnose wearing-off. Some caution must be taken when applying it to patients at very early stages of the disease.

### DISCUSSION

WOQ-19 is an important tool for clinical and research evaluation on PD. However, a proper validation for Portuguese was missing. Our linguistic validation followed Beaton’s guidelines, and there have been no deviations. Since there was a divergence in an item during the procedure, the need for proper validation might be reinforced when a scale or questionnaire built in other cultural or language setting is used. In the era of multicentric trials, cultural and language barriers can be overcome by proper validation tools. The Movement Disorder Society has been doing this with its scales.

The questionnaire showed to be reliable, with excellent ICC, similar to the Italian version. Also, kappa was used for categorical correlation, resulting in a moderate correlation. To our knowledge, this is the first study to apply kappa. The Cronbach’s alpha showed good internal consistency.

As expected, the WOQ-19 showed a statistical difference in relation to MDS-UPDRS part IV, which measures levodopa’s complications and equivalent dose, which is associated with WO. This reinforces the construct validation. When plotting the ROC curve, two positive items seemed to have the best accuracy. The same result was demonstrated by other authors. In addition, the 9-item questionnaire, which is a shortened version of this one, showed that the questionnaire loses specificity when a positive item is used as a cutoff point.

### Table 3. Discriminatory ability.

|                      | WOQ-19 (2 or more positive items) | WOQ-19 (1 or less positive items) |
|----------------------|-----------------------------------|-----------------------------------|
| Equivalent dose:     | 1,227.66 (475.76)                 | 1,006.83 (600.45)                 |
|                      | p=0.036                           | p=0.035                           |
| MDS-UPDRS 1:         | 12.41 (7.61)                      | 15.86 (8.56)                      |
| mean (±SD)           |                                   |                                   |
| MDS-UPDRS 2:         | 18.84 (10.20)                     | 22.46 (14.80)                     |
| mean (±SD)           |                                   |                                   |
| MDS-UPDRS 3:         | 36.84 (19.19)                     | 49.18 (25.07)                     |
| mean (±SD)           |                                   |                                   |
| MDS-UPDRS 4:         | 7.51 (3.65)                       | 2.68 (3.88)                       |
| mean (±SD)           |                                   | p<0.001                           |
| total: mean (±SD)    | 75.6 (32.38)                      | 90.18 (45.18)                     |

WOQ-19: 19-item wearing-off questionnaire; MDS-UPDRS: Movement Disorder Society Unified Parkinson’s Disease rating scale; SD: standard deviation. Student’s t-test was used for parametric variables.

### ACKNOWLEDGMENTS

CRM Rieder was supported by the Brazilian government research grant agency — Conselho Nacional de Desenvolvimento Científico e Tecnológico [CNPq] (No. 311970/2014-1). We thank Prof. Stacy and Duke University for granting permission for translation and support.
References

1. Aquino CC and Fox SH. Clinical spectrum of levodopa-induced complications. Mov Disord. 2015 Jan;30(1):80-9. https://doi.org/10.1002/mds.26126

2. Pahwa R, Lyons KE. Levodopa-related wearing-off in Parkinson's disease: identification and management. Curr Med Res Opin. 2009 Apr;25(4):841-9. https://doi.org/10.1186/0300902779319

3. Stacy M, Bowron A, Guttman M, Hauser R, Hughes K, Larsen JP, et al. Identification of motor and nonmotor wearing-off in Parkinson's disease: comparison of a patient questionnaire versus a clinician assessment. Mov Disord. 2005 Jun;20(6):726-33. https://doi.org/10.1002/mds.20383

4. Stacy M, Hauser R. Development of a patient questionnaire to facilitate recognition of motor and non-motor wearing-off in Parkinson's disease. J Neural Transm (Vienna). 2007 Feb;114(2):211-7. https://doi.org/10.1007/s00702-006-0554-y

5. Stacy M. The wearing-off phenomenon and the use of questionnaires to facilitate its recognition in Parkinson's disease. J Neural Transm (Vienna). 2010;117(7):837-46. https://doi.org/10.1007/s00702-010-0424-5

6. Antonini A, Martinez-Martín P, Chaudhuri RK, Merello M, Hauser R, Katzenschlager R, et al. Wearing-Off scales in Parkinson's disease: critique and recommendations. Mov Disord. 2011 Oct;26(12):2169-75. https://doi.org/10.1002/mds.23975

7. Mantese CE, Schumacher-Schuh A, Rieder CRM. Clinimetrics of the 9- and 19-item wearing-off questionnaire: a systematic review. Parkinosa Dis. 2018 Apr;2018:Article ID S308491.1-5. https://doi.org/10.1158/S0042-282X201000400007

8. Melo LM, Chien HF, Barbosa ER. Identification of wearing-off manifestations (reduction of levodopa effect) in Parkinson's disease using specific questionnaire and comparison of the results with routine ambulatory evaluations. Arq Neuro-Psiquiatr. 2010;68(4):506-10. https://doi.org/10.1590/S0004-282X2010000400007

9. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000;25(24):3186-91. https://doi.org/10.1097/00007632-200012150-00014

10. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical examination of Parkinson's disease: a clinical-pathological study of 100 cases. J Neurol Neurosurg Psychiatry. 1992;55(3):181-4. https://doi.org/10.1136/jnnp.55.3.181

11. Abbruzzese G, Antonini A, Barone P, Stocchi F, Tamburini T, Bernardi L, et al. Linguistic, psychometric validation and diagnostic ability assessment of an Italian version of a 19-item wearing-off questionnaire for wearing-off detection in Parkinson's disease. Neurol Sci. 2012 Dec;33(6):1319-27. https://doi.org/10.1007/s10072-012-0943-y

12. Brucki SMD, Nittriu R, Caramelli P, Bertolucci PHF, Okamoto H. Suggestions for utilization of the mini-mental state examination in Brazil. Arq Neuropsiquiatr. 2003 Sep;61(3B):777-81. https://doi.org/10.1590/s0004-282x2003000500014

13. Goetz CG, Tilley BC, Shaftman SR, Stebbings GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord. 2008 Nov;23(15):2129-70. https://doi.org/10.1002/mds.22340

14. Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord. 2010 Nov;25(15):2649-85. https://doi.org/10.1002/mds.23429

15. Bossuyt PM, Reitsma JB, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015 Oct;351:351-9. https://doi.org/10.1136/bmj.h5627

16. Martinez-Martin P, Tolosa E, Hernandez B, Badia X. The Patient Card questionnaire to identify wearing-off in Parkinson disease. Clin Neuropharmacol. 2007 Sep-Oct;30(5):266-75. https://doi.org/10.1097/wnf.0b013e3180413da7

17. Tomlinson CL. Early DEtection of wEaring off in Parkinson disease: the DEEP study. Parkinsonism Relat Disord. 2014 Feb;20(2):204-11. https://doi.org/10.1016/j.parkreldis.2013.10.027

18. Martinez-Martin P, Tolosa E, Hernandez B, Badia X. Validation of the "QUICK" Questionnaire — a tool for diagnosis of “wearing-off” in patients with Parkinson's disease. Mov Disord. 2008 Apr;23(6):830-6. https://doi.org/10.1002/mds.21944

19. Stacy MA, Murphy JM, Greely DR, Stewart RM, Murch H, Meng X, et al. The sensitivity and specificity of the 9-item Wearing-off Questionnaire. Parkinsonism Relat Disord. 2008;14(3):205-12. https://doi.org/10.1016/j.parkreldis.2007.07.013

20. Hauser R, McDermott MP, Messing S. Factors associated with the development of motor fluctuations and dyskinesias in Parkinson disease. Arch Neurol. 2006 Dec;63(12):1756-60. https://doi.org/10.1001/archneur.63.12.1756

21. Seki M, Takahashi K, Uematsu D, Mihara B, Morita Y, Isozumi K, et al. Clinical features and varieties of non-motor fluctuations in Parkinson's disease: A Japanese multicenter study. Parkinsonism Relat Disord. 2013 Jan;19(1):104-8. https://doi.org/10.1016/j.parkreldis.2012.08.004

22. Stocchi F, Antonini A, Barone P, Tinazzi M, Zappa M, Onofrj M, et al. Early DEtection of wEaring off in Parkinson disease: the DEEP study. Parkinsonism Relat Disord. 2014 Feb;20(2):204-11. https://doi.org/10.1016/j.parkreldis.2013.10.027

23. Bareš M, Rektorová I, Jech R, Farníková K, Roth J, Růžička E, et al. Does W0Q-9 help to recognize symptoms of non-motor wearing-off in Parkinson's disease? J Neural Transm (Vienna). 2012 Mar;119(3):373-80. https://doi.org/10.1007/s00702-011-0683-9