Published in final edited form as:

Mov Disord. 2009 April 15; 24(5): 684–688. doi:10.1002/mds.22437.

How Cautious Should We Be When Assessing Apathy with the Unified Parkinson's Disease Rating Scale?

Lindsey Kirsch-Darrow, MS^{1,*}, Laura B. Zahodne, MS¹, Chris Hass, PhD², Ania Mikos, PhD¹, Michael S. Okun, MD³, Hubert H. Fernandez, MD³, and Dawn Bowers, PhD¹

¹Department of Clinical and Health Psychology, College of Public Health and Health Professions, University of Florida, Gainesville, Florida, USA

²Department of Applied Physiology and Kinesiology, University of Florida, Gainesville, Florida, USA

³Department of Neurology, University of Florida, Gainesville, Florida, USA

Abstract

Current practice often assesses apathy with a single item from the Unified Parkinson's Disease Rating Scale (UPDRS, item 4). Yet, the relationship between the UPDRS item 4 and the validated Apathy Scale (AS) is unknown. The purpose of this study was to evaluate the operating characteristics of UPDRS item 4 in relation to the AS. Three hundred and one patients with PD were administered the AS and the UPDRS. We compared the UPDRS item 4 to the standard AS classification of ≥ 14 as apathetic. A receiver operating characteristics (ROC) curve was obtained, and sensitivity, specificity, positive, and negative predictive power were calculated. The ROC curve showed area under the curve as 0.75. A cut-off of 1 had good sensitivity (81%) but poor specificity (53%; high false positive rate). A cut-off point of 2 had acceptable specificity (87%) but poor sensitivity (52%, high false negative rate). Continuing to increasing the cut-off point (e.g., 3, 4) continues to increase specificity at the expense of dramatically reducing sensitivity. These findings suggest the use of caution when screening for apathy with item 4 due to its poor sensitivity in relation to the AS.

Keywords

Parkinson's disease; apathy; mood disorders; UPDRS; assessment

Apathy, defined as a primary lack of motivation, is a common mood disturbance in Parkinson's disease (PD). Prevalence estimates range between 16 and 51%.^{1–4} Apathy manifests in several different domains, including behavioral (e.g., decrease in goal directed activities), cognitive (e.g., loss of interest), and emotional (e.g., blunted emotional experience) domains.^{5,6} Thus, apathy can have a negative effect on multiple aspects of patients' lives. There are several rating scales used to assess apathy in PD; these include the Apathy Evaluation Scale (AES),⁵ the Apathy Scale (AS),² the Lille Apathy Rating Scale (LARS),⁷ and a single item from the Unified Parkinson's Disease Rating Scale (UPDRS item 4).⁸ This single item, item 4, assesses motivation/initiative on a 0–4 point Likert scale. Indeed, a single item that captures apathy severity in PD and that could be performed

quickly, and within the standard of care, would provide an attractive means to track this mood disturbance. However, the diagnostic accuracy of item 4 in assessing apathy is unknown. Studies that examine the validity of UPDRS item 4 are limited.^{9,10} In this study, we sought to use a large, consecutive series of patients to determine the utility of UPDRS item 4 in discriminating between patients with PD with or without apathy. To do this, we investigated the concurrent validity of UPDRS item 4 by comparing it to the full 14-item AS.²

PATIENTS AND METHODS

Candidates for inclusion in this study were 334 patients with PD who attended the University of Florida's Movement Disorders Center Clinics. To be included, participants had to meet diagnostic criteria for idiopathic PD in accordance with U.K. Brain Bank criteria¹¹ and be between 30 and 90 years of age. Patients were excluded if they had any neurosurgical treatments for PD such as deep brain stimulation or pallidotomy. The final sample included 301 participants; 33 either did not meet inclusion/exclusion criteria or did not complete one of the questionnaires administered for the study. Prior to participation, informed consent was obtained according to university and federal guidelines.

Participants included 189 men and 112 women who ranged in age from 30 to 90 years ($M = 67.9$, $SD = 10.6$). These participants had obtained an average of 14.8 years of education ($SD = 3.2$, range 5–20). On average, their motor symptoms were in the moderate range when assessed “on” medications (UPDRS motor = 29.6, $SD = 12.2$; Hoehn and Yahr = 2.45, $SD = 0.68$). See Table 1.

Participants completed the UPDRS (while “on” their dopaminergic medications), the Apathy Evaluation Scale (AS), and the Beck Depression Inventory (BDI). The BDI is a 21 item scale measuring depression experienced over the last week. Classification of clinically significant depression was based on Lezak's criteria of BDI score ≥ 14 . The UPDRS is a 42 item clinician administered rating scale that has four parts: (1) Mentation, Behavior, and Mood, (2) Activities of Daily Living, (3) Motor, and (4) Dyskinesia.⁸ The current study focused on a single item, item 4, which assesses motivation/initiative and is in the Mentation, Behavior, and Mood section. UPDRS-I-item 4 is rated on a 0–4 Likert scale and scores include the following: 0 = Normal, 1 = Less assertive than usual; more passive, 2 = Loss of initiative or disinterest in elective (non-routine) activities, 3 = Loss of initiative or disinterest in day-to-day (routine) activities, and 4 = Withdrawn, complete loss of motivation. Participants also completed the modified AS. The AS is a 14-item scale measuring cognitive, emotional, and behavioral symptoms of apathy. Sample items include the following: “Are you indifferent to things?” “Are you interested in learning new things?” “Do you have plans and goals for the future?” Items are rated on a 0–3 Likert scale and total scores range from 0 to 42. The scale was abridged and modified from the original 18-item version⁵ and wording was simplified in 1992.² This modified AS was reported to have good psychometric properties in PD (internal consistency reliability = 0.76, test–retest at 1 week = 0.90). Additionally, a cut-off score of 14 points was established as indicating clinically significant apathy.²

In this study, the diagnostic accuracy of UPDRS item 4 was assessed by comparing it to the AS, with scores ≥ 14 on the AS indicating clinical levels of apathy. This was done by creating a receiver operating characteristic (ROC) curve for item 4 and calculating the area under the curve (AUC). Sensitivity, specificity, positive, and negative predictive power were then calculated for various cut-off points along the ROC curve. Each cut-off point/decision threshold corresponds to a particular sensitivity/specificity pair. Sensitivity is the proportion of people with the condition that are correctly classified by the test, Sensitivity = True

positives/(True positives + False negatives). Specificity is the proportion of people who do not have the condition that are correctly classified by the test, $\text{Specificity} = \text{True negatives} / (\text{True negatives} + \text{False positives})$. Positive predictive value (PPV) and negative predictive value (NPV) are the probability that an obtained result is accurate. PPV is the probability that a person has the condition given a positive test result ($\text{PPV} = \text{True positives} / (\text{True positives} + \text{False positives})$). NPV is the probability that a person does not have the condition given a negative result ($\text{NPV} = \text{True negatives} / (\text{True negatives} + \text{False negatives})$).¹²

Sensitivity and specificity are considered as tradeoffs. The UPDRS Part I examining Mentation, Behavior, and Mood is considered a screening tool for further psychiatric assessment. The goal is to minimize false negative results using an instrument to screen with high sensitivity. However, given the time and expense of further assessment and evaluation for both the patient and the clinician, excessive false positives also need to be avoided. Some authors have suggested that a useful two-stage screening for psychiatric disorders should include a cut-off for patients where the sensitivity and specificity is at least 75%.^{10,13}

RESULTS

Scores on the AS averaged 13.7 (SD = 6.9) and ranged from 0 to 31. Fifty percent (153/301) of patients with PD scored ≥ 14 on the AS. This included 31.4% with apathy alone (e.g., apathy ≥ 14 without depression ≥ 14) and 18.6% with both apathy and depression (apathy ≥ 14 and depression ≥ 14). Only 3.4% had depression and 46.4% did not have clinically significant apathy or depression scores. Scores on the UPDRS item 4 averaged 1.14 (SD = 1.1) and ranged from 0 to 4. The correlation between the AS and the UPDRS item 4 was 0.53 ($P < 0.01$). As for percentages of individuals in each UPDRS score category: scoring a 0 = 35.9% (108/301), scoring a 1 = 31.2% (94/301), scoring a 2 = 19.6% (59/301), scoring a 3 = 9.3% (28/301), and scoring a 4 = 4% (12/301). The ROC curve for UPDRS item 4 was calculated and is shown in Figure 1, displaying sensitivity on the y-axis and 1-specificity on the x-axis.¹² Additionally, a global measure of the diagnostic accuracy of a test is the AUC. The AUC for this study is 0.75, indicating half way between perfect (1.0) and worse possible discrimination (0.5) between groups (CI = 0.695–0.805, $P < 0.01$).

Sensitivity, specificity, PPV, and NPV were calculated for UPDRS item 4 cut-off points. A cut-off score of 1 resulted in a sensitivity of 81% and a specificity of 53.4%. PPV was 64.2% and NPV was 73.1%. The cut-off was then raised to 2 (See Table 2). This greatly reduced sensitivity to 52.3%, but increased specificity to 87.2%. The PPV was 80.7% and NPV was 63.9%. When the cut-off was raised even higher to 3, the specificity increased to 97.3%, but the sensitivity fell to a low 23.1%. Thus, there were no UPDRS item 4 cut-off points where both sensitivity and specificity were greater than 75%.

DISCUSSION

Summary and Comparison to Literature

Our ROC curve findings suggest that the UPDRS item 4 does not have a cut-off with an adequate balance of sensitivity and specificity. A cut-off of 1 has good sensitivity (81%) at the expense of poor specificity (53.4%), whereas a cut-off of 2 has good specificity (87.2%) at the expense of poor sensitivity (52.3%). Further, while the total AUC = 0.75 is statistically significant, it falls halfway between the worst (0.5) and the best (1) possible discrimination, indicating mediocre discrimination.

Our results can be compared and contrasted with other studies examining UPDRS item 4 (see Table 2). Pedersen et al. (2007)⁹ examined a consecutive series of 58 patients with PD with the UPDRS item 4 and a shortened version of the AS that used approximately half the

items on the scale. They found the AUC to be 0.70 ($P = 0.05$); a cut-off of 1 had sensitivity of 80% and a specificity of 23%, whereas a cut-off of 2 had sensitivity of 70% and a specificity of 75%. On the basis of these results, they concluded that a 2 or more on UPDRS item 4 was adequate for screening apathy in PD. Starkstein and Merello (2007)¹⁰ used similar methods and examined a consecutive series of 164 patients with PD with the UPDRS item 4 versus the same shortened version of the AS. They found the AUC to be 0.76 ($P < 0.01$), a cut-off of 1 had a sensitivity of 91% and specificity of 46%, whereas a cut-off of 2 indicated sensitivity of 73% and specificity of 65%. Starkstein and Merello (2007),¹⁰ like Pedersen et al. (2007),⁹ concluded that a cut-off of 2 is adequate for screening, but not for diagnostic purposes in PD, despite sensitivity values falling below 75% in both cases.

Our results were similar to both previous studies in the value of the area under the curve (AUC, ours = 0.75 versus 0.70 and 0.74). This indicates that all three studies showed comparable overall discrimination rates of item 4. Comparing across studies, our results were also similar to the others in that a cut-off of 1 has good sensitivity but poor specificity. Our results differed in relation to a cut-off value of 2. We found that a cut-off of 2 had much lower sensitivity (52.3%) than the other studies (70% and 73% for Pedersen et al. (2007)⁹ and Starkstein and Merello (2007),¹⁰ respectively).

Different results regarding the cut-off of 2 may be due to different methods of classifying apathy. This study uses the full AS and the validated criterion of ≥ 14 on the AS to distinguish between apathetic and nonapathetic individuals. The other two studies use a shortened version of the scale based on an algorithm to convert the AS into proposed diagnostic criteria for apathy.^{10,9,14} The scores are used on approximately half of the 14 items in a dichotomous way (i.e., did the patient get a 2 or 3 on item 7 and a score of 2 or 3 on items 4 or 9, etc) to determine if patients were considered apathetic or nonapathetic.

Limitations

A major limitation of our study is the lack of a gold standard of established diagnostic criteria for apathy. Although the AS has good psychometric properties in PD, it has not been validated against established consensual diagnostic criteria for apathy. DSM-IV criteria for a “syndrome of apathy” have been proposed, they have not been validated with a consensus or large scale empirical studies.^{14,15} The previous studies have addressed this by converting the AS into a version of the proposed diagnostic criteria as described earlier.^{9,10} Yet, there are difficulties associated with this conversion. It only uses between 4 and 7 items on the AS and makes them dichotomous items. This limits the amount of information that can be obtained from the scale as well as decreases the range of each item from 0–4 to 0–1. Thus, less information about the nature and severity of the apathy being experienced by the patient is taken into consideration. A final solution to these problems may be to re-examine the operating characteristics of the AS and item 4 once well-validated criteria are established for the syndrome of apathy. We believe this will be an important next step for future studies. Another limitation of our study is that the patients were cognitively unselected (i.e., the sample could include dementia patients). Due to the link between dementia and apathy, this would be an important characteristic to track in the future. Although we collected BDI scales, we did not collect formal DSM-IV depression diagnoses. This is also an important consideration for future studies.

Overall, our study indicates low specificity with a cut-off of 1 and low sensitivity with a cut-off of 2 or higher. This suggests that item 4 is not an adequate screening measure. If the goal is to use this single item to triage patients into further assessment and treatment, then missing nearly half the patients is inadequate. Therefore, on the basis of our results, we recommend using caution when screening solely based on the single item, item 4, from the UPDRS I.

Acknowledgments

Supported by F31-NS059142 and in part by R01-NS05063 and K23-NS044997.

Author Roles: L. Kirsch-Darrow MS first author, conception, organization, execution, writing of manuscript. L. B. Zahodne MS organization, execution, review and critique. C. Hass PhD conception, review and critique. A. Mikos MS review and critique. M. S. Okun MD review and critique, organization. H. H. Fernandez MD review and critique. D. Bowers PhD done in this author's laboratory, conception, organization, execution, review and critique.

REFERENCES

1. Aarsland D, Larsen J, Lim N, et al. Range of neuropsychiatric disturbances in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1999;67:492–496. [PubMed: 10486397]
2. Starkstein S, Mayberg H, Preziosi T, Andrezejewski P, Leiguarda R, Robinson RG. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 1992;4:134–139. [PubMed: 1627973]
3. Isella V, Melzi P, Grimaldi M, et al. Clinical, neuropsychological, and morphometric correlates of apathy in Parkinson's disease. *Mov Disord* 2002;17:366–371. [PubMed: 11921125]
4. Kirsch-Darrow L, Fernandez HH, Marsiske M, Okun MS, Bowers D. Dissociating apathy and depression in Parkinson disease. *Neurology* 2006;67:33–38. [PubMed: 16832074]
5. Marin R. Apathy: a neuropsychiatric syndrome. *J Neuropsychiatry Clin Neurosci* 1991;3:243–254. [PubMed: 1821241]
6. Pluck G, Brown R. Apathy in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2002;73:636–642. [PubMed: 12438462]
7. Sockeel P, Dujardin K, Devos D, Denève C, Destée A, Defebvre L. The Lille apathy rating scale (LARS), a new instrument for detecting and quantifying apathy: validation in Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 2006;5:579–584.
8. Fahn, S.; Elton, R. Members of the UPDRS Development Committee. Unified Parkinson's disease rating scale. In: Fahn, S.; Marsden, C.; Calne, D.; Liberman, A., editors. Recent developments in Parkinson's disease. Florham Park, NJ: Macmillan Health Care Information; 1987. p. 153-163.
9. Pedersen KF, Larsen JP, Aarsland D. Validation of the United Parkinson's Disease Rating Scale (UPDRS) section I as a screening and diagnostic instrument for apathy in patients with Parkinson's disease. *Parkinsonism Relat Disord* 2008;14:183–186. [PubMed: 17889589]
10. Starkstein SE, Merello M. The Unified Parkinson's Disease Rating Scale: Validation study of the mentation, behavior, and mood section. *Mov Disord* 2007;22:2156–2161. [PubMed: 17721877]
11. 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;55:181–184. [PubMed: 1564476]
12. Zweig MH, Campbell G. Receiver-operating characteristics (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;39:561–577. [PubMed: 8472349]
13. Lowe B, Spitzer RL, Grafe K, et al. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord* 2004;78:131–140. [PubMed: 14706723]
14. Starkstein S. Apathy and withdrawal. *Int Psychogeriatr* 2000;12:135–137. [PubMed: 10937535]
15. Starkstein S, Leetjens A. The nosological position of apathy in clinical practice. *J Neurol Neurosurg Psychiatry* 2008;79:1088–1092. [PubMed: 18187477]

FIG. 1.
ROC curve for UPDRS item 4 based on diagnosis with the Apathy Evaluation Scale.

TABLE 1

Patient characteristics

Characteristic	Average (standard deviation)	Range
Age	67.98 (10.6)	30–90
Sex (% male)	189 M/112 F (63% male)	
Education	14.8 (3.2)	5–20
On antidepressant med (%)	31%	
UPDRS-motor score (on meds)	29.6 (12.2)	4–72
Disease duration (yrs)	8.1 (5.9)	0.5–36
Ave. item 4 score	1.14 (1.1)	0–4
Ave. AES score	13.7 (6.9)	0–31

TABLE 2

Operating characteristics compared across studies

	Current study cut-off = 1	Current study cut-off = 2	Pedersen study cut-off = 1	Pedersen study cut-off = 2	Starkstein study cut-off = 1	Starkstein study cut-off = 2
Sensitivity	81.0	52.3	80	70	91	73
Specificity	53.4	87.2	23	75	46	65
PPV	64.2	80.7	18	37	n/a	51
NPV	73.1	63.9	85	92	n/a	83

n/a = not reported in the study.