Article published online: 2023-06-13

https://doi.org/10.1590/0004-282X20200125

ARTICLE

# Validation of the Brazilian Portuguese version of the Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire (RBDSQ-BR)

## Validação da Versão Brasileira do Questionário de Triagem do Transtorno Comportamental do Sono REM (QT-TCSREM-BR)

*Márcio Alexandre PENA-PEREIRA[1], Manoel Alves SOBREIRA-NETO[1,2], Emmanuelle SOBREIRA[1], Marcos Hortes Nisihara CHAGAS[1,3], Daniel Sabino de OLIVEIRA[1], Guilherme Riccioppo RODRIGUES[1], Carolina Pinto de SOUZA[1], Alan Luiz ECKELI[1], Regina Maria França FERNANDES[1], Vitor TUMAS[1]*

**ABSTRACT**

**Introduction:** A diagnosis of rapid eye movement sleep behavior disorder (RBD) currently requires confirmation with polysomnography (PSG). However, PSG may not be sufficiently available. In these situations, a clinical diagnostic measure might be useful. **Objective:** To validate the Brazilian Portuguese version of RBD screening questionnaire (RBDSQ) for patients with Parkinson's disease (PD). **Methods:** Using detailed clinical interviews and PSG analysis (diagnostic gold standard), a convenience sample of 69 subjects was divided into the following subgroups: patients with PD and RBD (PD+RBD; n=50) and patients with PD alone (PD-RBD; n=19). **Results:** RBDSQ-BR showed adequate internal consistency (Cronbach's $\alpha$=0.809) and, except for item 8, adequate item-test correlation. The retest performed in a second sample (n=13, consecutive) showed high agreement for total score (intraclass correlation coefficient, ICC=0.863) and acceptable agreement for items 2, 3, 6.2, 6.3, 7, and 8 (K>0.60). The receiver operating characteristic (ROC) curve analysis had an area under the curve (AUC) of 0.728. A cut-off score of 4 enabled the correct diagnosis of 76.8% subjects and provided the best balance between sensitivity (84%) and specificity (57.9%), with a 2.0 likelihood ratio of a positive result (LR+) and a 0.3 likelihood ratio of a negative result (LR-). Items 2 and 6.2 had 84.2% specificity and 3.2 LR+. Combined items 1+2+6.2, 2+6.1, and 6.1+6.2 increased the specificity to 94.7%, with LR+ ranging from 6.1 to 7.6. **Conclusions:** RBDSQ-BR is a reliable instrument, which may be useful for RBD diagnosis of Brazilian patients with PD. The instrument is also valid and may help in a better selection of cases for a more detailed clinical evaluation or even PSG analysis.

**Keywords:** Parkinson's Disease; REM Sleep Behavior Disorder; Brazil; Polysomnography; Surveys and Questionnaires.

**RESUMO**

**Introdução:** O diagnóstico do transtorno comportamental do sono REM (TCSREM) implica na realização da polissonografia (PSG), mas sua disponibilidade pode não ser suficiente. Portanto, meios clínicos para o diagnóstico podem ser úteis. **Objetivo:** Validar para a língua portuguesa falada no Brasil o questionário de triagem do TCSREM (QT-TCSREM) em pacientes portadores de doença de Parkinson (DP). **Métodos:** Uma amostra por conveniência composta de 69 indivíduos foi dividida em portadores de DP com TCSREM (n=50) e DP sem TCSREM (n=19) através de entrevista clínica detalhada e análise da PSG. **Resultados:** QT-TCSREM-BR apresentou consistência interna adequada ($\alpha$ de Cronbach=0,809) e, exceto pelo item 8, correlação item-total adequada. Reteste feito em uma segunda amostra (n=13, consecutivos) evidenciou concordância elevada para o escore total (coeficiente de correlação intraclasse, CCI=0,863) e aceitável para os itens 2, 3, 6.2, 6.3, 7 e 8 (K>0,60). Análise da curva característica de operação do receptor (COR) obteve uma área sob a curva de 0,728. O corte 4 permitiu o diagnóstico correto de 76,8% dos indivíduos e apresentou o melhor equilíbrio entre sensibilidade (84%) e especificidade (57,9%), com uma razão de verossimilhança de um resultado positivo (RV+) 2,0 e de um resultado negativo (RV-) 0,3. Os itens 2 e 6.2 obtiveram especificidade

[1]Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Departamento de Neurociências e Ciências do Comportamento, Ribeirão Preto, São Paulo SP, Brazil.

[2]Universidade Federal do Ceará, Faculdade de Medicina, Fortaleza CE, Brazil.

[3]Universidade Federal de São Carlos, Centro de Ciências Biológicas e da Saúde, São Carlos SP, Brazil.

Márcio Alexandre PENA-PEREIRA https://orcid.org/0000-0001-6468-3582; Manoel Alves SOBREIRA-NETO https://orcid.org/0000-0001-9619-6551; Emmanuelle SOBREIRA https://orcid.org/0000-0002-6017-5107; Marcos Hortes Nisihara CHAGAS https://orcid.org/0000-0003-3752-7984; Daniel Sabino de OLIVEIRA https://orcid.org/0000-0001-6287-4756; Guilherme Riccioppo RODRIGUES https://orcid.org/0000-0003-1475-1908; Carolina Pinto de SOUZA https://orcid.org/0000-0002-0724-5276; Alan Luiz ECKELI https://orcid.org/0000-0001-5691-7158; Regina Maria França FERNANDES https://orcid.org/0000-0003-0436-1236; Vitor TUMAS https://orcid.org/0000-0003-2402-2709

**Correspondence:** Vitor Tumas; E-mail: tumasv@fmrp.usp.br

**Conflicts of interest:** There is no conflict of interest to declare.

**Authors' contributions:** MAPP, ALE, RMFF and VT conceived the original idea. MAPP, MASN, DSO, GRR, CPS and ALE participated in the cultural adaptation of the questionnaire. MAPP, MASN, MAPP, MASN, ES and MHNC collected the data. MAPP, MASN, ALE and VT analyzed the data. MAPP wrote the paper. All the authors revised the work critically and approved the final version of the paper.

**Ethical standards:** All participants provided a signed informed consent form, which was previously approved by the local ethics committee (protocols 2213/2009 and 13410/2009).

Received on February 25, 2020; Received in its final form on April 16, 2020; Accepted on April 28, 2020.

84,2% e RV+ 3,2. Itens combinados 1+2+6,2, 2+6,1 e 6,1+6,2 aumentaram a especificidade para 94,7%, com RV+ variando de 6,1 até 7,6. **Conclusões:** O QT-TCSREM-BR é um instrumento confiável que pode ser útil para o diagnóstico do TCSREM em pacientes com DP no Brasil. O instrumento também é válido e pode auxiliar numa melhor seleção de casos a serem submetidos a uma avaliação mais detalhada ou até mesmo a uma análise de PSG.

**Palavras-chave:** Doença de Parkinson; Transtorno do Comportamento do Sono REM; Brasil; Polissonografia; Inquéritos e Questionários.

## INTRODUCTION

Rapid eye movement (REM) sleep behavior disorder (RBD) is characterized by the loss of atonia during REM sleep, due to excessive motor activity or even dream performance[1,2]. The prevalence of RBD among the general population is approximately 1.8%[3], whereas it may reach 55.7% in patients with Parkinson disease (PD)[4].

A diagnosis of RBD currently requires confirmation with polysomnography (PSG)[2,5]. However, PSG may not be sufficiently available, especially in the public health system[6]. In these situations, a clinical diagnostic measure might be useful[7,8,9,10,11]. The RBD screening questionnaire (RBDSQ) is a 13-item, self-administered, dichotomous instrument, which also allows for the input from the patient's companion. It was originally published in German and English[7] and subsequently validated in Japan[12], China[13], South Korea[14], Turkey[15], and Italy[16]. Its use has already been validated for patients with PD[17,18].

The objective of the present study was to validate the Brazilian Portuguese version of the RBDSQ (RBDSQ-BR) for patients with PD.

## METHOD

The current study had a cross-sectional, observational design. The sample of 82 patients with PD was selected from a cohort of subjects participating in a follow-up study in outpatients at the University Hospital of Ribeirão Preto School of Medicine of *Universidade de São Paulo*. The following inclusion criteria used were: of Brazilian ethnicity, ≥18 years of age, diagnosed with PD according to the United Kingdom Brain Bank criteria[19] (allowing the presence of family history of PD). They also needed to be available for PSG and to provide verbal and written consent to participate in the study. The inclusion periods were from February 2010 to November 2011 and from July to September 2014. The study was approved by the Research Ethics Committee under protocol numbers 2213/2009 and 13410/2009.

### Questionnaire translation and cultural adaptation

The original author of the RBDSQ authorized the current study, which was performed according to previously published protocols[20,21]. The English questionnaire was initially translated by three bilingual Brazilian natives separately (*i.e.*, one physician experienced in sleep disorders, one physician experienced in movement disorders, and one engineer). The three versions were then analyzed by a committee comprising bilingual physicians (three physicians experienced in movement disorders, one physician experienced in sleep disorders, and one physician experienced in both disorders, who were informed about the objective of the study, but were not involved in the translation) who prepared a single version. This version was administered to 10 subjects (patients or their companions) as a pre-test. The committee then conducted a revision of this version (RBDSQ-BR, see Online resource) toward identifying any inconsistencies.

Subsequently, independent back-translation was performed by two native English teachers (school teachers). These versions were reviewed by a committee of three professionals (one experienced in movement disorders, one experienced in sleep disorders, and one experienced in both), preparing a single, back-translated version that was sent to the original author along with its Brazilian version for review and consent.

### *Clinical evaluation and polysomnography*

Of the 82 patients, 69 patients with clinical and PSG data were selected for analysis. The remaining 13 consecutive patients only had the RBDSQ-BR results available for analysis.

The evaluation started with patients filling the RBDSQ-BR with the help of their companion, immediately after the PSG or on the following day. The patient was then subjected to a detailed sleep evaluation by an experienced physician who used the diagnostic criteria from the 3rd edition of the International Classification of Sleep Disorders[5] and was blinded to the results from the questionnaire and PSG. This detailed evaluation occurred within 30 days of the PSG. Evaluation of PD was performed by another experienced physician who was blinded to the questionnaire, sleep evaluation, and PSG results, using the Unified Parkinson's disease rating scale[22], the Hoehn and Yahr Staging Scale (HY)[23], and the Schwab & England Functional Scale (S&E)[24]. All evaluations were performed with the patients in the *on state*.

PSG was performed using a digital polygraph (Biologic Sleepscan VISION PSG, Natus Bio-logic Systems Inc., San Carlos, CA) using an extended 10–20 system electroencephalogram (EEG), electro-oculogram, surface electromyogram (chin, masseter muscles, finger extensors, and tibialis anterior muscle), nasal-cannula pressure transducer, thermocouple nasal/oral airflow sensor chest and abdominal

respiratory inductive plethysmography band transducers, peripheral oximetry, electrocardiogram (ECG), snore and body position sensors, and synched audio/video. The analysis was performed by two experienced physicians blinded to the RBDSQ-BR results. All of the technical parameters used were performed in accordance with the AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Technical Specification[25].

### Statistical analysis

Categorical variables were assessed using Fisher's exact test, while quantitative variables were evaluated using Student's t-tests or Mann-Whitney tests to compare the groups of patients with PD, with (PD+RBD), and without (PD-RBD) RBD. Normality assessment was performed using the Shapiro-Wilk test[26].

RBDSQ-BR reliability was assessed by determining both the temporal stability (re-test) and homogeneity (internal consistency) using Spearman's rank correlation coefficient (p) for the re-test (21 days of interval) and the intraclass correlation coefficient (ICC) for the total score (type III, two-factor, mixed analysis of variance [ANOVA])[27], and Cohen's kappa coefficient (K) for each item. Item-test correlation was used to estimate the coefficient discrimination of the items, accepting values >0.3. Internal consistency was determined using the Cronbach's α reliability estimate[27,28,29,30].

The validity of the RBDSQ-BR was assessed using receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) analysis, with confidence intervals determined using the exact binomial method[31]. Fisher's exact test was used to evaluate each item. Sensitivity, specificity, positive (PPV) and negative (NPV) predictive values, accuracy, and the likelihood ratios of positive (LR+) and negative (LR-) results were calculated for each cut-off score and item. Pre-test probability of RBD was 72.5% (50 PD+RBD subjects/69 total subjects)[32,33].

P<0.05 were considered significant, with test power >80%. Microsoft Office Excel®, IBM SPSS Statistics 19, and R 3.1.0 software were used for statistical analysis.

## RESULTS

### Clinical data

Table 1 outlines the clinical data of the 69 patients. Only the presence of companions during evaluation and dose-equivalent levodopa levels, which were higher in the PD+RBD group, significantly differed between the groups.

### Questionnaire reliability

RBDSQ-BR showed adequate internal consistency (0.809 overall Cronbach's α; 95%CI 0.74–0.89; n=69). Independent withdrawal of each item from the questionnaire

**Table 1.** Sample characteristics.

| | All n=69 | PD(-)RBD n=19 | PD(+)RBD n=50 | p-value[a] |
|---|---|---|---|---|
| Males, n (%) | 39 (56.5) | 10 (52.6) | 29 (58.0) | 0.79 |
| Age (years; mean, SD) | 60.4 (12.0) | 59.7 (15.1) | 60.7 (10.7) | 0.79 |
| PD duration (years; mean, SD) | 8.6 (5.4)[b] | 7.9 (6.4)[b] | 8.9 (5.1) | 0.57 |
| Patient education (years; mean, SD) | 6.1 (5.0)[b] | 5.5 (4.1)[b] | 6.3 (5.3) | 0.72 |
| Patients with companion, n (%) | 54 (79.4)[b] | 11 (61.1)[b] | 43 (86) | 0.04* |
| Companion is a roommate, n (%) | 13 (19.1)[b] | 2 (18.2)[b] | 11 (25.6) | 0.10 |
| Companion education (years; mean, SD) | 8.2 (4.9) | 8.5 (4.8) | 8.2 (5.0) | 0.59 |
| Part III UPDRS ("on"; mean, SD) | 16.3 (9.4) | 16.0 (7.9) | 16.4 (10.0) | 0.87 |
| Total UPDRS ("on"; mean, SD) | 33.3 (17.5) | 30.0 (14.2) | 34.5 (18.7) | 0.85 |
| HY ("on"; median, variance) | 2 (0–3) | 2 (1–2) | 2 (0–3) | 0.17 |
| S&E ("on"; mean, SD) | 84.1 (14.2) | 86.3 (12.1) | 83.2 (15.0) | 0.55 |
| Levodopa dose-equivalent (mg; mean, SD) | 797.5 (451.2)[b] | 534.7 (367.7)[b] | 892.2 (443.8) | 0.003* |
| Presence of other sleep disorders, n (%) | | | | |
| OSAS | 47 (68.1) | 15 (78.9) | 32 (64.0) | 0.26 |
| RLS | 7 (10.1) | 0 (0.0) | 7 (14.0) | 0.18 |
| Insomnia | 39 (56.5) | 12 (63.2) | 27 (54.0) | 0.59 |
| >1 of the above disorders (with or without RBD) | 29 (42.0) | 9 (47.4) | 20 (40.0) | 0.91 |

HY: Hoehn and Yahr staging scale; OSAS: obstructive sleep apnea syndrome; PD: Parkinson's disease; PD(-)RBD: patient with PD and without RBD; PD(+)RBD: patients with PD and RBD; RBD: REM sleep behavior disorder; REM: rapid eye movement; RLS: restless leg syndrome; SD: standard deviation; S&E: Schwab and England Functional Scale; UPDRS: Unified Parkinson's Disease Rating Scale. [a]Comparative analysis between the PD(+)RBD and the PD(-)RBD groups. Fisher's test was used for categorical variables, while Mann-Whitney test and Student's t-test was used for quantitative variables; normality assessment was conducted using the Shapiro-Wilk test; [b]missing data for one patient. *Significant (p<0.05).

caused no significant changes in that value. The coefficient of discrimination of the item-test correlation test was adequate for all items, except item 8 (0.26; p=0.03). Item 10 was inconclusive (p>0.05). Re-test (n=13, after 21 days) showed moderate-to-strong correlation (p=0.764; 95%CI 0.30–0.97; p=0.002; 0.88 test power) and high agreement for total score (ICC=0.863; 95%CI 0.611–0.956; p<0.001). The evaluation of each item separately in the re-test (n=13) showed complete agreement for items 2 and 3 (K=1.0), and acceptable agreement (K>0.6) for items 6.2, 6.3, 7, and 8. The agreement of items 1 and 10 could not be calculated because the test only elicited positive answers. The analysis of combined items showed complete agreement for 1+2, and acceptable for 2+6.1, 6.1+6.2, 1+2+6.2, 1+6.2, 1+2+6.1, 1+2+6.1+6.2, 2+6.2, 1+6.1+6.2, and 2+6.1+6.2. No conclusive indication (p>0.05) was assessed for the combined items 1+6.1.

## Questionnaire validity

The RBDSQ-BR score was assessed using the ROC curve (Figure 1), showing discriminatory power between groups (n=69; 0.728 AUC; 95%CI 0.588–0.867; p=0.004). According to the shortest distance principle, a cut-off score of 4 was optimum for balancing sensitivity (84.0%; 95%CI 70.9–92.8) and specificity (57.9%; 95%CI 33.5–79.7), determining 2.0 LR+ (95%CI 1.2–3.4) and 0.3 LR- (95%CI 0.1–0.6). A cut-off score 3 had the highest sensitivity (90.0%, 95% CI: 80.0, 98.0), with the lowest LR- 0.2 (95% CI: 0.1, 0.6). Cut-off score 7 had 78.9% specificity (95% CI: 57.9, 94.7), with 2.7 LR+ (95% CI: 1.1, 6.6). (Table 2).

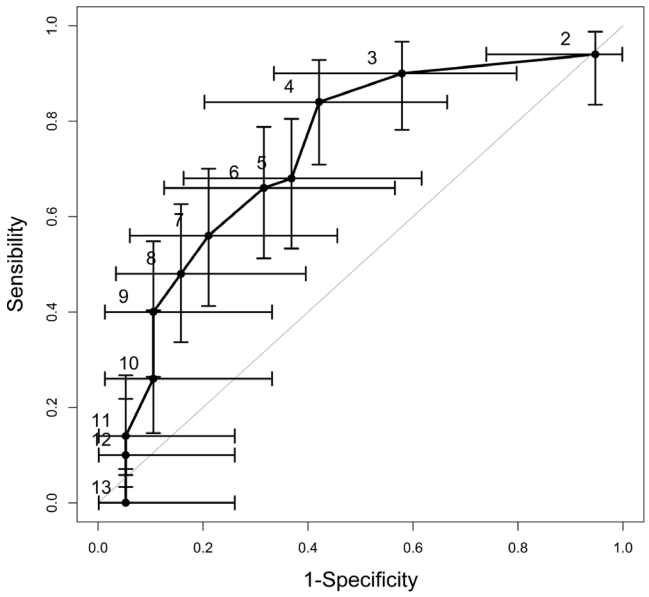


Figure 1. Receiver operating characteristic curve of the RBDSQ-BR (n=69; 0.728 AUC; 95%CI 0.588–0.867; p=0.004). A cut-off score of 4 showed the shortest distance from the upper left corner, balancing the sensitivity and specificity values, with confidence intervals avoiding the main diagonal line (S=84%; 95%CI 74.0–94.0; E=57.9%; 95%CI 36.85–78.9). The confidence intervals of cut-off scores 3, 5, 6, 7, 8, and 9 also avoided the main diagonal line. All patients scored at least 1 in the questionnaire, precluding calculating the values of that cut-off score (AUC: area under the curve; CI: confidence interval; E: specificity; RBDSQ-BR: REM Sleep Behavior Disorder Screening Questionnaire - Brazilian Portuguese version; REM: rapid eye movement; S: sensibility).

Table 2. Analysis of Brazilian Portuguese version of the Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire cut-off scores, n=69.

| Cut-off scores | S% (95%CI) | E% (95%CI) | PPV% (95%CI)[a] | NPV% (95%CI)[a] | Ac% (95%CI)[a] | LR+ (95%CI) | LR- (95%CI) |
|---|---|---|---|---|---|---|---|
| 2 | 94.0 (88.0–100.0) | 5.3 (0.0–15.8) | 72.3 (59.8–82.7) | 25.0 (0.6–80.6) | 69.6 (57.3–80.1) | 1.0 (0.9–1.1) | 1.1 (0.1–10.3) |
| 3[b] | 90.0 (80.0–98.0) | 42.1 (21.0–63.2) | 80.4 (70.9–92.8) | 61.5 (31.6–86.1) | 76.8 (65.1–86.1) | 1.6 (1.0–2.3) | 0.2 (0.1–0.6)* |
| 4[b,c] | 84.0 (74.0–94.0) | 57.9 (36.8–78.9) | 84.0 (70.9–92.8) | 57.9 (33.5–79.7) | 76.8 (65.1–86.1) | 2.0 (1.2–3.4) | 0.3 (0.1–0.6) |
| 5[b] | 68.0 (56.0–80.0) | 63.2 (42.1–84.2) | 82.9 (67.9–92.8) | 42.9 (24.5–62.8) | 66.7 (54.3–77.6) | 1.8 (1.0–3.4) | 0.5 (0.3–0.9) |
| 6[b] | 66.0 (54.0–78.0) | 68.4 (47.4–89.5) | 84.6 (69.5–94.1) | 43.3 (25.5–62.6) | 66.7 (54.3–77.6) | 2.1 (1.0–4.2) | 0.5 (0.3–0.8) |
| 7[b] | 56.0 (43.9–70.0) | 78.9 (57.9–94.7) | 87.5 (71.0–96.5) | 40.5 (24.8–57.9) | 62.3 (49.8–73.7) | 2.7 (1.1–6.6)** | 0.7 (0.4–0.8) |
| 8[b] | 48.0 (34.0–64.0) | 84.2 (63.2–100.0) | 88.9 (70.8–97.6) | 38.1 (23.6–54.4) | 58.0 (45.5–69.8) | 3.0 (1.0–8.9) | 0.6 (0.4–0.9) |
| 9[b] | 40.0 (26.0–54.0) | 89.5 (73.7–100.0) | 90.9 (70.8–98.9) | 36.2 (22.7–51.5) | 53.6 (41.2–65.7) | 3.8 (1.0–14.7) | 0.7 (0.5–0.9) |
| 10 | 26.0 (14.0–38.0) | 89.5 (73.7–100.0) | 86.7 (59.5–98.3) | 31.5 (19.5–45.6) | 43.5 (31.6–56.0) | 2.5 (0.6–9.9) | 0.8 (0.7–1.0) |
| 11 | 14.0 (6.0–24.0) | 94.7 (84.2–100.0) | 87.5 (47.3–99.7) | 29.5 (18.5–42.6) | 36.2 (25.0–48.7) | 2.7 (0.3–20.1) | 0.9 (0.8–1.1) |
| 12 | 10.0 (2.0–18.0) | 94.7 (84.2–100.0) | 83.3 (35.9–99.6) | 28.6 (17.9–41.3) | 33.3 (22.4–45.7) | 1.9 (0.2–15.2) | 0.9 (0.8–1.1) |
| 13 | 0.0 (0.0–0.0) | 94.7 (84.2–100.0) | 0 (0–98.7) | 26.5 (16.5–38.6) | 26.1 (16.3–38.1) | 0 (-) | 1.1 (0.9–1.2) |

Ac: accuracy; CI: confidence interval; E: specificity; LR+: likelihood ration of a positive result; LR-: likelihood ration of a negative result; NPV: negative predictive value; PPV: positive predictive value; REM: rapid eye movement; S: sensitivity. [a]Pre-test probability: 72.5% (50/69=0,725); [b]cut-off scores whose CI avoid the main diagonal line of the receiver operating characteristic (ROC) curve; [c]cut-off score with the shortest distance from the upper left corner. *Lowest LR- value within the cut-off scores, whose CI avoided the main diagonal line of the ROC curve. **Highest LR+ value within the cut-off scores, whose CI avoided the main diagonal line of the ROC curve. The LR+ CI also excluded the value "1."

The value of each item for RBD diagnosis is outlined in Table 3. Items 1, 2, 6.1, and 6.2 were the only items with discriminatory power between groups (p<0.05), separately. Among them, item 1 had the highest sensitivity (86%; 95%CI 73.3–94.2), with the lowest LR- (0.4; 95%CI 0.1–0.9).

Table 4 outlines the combined values of positive answers for those items. The combinations 1+2, 1+6.1, 1+2+6.1, 2+6.1, 2+6.2, 6.1+6.2, 1+2+6.2, and 1+6.1+6.2 had discriminatory power (p<0.05).

The AUC of the ROC curve of each condition was calculated to assess whether other sleep disorders might be confounders. Only insomnia was significant, albeit with a smaller AUC than that of RBD (AUC 0.652; 95%CI 0.519–0.785; p=0.031; Table 5).

**Table 3.** Analysis of each Brazilian Portuguese version of the Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire item[a], n=69.

| Item | S% (95%CI) | E% (95%CI) | PPV% (95%CI)[b] | NPV (95%CI)[b] | Ac% (95%CI)[b] | LR+ (95%CI) | LR- (95%CI) | p-value[c] |
|---|---|---|---|---|---|---|---|---|
| 1 | 86.0 (73.3–94.2)** | 36.8 (16.3–61.6) | 78.2 (65.0–88.2) | 50.0 (23.0–77.0) | 72.5 (60.4–82.5) | 1.4 (0.9–1.9) | 0.4 (0.1–0.9)** | 0.048* |
| 2[d] | 50.0 (35.5–64.5) | 84.2 (60.4–96.6) | 89.3 (71.8–97.7) | 39.0 (24.2–55.5) | 59.4 (46.9–71.1) | 3.2 (1.1–9.3)*** | 0.6 (0.4–0.8) | 0.013* |
| 3[d] | 50.0 (35.5–64.5) | 73.7 (48.8–90.9) | 83.3 (65.3–94.4) | 35.9 (21.2–52.8) | 56.5 (44.0–68.4) | 1.9 (0.8–4.2) | 0.7 (0.5–1.0) | 0.105 |
| 4 | 58.0 (43.2–71.8) | 57.9 (33.5–79.7) | 78.4 (61.8–90.2) | 34.4 (18.6–53.2) | 58.0 (45.5–69.8) | 1.4 (0.8–2.4) | 0.7 (0.4–1.2) | 0.286 |
| 5 | 32.0 (19.5–46.7) | 84.2 (60.4–96.6) | 84.2 (60.4–96.6) | 32.0 (19.5–46.7) | 46.4 (34.3–58.8) | 2.0 (0.7–6.2) | 0.8 (0.6–1.1) | 0.235 |
| 6.1 | 58.0 (43.2–71.8) | 84.2 (60.4–96.6) | 90.6 (75.0–98.0) | 43.2 (27.1–60.5) | 65.2 (52.8–76.3) | 3.7 (1.3–10.6)*** | 0.5 (0.3–0.7) | 0.002* |
| 6.2[e] | 50.0 (35.5–64.5) | 84.2 (60.4–96.6) | 89.3 (71.8–97.7) | 39.0 (24.2–55.5) | 59.4 (46.9–71.1) | 3.2 (1.1–9.3)*** | 0.6 (0.4–0.8) | 0.013* |
| 6.3[e] | 26.0 (14.6–40.3) | 94.7 (74.0–99.9) | 92.9 (66.1–99.8) | 32.7 (20.7–46.7) | 44.9 (32.9–57.4) | 4.9 (0.7–35.2) | 0.8 (0.6–0.9) | 0.091 |
| 6.4 | 28.0 (16.2–42.5) | 94.7 (74.0–99.9) | 93.3 (68.1–99.8) | 33.3 (21.1–47.5) | 46.4 (34.3–58.8) | 5.3 (0.7–37.7) | 0.8 (0.6–0.9) | 0.052 |
| 7[e] | 60.0 (45.2–73.6) | 68.4 (43.4–87.4) | 83.3 (67.2–93.6) | 39.4 (22.9–57.9) | 62.3 (49.8–73.7) | 1.9 (0.9–3.8) | 0.6 (0.4–0.9) | 0.058 |
| 8[e] | 62.0 (47.2–75.3) | 52.6 (28.9–75.6) | 77.5 (61.5–89.2) | 34.5 (17.9–54.3) | 59.4 (46.9–71.1) | 1.3 (0.8–2.2) | 0.7 (0.4–1.3) | 0.290 |
| 9 | 52.0 (37.4–66.3) | 73.7 (48.8–90.9) | 83.9 (66.3–94.5) | 36.8 (21.8–54.0) | 58.0 (45.5–69.8) | 2.0 (0.9–4.4) | 0.6 (0.4–1.0) | 0.064 |
| 10 | 88.0 (75.7–95.5) | 15.8 (3.4–39.6) | 73.3 (60.3–83.9) | 33.3 (7.5–70.1) | 68.1 (55.8–78.8) | 1.0 (0.8–1.3) | 0.8 (0.2–2.7) | 0.699 |

Ac: accuracy; 95%CI: 95% confidence interval; E: specificity; LR+: likelihood ration of a positive result; LR: likelihood ration of a negative result; NPV: negative predictive value; PPV: positive predictive value; REM: rapid eye movement; S: sensitivity. [a]Answer "yes" in the item in question; [b]pre-test probability: 72.5% (50/69=0.725); [c]Fisher's exact test; [d]items with complete agreement (K=1,0) in the re-test (n=13); [e]items with acceptable agreement (K>0.6) in the re-test (n=13). *Significant (p<0.05). **Highest S and lowest LR- among the items with discriminatory power. ***Highest LR+ values among the items with discriminatory power.

**Table 4.** Analysis of combined Brazilian Portuguese version of the Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire items[a], n=69.

| Items | S% (95%CI) | E% (95%CI) | PPV% (95%CI)[b] | NPV% (95%CI)[b] | Ac% (95%CI)[b] | LR+ (95%CI) | LR- (95%CI) | p-value[c] |
|---|---|---|---|---|---|---|---|---|
| 1+2 | 50.0 (35.5–64.5) | 84.2 (60.4–96.6) | 89.3 (71.8–97.7) | 39.0 (24.2–55.5) | 59.4 (46.9–71.1) | 3.2 (1.1–9.3) | 0.6 (0.4–0.8) | 0.013* |
| 1+6.1 | 52.0 (37.4–66.3) | 94.7 (74.0–99.9) | 96.3 (81.0–99.9) | 42.9 (27.7–59.0) | 63.8 (51.3–75.0) | 9.9 (1.4–67.8)** | 0.5 (0.4–0.7) | 0.003* |
| 1+6.2 | 42.0 (28.2–56.8) | 84.2 (60.4–96.6) | 87.5 (67.6–97.3) | 35.6 (21.9–51.2) | 53.6 (41.2–65.7) | 2.7 (0.9–7.9) | 0.7 (0.5–0.9) | 0.051 |

Continue...

Table 4. Continuation.

| Items | S% (95%CI) | E% (95%CI) | PPV% (95%CI)[b] | NPV% (95%CI)[b] | Ac% (95%CI)[b] | LR+ (95%CI) | LR- (95%CI) | p-value[c] |
|---|---|---|---|---|---|---|---|---|
| 1+2+6.1 | 40.0 (26.4–54.8) | 94.7 (74.0–99.9) | 95.2 (76.2–99.9) | 37.5 (24.0–52.6) | 55.1 (42.6–67.1) | 7.6 (1.1–52.8)** | 0.6 (0.5–0.8) | 0.007* |
| 1+2+6.1+6.2 | 30.0 (17.9–44.6) | 94.7 (74.0–99.9) | 93.8 (69.8–99.8) | 34.0 (21.5–48.3) | 47.8 (35.6–60.2) | 5.7 (0.8–40.2) | 0.7 (0.6–0.9) | 0.052 |
| 2+6.1 | 40.0 (26.4–54.8) | 94.7 (74.0–99.9) | 95.2 (76.2–99.9) | 37.5 (24.0–52.6) | 55.1 (42.6–67.1) | 7.6 (1.1–52.8)** | 0.6 (0.5–0.8) | 0.007* |
| 2+6.2 | 32.0 (19.5–46.7) | 94.7 (74.0–99.9) | 94.1 (71.3–99.9) | 34.6 (22.0–49.1) | 49.3 (37.0–61.6) | 6.1 (0.9–42.7) | 0.7 (0.6–0.9) | 0.027* |
| 6.1+6.2 | 40.0 (26.4–54.8) | 94.7 (74.0–99.9) | 95.2 (76.2–99.9) | 37.5 (24.0–52.6) | 55.1 (42.6–67.1) | 7.6 (1.1–52.8)** | 0.6 (0.5–0.8) | 0.007* |
| 1+2+6.2 | 32.0 (19.5–46.7) | 94.7 (74.0–99.9) | 94.1 (71.3–99.9) | 34.6 (22.0–49.1) | 49.3 (37.0–61.6) | 6.1 (0.9–42.7) | 0.7 (0.6–0.9) | 0.027* |
| 1+6.1+6.2 | 34.0 (21.2–48.8) | 94.7 (74.0–99.9) | 94.4 (72.7–99.9) | 35.3 (22.4–49.9) | 50.7 (38.4–63) | 6.5 (0.9–45.2) | 0.7 (0.6–0.9) | 0.015* |
| 2+6.1+6.2 | 30.0 (17.9–44.6) | 94.7 (74.0–99.9) | 93.8 (69.8–99.8) | 34.0 (21.5–48.3) | 47.8 (35.6–60.2) | 5.7 (0.8–40.2) | 0.7 (0.6–0.9) | 0.052 |

Ac: accuracy; 95%CI: 95% confidence interval; E: specificity; LR+: likelihood ratio of a positive result; LR-: likelihood ratio of a negative result; NPV: negative predictive value; PPV: positive predictive value; REM: rapid eye movement; S: sensitivity. [a]Answer "yes" in the combined items; [b]pre-test probability: 72.5% (50/69=0.725); [c]Fisher's exact test. *Significant (p<0.05). **LR+ whose confidence intervals excluded the value 1.

Table 5. Other sleep disorders as Brazilian Portuguese version of the Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire confounders: comparison between area under the curve, n=69.

| | AUC | 95%CI | p-value |
|---|---|---|---|
| RBD[a] | 0.728 | 0.588, 0.867 | 0.004* |
| OSAS | 0.456 | 0.317, 0.596 | 0.562 |
| RLS | 0.674 | 0.522, 0.826 | 0.133 |
| Insomnia | 0.625 | 0.519, 0.785 | 0.031* |

AUC: area under the curve; 95%CI: 95% confidence interval; OSAS: obstructive sleep apnea syndrome; RLS: restless leg syndrome. [a]Highest AUC value for Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire use. *Significant (p<0.05).

## DISCUSSION

This study determined the diagnostic value of the Brazilian Portuguese version of the RBDSQ to validate its use for patients with PD. All participants were subjected to the gold standard of RBD diagnosis (clinical interview+PSG)[5]. The original RBDSQ study[7] and other studies that have sought to assess the diagnostic value of the questionnaire[12,13,14,15,16,17,18] have also used this gold standard to evaluate participants (except healthy controls).

The sample consisted of 50 participants in the PD+RBD group and 19 participants in the PD-RBD group. As in a previous study conducted by Stiasny-Kolster et al.[18], the sample had a slight predominance of elderly men, although they were slightly younger (*i.e.*, 68 years *versus* 60.4±12 years, respectively). Although the disease duration was similar (approximately 8.5 years), PD was less advanced in the current study (*i.e.*, 2.0 *versus* 3.0) according to the HY staging.

RBDSQ-BR proved reliable, with adequate internal consistency (overall Cronbach's α=0.809)[28]. The item-test correlation was adequate (>0.3), except for item 8 (0.26), suggesting the low discriminatory power of this item. Item 10 was inconclusive (p>0.05). The total score of the instrument showed high agreement (ICC=0.863). Items 2, 3, 6.2, 6.3, 7, and 8 had adequate agreement when assessed separately (K>0.60)[30]. The other items were inconclusive in this regard.

The original study by Stiasny-Kolster et al.[7] already had adequate internal consistency (0.885 Cronbach's α), which was also shown in other validation studies conducted in Japan[12], China[13], South Korea[14], Turkey[15], and Italy[16]. The first study, which focused on patients with PD in Japan, found an adequate internal consistency (0.73 Cronbach's α)[17]. The item-test correlation of all items in the original study[7] was satisfactory (>0.3); however, the South Korean and Italian studies reported inadequate item-test correlations for item 10[14,16]. Our study showed low discriminatory power for item 8. This item indicates an awakening while dreaming due to the motor activity reported in item 7 (see Online resource). If this relationship is not recognized, it can be assumed that item 8 is about remembering the dreams that occurred the night before. Considering the low schooling (Table 1), a misinterpretation could explain the low discriminatory power of this item in our sample.

Agreement analysis of the re-test total score was reported in some studies[12,13,14,15] and proved satisfactory, albeit with variable agreement between the instrument items. Miyamoto et al.[12] showed good agreement for items

1, 2, 5, and 6.1 (K>0.60). Conversely, the South Korean study[14] only reported this for items 7 and 10. Thus far, no study had focused on a population of patients with PD for assessing the RBDSQ agreement of either its total score or its items.

In our study, a cut-off score of 4 enabled the correct diagnosis of 76.8% subjects and provided the best balance between sensitivity (84%) and specificity (57.9%), with 2.0 LR+ and 0.3 LR-. A cut-off score of 3 also correctly diagnosed 76.8% individuals and provided increased sensitivity (90.0%) with 0.2 LR- (an approximately 5-fold reduction in the probability of having RBD), whereas a cut-off score of 7 provided increased specificity (78.9%), with a correct diagnosis of 62.3% subjects and 2.7 LR+. Thus, total scores of <3 in our study might be useful to exclude RBD (LR- from 0.5 to 0.2 may generate small-to-moderate changes in post-test probability)[33]. The separate analysis of each item in our study enabled us to increase its specificity. Both items 2 and 6.2 had a specificity of 84.2%, with 3.2 LR+ (an increase of slightly more than 3 times the probability) and may be regarded as of little (but not necessarily unimportant) utility for RBD diagnosis[33]. The use of combined items also enabled us to further increase the specificity. The combination of positive answers in items 1+2+6.2, 2+6.1, and 6.1+6.2 determined a specificity of 94.7%, rendering more robust LR+ (6.1 up to 7.6, indicating an approximately 6-to-8-fold increase in the probability of having RBD) with acceptable reliability.

The original study by Stiasny-Kolster et al.[7] indicated a higher value (cut-off score 5) of the balance between sensitivity and specificity (96% sensitivity and 56% specificity). The low specificity of the total score of the instrument was attributed to the presence of comorbidities associated with excessive motor activity during sleep, including restless leg syndrome (RLS), obstructive sleep apnea syndrome (OSAS), periodic limb movements, and narcolepsy. Those comorbidities would render positive answers in items indicating limb movements (sub-items 4, 5, 6.2, and 7), inflating the resulting score. Another explanation was the presence of other sleep and neurological disorders, which would have rendered positive answers in sub-items 9 and 10, increasing the final score. As in the present study, the analysis of separate items in the Stiasny-Kolster study also allowed an increase of specificity, with 85.3–91.1% specificity when items 5, 6.3, and 6.4 were answered positively.

In our study, both the PD+RBD and PD-RBD subgroups included patients with other sleep disorders (such as RLS, OSAS, and insomnia), at similar ratios between groups. Thus, the low specificity found in the current study may be explained using the same arguments. The RBDSQ-BR score, however, showed no potential to select those patients, except in the case of insomnia. Even in this case, the resulting AUC (0.625) was less than that related to RBD (0.728), which was the target of the questionnaire.

The RBDSQ versions that were validated in Far East Asia[12,13,14,15] confirmed the cut-off score of 5 reported in the original study[7], with sensitivity ranging from 88.5 to 100%, albeit with markedly high specificities (more often >90%). A new validation in Europe (Italy)[16] indicated a higher value as the best cut-off score (*i.e.*, 8), with sensitivity of 84.2% and lower specificity of 78.0%. The validation study in Japan reported item 5 as the most specific one (92.3–96.4%) when the items were analyzed separately[12]. Tari et al.[15] found higher specificities in the Turkish population, not only for item 5 (85.7–97.4%), but also for items 7 (94.9%) and 10 (93.6%). Curiously, the Italian study[16] not only had higher specificities in some items separately (83.2–83.5% for items 5, 6.3, and 6.4), but also higher sensitivities than the cut-off total score (90.8–92.1% for items 1, 3, and 6.1).

The samples of those studies, however, were different from ours in that they were sometimes highly heterogeneous. Few studies have included patients with PD[7,15], and except for that by Wang et al.[13], the number of subjects with PD in these studies was minimal. Further, the method of selecting participants was not homogeneous between studies. Moreover, although the RBDSQ allows the companion to help, not all studies included that participation[12,15].

The study by Nomura et al.[17] from 2011 was the first with the primary objective of evaluating RBDSQ performance specifically in patients with PD. Using the version validated in Japan[12], consecutive patients with PD and patients with RBD alone were evaluated. The ROC curve of patients with PD showed a value of 6 as the best cut-off score, with a sensitivity similar to that observed in our study (84.2%), albeit with a considerably higher specificity (96.2%), which was comparable to other studies conducted in Far East Asia. The increase of 1 point in the cut-off score was explained by the positivity necessarily present in item 10 of the instrument. A possible explanation for the high value of specificity may be the fact that apparently no other sleep disorders were identified in the patients. Otherwise, they would have tended to score several items of the instrument, as previously mentioned.

In 2015, Wang et al.[13] included a sub-sample of patients with PD, with and without RBD. The cut-off score of 6 was also the most adequate for those subjects, with a sensitivity and specificity of 90.9 and 91.9%, respectively, which determined the diagnostic accuracy in 91.52% of patients. Their results did not show other sleep disorders in those patients, which might explain the high specificity.

In the same year, Stiasny-Kolster et al.[18] evaluated consecutive patients with PD, with no help from the companion, and concluded that the "learning" effect resulting from a detailed clinical evaluation before filling in the questionnaire may have significantly affected the diagnostic power of the instrument, at least in the case of PD.

In our study, the patients filled in the questionnaire before the clinical evaluation, which may explain why the overall RBDSQ-BR performance was clearly similar to the poor performance of the group that filled the instrument first in the previous study[18]. Nonetheless, the performance of

the patients in the current study cohort was relatively better, which may be explained by the fact that most of our participants were helped by their companions. However, it should be noted that a possible epidemiological survey in our setting would most likely include the companion alongside the patient with PD filling in the questionnaire.

A few limitations in the current study must be noted. First, there was a lack of a control group of healthy subjects and a group of subjects with RBD and without PD (RBD alone), which would have improved the evaluation of the questionnaire performance and increased the generalization of the findings. Second, despite the similarity in education, either the patients' or their companions', between the analysis groups, its effect on the instrument performance cannot be ruled out. Further, the effect of companion assistance on instrument accuracy could also not be ruled out. Finally, the imbalance between the sizes of the groups, *i.e.*, the small *n* in the group of patients without RBD, may have affected the determination of a low cut-off total score compared with the cut-off scores of other studies on patients with PD.

In conclusion, the results of the current study demonstrated that RBDSQ-BR was a valid and reliable instrument and that it may be useful for diagnosing RBD in Brazilian patients with PD. The instrument may also help in improving the selection of cases for a more detailed clinical evaluation or even polysomnography.

## ACKNOWLEDGMENTS

We thank Hayala Cavenague de Souza for her help in statistical analysis.

## References

1. Schenck CH, Bundlie SR, Ettinger MG, Mahowald MW. Chronic behavioral disorders of human REM sleep: a new category of parasomnia. Sleep. 1986 Jun;9(2):293-308. https://doi.org/10.1093/sleep/9.2.293
2. Boeve BF. REM sleep behavior disorder: updated review of the core features, the RBD-Neurodegenerative Disease Association, evolving concepts, controversies, and future directions. Ann N Y Acad Sci. 2010 Jan;1184:15-54. https://doi.org/10.1111/j.1749-6632.2009.05115.x
3. Eisensehr I, v Lindeiner H, Jäger M, Noachtar S. REM sleep behavior disorder in sleep-disordered patients with versus without Parkinson's disease: is there a need for polysomnography? J Neurol Sci. 2001 May;186(1-2):7-11. https://doi.org/10.1016/s0022-510x(01)00480-4
4. Sobreira-Neto MA, Pena-Pereira MA, Sobreira EST, Chagas MHN, Fernandes RMF, Tumas V, et al. High frequency of sleep disorders in Parkinson's disease and its relationship with quality of life. Eur Neurol 78:330-7. https://doi.org/10.1159/000481939
5. Sateia M. International Classification of Sleep Disorders. 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
6. Durán-Cantolla J, Mar J, La Torre GD, Aramendi RR, Guerra L. El síndrome de apneas-hipopneas durante el sueño en España. Disponibilidad de recursos para su diagnóstico y tratamiento en los hospitales del Estado español. Arch Bronconeumol. 2004;40(6):259-67. https://doi.org/10.1016/S0300-2896(04)75517-9
7. Stiasny-Kolster K, Mayer G, Schäfer S, Möller JC, Heinzel-Gutenbrunner M, Oertel WH. The REM sleep behavior disorder screening questionnaire - A new diagnostic instrument. Mov Disord. 2007 Dec;22(16):2386-93. https://doi.org/10.1002/mds.21740
8. Li SX, Wing YK, Lam SP, Zhang J, Yu MW, Ho CKW, et al. Validation of a new REM sleep behavior disorder questionnaire (RBDQ-HK). Sleep Med. 2010 Jan;11(1):43-8. https://doi.org/10.1016/j.sleep.2009.06.008
9. Boeve BF, Molano JR, Ferman TJ, Lin SC, Bieniek K, Tippmann-Peikert M, et al. Validation of the Mayo Sleep Questionnaire to screen for REM Sleep Behavior Disorder in a community-based sample. J Clin Sleep Med. 2013 May;9(5):475-80. https://doi.org/10.5664/jcsm.2670
10. Frauscher B, Ehrmann L, Zamarian L, Auer F, Mitterling T, Gabelia D, et al. Validation of the Innsbruck REM sleep behavior disorder inventory. Mov Disord. 2012 Nov;27(13):1673-8. https://doi.org/10.1002/mds.25223
11. Postuma RB, Arnulf I, Hogl B, Iranzo A, Miyamoto T, Dauvilliers Y, et al. A single-question screen for rapid eye movement sleep behavior disorder: a multicenter validation study. Mov Disord. 2012 Jun;27(7):913-6. https://doi.org/10.1002/mds.25037
12. Miyamoto T, Miyamoto M, Iwanami M, Kobayashi M, Nakamura M, Inoue Y, et al. The REM sleep behavior disorder screening questionnaire: Validation study of a Japanese version. Sleep Med. 2009 Dec;10(10):1151-4. https://doi.org/10.1016/j.sleep.2009.05.007
13. Wang Y, Wang Z, Yang Y, Wu H, Zhao HY, Zhao Z. Validation of the rapid eye movement sleep behavior disorder screening questionnaire in China. J Clin Neurosci. 2015;22(9):1420-24. https://doi.org/10.1016/j.jocn.2015.03.008
14. Lee S, Paek J, Han S, Ryu H. The utility of a Korean version of the REM sleep behavior disorder screening questionnaire in patients with obstructive sleep apnea. J Neurol Sci. 2015 Nov;358(1-2):328-32. https://doi.org/10.1016/j.jns.2015.09.347.
15. Tari Cömert I, Pelin Z, Aricak T, Yapan S. Validation of the Turkish Version of the Rapid Eye Movement Sleep Behavior Disorder Questionnaire. Behav Neurol. 2016;2016:8341651. https://doi.org/10.1155/2016/8341651
16. Marelli S, Rancoita PMV, Giarrusso F, Galbiati A, Zucconi M, Oldani A, et al. National validation and proposed revision of REM sleep behavior disorder screening questionnaire (RBDSQ). J Neurol. 2016 Dec;263(12):2470-5. https://doi.org/10.1007/s00415-016-8285-y
17. Nomura T, Inoue Y, Kagimura T, Uemura Y, Nakashima K. Utility of the REM sleep behavior disorder screening questionnaire (RBDSQ) in Parkinson's disease patients. Sleep Med. 2011 Aug;12(7):711-3. https://doi.org/10.1016/j.sleep.2011.01.015
18. Stiasny-Kolster K, Sixel-Döring F, Trenkwalder C, Heinzel-Gutenbrunner M, Seppi K, Poewe W, et al. Diagnostic value of the REM sleep behavior disorder screening questionnaire in Parkinson's disease. Sleep Med. 2015 Jan;16(1):186-9. https://doi.org/10.1016/j.sleep.2014.08.014
19. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry. 1992 Mar;55(3):181-4. https://doi.org/10.1136/jnnp.55.3.181
20. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993 Dec;46(12):1417-32. https://doi.org/10.1016/0895-4356(93)90142-n

21. 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 Dec;25(24):3186-91. https://doi.org/10.1097/00007632-200012150-00014

22. Fahn S, Elton R. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Calne DB, Goldstein M, editors. Recent developments in Parkinson's disease. v. 2. Florham Park: Macmillan Health Care Information; 1987. p.153-63,293-304.

23. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. Neurology. 1967;17(5):427-42. https://doi.org/10.1212/WNL.17.5.427

24. Schwab RS. Projection techniques for evaluating surgery in Parkinson's Disease. In: Gillingham FJ, Donaldson IML, editors. Third Symposium on Parkinson's Disease, Royal College of Surgeons in Edinburgh. Edinburgh & London: E&S Livingstone; 1969. p. 152-7.

25. Iber C, Ancoli-Israel S, Chesson Jr AL, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: Rules Terminology and Technical Specifications. Westchester, IL: American Academy of Sleep Medicine; 2007.

26. Armitage P, Berry G, Matthews JN. Statistical Methods in Medical Research. 4. ed. Malden, MA: Blackwell Publishing; 2002.

27. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979 Mar;86(2):420-8. https://doi.org/10.1037//0033-2909.86.2.420

28. Bland JM, Altman DG. Cronbach's alpha. BMJ. 1997 Jan;314(7080):572. https://doi.org/10.1136/bmj.314.7080.572

29. Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007 Jan;60(1):34-42. https://doi.org/10.1016/j.jclinepi.2006.03.012

30. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury. 2011 Mar;42(3):236-40. https://doi.org/10.1016/j.injury.2010.11.042

31. Fawcett T. An introduction to ROC analysis. Pattern Recognitt Lett. 2006;27(8):861-74. https://doi.org/10.1016/j.patrec.2005.10.010

32. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? JAMA. 1994;271(5):389-91. https://doi.org/10.1001/jama.271.5.389

33. Jaeschke R, Guyatt G, Sackett DL. Users' Guides to the Medical Literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? JAMA. 1994;271(9):703-7. https://doi.org/10.1001/jama.271.9.703