Citation for published version (APA):
Accuracy of screening instruments for detection of neuropsychiatric syndromes in Parkinson’s disease.

Authors: Pablo Martinez-Martin, MD, PhD;1 Albert F.G. Leentjens, MD, PhD;2 Jesus de Pedro-Cuesta, MD, PhD;1 Kallol Ray Chaudhuri, MD;3,4 Anette E. Schrag, MD;5 Daniel Weintraub, MD, PhD6

Centres:

(1) National Center of Epidemiology and CIBERNED, Carlos III Institute of Health, Madrid, Spain.

(2) Department of Psychiatry, Maastricht University Medical Center, Maastricht, the Netherlands

(3) National Parkinson Foundation International Centre of Excellence, King’s College London, London, UK.

(4) National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre and Dementia Unit at South London and Maudsley NHS Foundation Trust and King’s College London, London, UK.


(6) Department of Psychiatry and Departments of Neurology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, USA; and Philadelphia Veterans Affairs Medical Center, Philadelphia, USA
**Corresponding author:** Dr. P. Martinez-Martin  
National Center of Epidemiology  
Carlos III Institute of Health  
Avdª. Monforte de Lemos, 5  
28029 – Madrid  
Spain

Phone: +34 91 822 2618  
Fax: +34 91 387 7815  
E-mail: pmartinez@isciii.es

**Running title:** Screening for neuropsychiatric symptoms in PD

**Text:** 4698 words  
**Abstract:** 244 words

**Key words**  
Parkinson’s disease; Neuropsychiatric symptoms; Screening; Instruments; Accuracy.

**Conflict of interest:** Authors declare no conflict of interest with this manuscript.
Abstract

Parkinson’s disease includes neuropsychiatric manifestations such as depression, anxiety, apathy, psychosis, and impulse control disorders that frequently remain undeclared by patients and caregivers or undetected by doctors. Given their substantial impact on patients and caregivers, as well as the existence of effective therapies for some of these disorders, screening has an important role. Instruments for screening have a particular methodology for validation and their performance is expressed in terms of accuracy compared with the diagnostic criterion. The present study reviews the attributes of the screening instruments applied for detection of the abovementioned neuropsychiatric symptoms in Parkinson’s disease.

A “quasi-systematic” review was carried out on the basis of previous systematic reviews commissioned by the American Academy of Neurology and the Movement Disorder Society, with an update from a literature search (2005-2014).

For depression, eleven scales and questionnaires were shown to be valid for screening use in Parkinson’s disease. The recently developed Parkinson Anxiety Scale and the Geriatric Anxiety Inventory have shown satisfactory properties as screening instruments for anxiety and the Lille Apathy Rating Scale for detection of apathy in Parkinson’s disease. No scale covers the needs of a screening for psychosis in this setting and, therefore, a specific scale should be developed. The Questionnaire for Impulsive-Compulsive Disorders in Parkinson’s Disease and the Questionnaire for Impulsive-Compulsive Disorders-Rating Scale are valid for a comprehensive screening of impulse control disorders and the Parkinson’s Disease-Sexual Addiction Screening Test for hypersexuality.
Introduction

Neuropsychiatric symptoms (NPS) are an important component of the clinical manifestations of Parkinson’s disease (PD). Furthermore, they are highly prevalent and increase the impact of the disease on patients, their caregivers, and society as a whole.

Studies on prevalence indicate that major depression is present in 5 to 20% of patients and minor depression in an additional 1 to 30%; anxiety in about 40%, with panic disorder from 13 to 30%; apathy in 40%; psychotic manifestations in 10 to 40% of patients on dopaminergic medication and in 5 to 10% of those not receiving this treatment; and impulse control disorders (ICDs) in 14%.

In spite of their importance, NPS associated with PD frequently remain undetected by doctors and undeclared by patients, preventing interventions aimed at alleviating these symptoms.

It is for these reasons that routine screening has a role. Screening is carried out for detecting disorders before they are advanced, when earlier treatment may be associated with symptom improvement or a better prognosis, or while they remain hidden or confounded with another condition. However, screening is only effective if several criteria are met. Disease-related requirements for screening are that the disorder: (1) has a high enough prevalence or a long lasting preclinical phase, so that the chances of detection of the disorder are increased; (2) has potentially serious consequences for the patient if undetected and untreated; (3) has an acceptable and effective treatment or intervention; and (4) has a better course if treated after detection by screening as opposed to delayed treatment after routine clinical detection.

Screening instruments should not be used as definitive “diagnostic instruments”. When the gold standard for diagnosis (e.g., a detailed diagnostic interview) is too complex to apply, the use of a screening tool can act as a substitute in certain circumstances. Tests used for screening should have been validated before they are used, the validity being defined as the “extent to which its scores are related to an independent diagnostic criterion or gold standard”. In short, a screening test is valid if it efficiently
distinguishes healthy from sick individuals. However, this distinction is not easy, as many disorders or health conditions (e.g., disability) develop on a continuum between two extremes (healthy – sick) with an area of transition in which the true condition of the individual is unclear.

Jaeschke et al. provided guidelines for validation studies of screening instruments, including: (1) independent, blind comparison with the reference gold standard of diagnosis; (2) evaluation in an appropriate spectrum of patients, such as those usually seen in clinical practice; (3) application of the reference standard independently of the result with the index diagnostic test; and (4) test description in detail sufficient to allow replication or independent validation in a different group of patients. The non-observance of these rules gives room for bias in the outcomes of the study. Overestimation of test performance, for example, is typically observed in the following scenarios: (1) case-control studies; (2) when different reference tests are applied to subjects who are positive and negative for the assessed test; (3) when the reference test is not applied in a blind manner; or (4) when there is no description of criteria for the test.

The outcome of the screening in relation with the gold standard is represented in a contingency (2x2) table for a binary situation. The parameters indicative of the quality of a screening test are derived from the data displayed in this table (Supplementary material).

Usually, in the field of the NPS, screening instruments are rating scales that were initially designed for measurement of a continuous condition (e.g., the severity of a disorder) and later were compared with diagnostic criteria to determine a cut-off value using comparison to the gold standard. In clinical practice, identification of NPS in PD patients using extensive diagnostic scales is impractical, but the application of such scales on a subset of screened positive patients may reach higher performance standards. Cut-off points with appropriate balance between sensitivity and specificity can be useful for therapeutic decisions, an aspect related with the aforesaid disease-related requirements (3 and 4) for screening procedure.
The objective of this manuscript is to present the results of a “quasi-systematic” review (selection criteria were explicit, but the quality of the studies was not evaluated) on the accuracy of the instruments used for screening of common and significant neuropsychiatric disorders in PD: depression, anxiety, apathy, psychosis, and ICDs. Screening tools for cognitive-, sleep-, and sexual-related disorders will be not considered in the present study to allow a more detailed review of those focused on the abovementioned disorders.

**Methods**

We reviewed the articles published by the Movement Disorder Society (MDS) Task Forces on systematic reviews of scales depression, anxiety, apathy, and psychosis, as well as the corresponding reports of the Quality Standards Subcommittee of the American Academy of Neurology. In addition, a search was carried out in PubMed for papers published since 2005 to December 2014 including the following terms: (Parkinson’s disease) AND (screening) AND (neuropsychiatric symptoms OR depression OR anxiety OR apathy OR psychosis OR psychotic OR impulse control disorders). Also, articles from reference lists of core articles and authors’ files were reviewed.

In the search, 2296 articles with a related title were obtained; 146 were selected for abstract review; of which 69 were reviewed in detail (21 related with depression, 9 of them not included in previous reviews; 9 with anxiety, 4 of them not previously reviewed; 8 with apathy, 1 of them not included in previous reviews; 10 with psychosis, 4 of them not previously reviewed; and 21 with ICDs). Six articles were previous systematic reviews; 9 were related with diagnostic criteria; and 54 were studies potentially providing screening outcomes.

Selection criteria for articles review were: (1) systematic reviews or research articles; (2) focusing on testing screening properties of the instrument; (3) and providing, at least, cut-off point, sensitivity, and specificity values of the screening instrument for the aforementioned neuropsychiatric disorders in PD populations.
Sensitivity and specificity reflect the accuracy of the test and are intrinsic of the test, but standard values for these attributes have not been determined for generalized use. The predictive values indicate how the test performs in the population to which it is applied and are influenced by the prevalence of the condition in that population. Prioritizing efficiency of screening, a high sensitivity (for avoiding loss of affected cases) and positive predictive value (at a low “cost”) are favored. Subjects identified as “positive” by the screening test will be later assessed with the gold standard to confirm or refute the diagnosis.

1. Depression

Depressive symptoms are common in patients with PD, and a major contributor to quality of life. The characteristics of depression differ in some respects from those of depression in patients without PD, and new criteria have been proposed. In 2006, the Quality Standards Subcommittee of the American Academy of Neurology (AAN) reviewed the instruments for evaluation of depression, psychosis and dementia in PD. In this paper, the Subcommittee established that “validated criteria for depression, psychosis, and dementia in PD do not exist. Hence, the identification of clinically relevant screening and diagnostic tools for depression, psychosis, and cognitive decline validated specifically in the PD population is necessary”. They reported the cut-off points providing the highest diagnostic accuracy of three frequently used scales: Beck Depression Inventory-I (BDI-I), Hamilton Depression Rating Scale (HAMD-17) and Montgomery-Asberg Depression Rating Scale (MADRS) (Table 1). The authors advice that HAMD-17 and MADRS appear to perform better than the BDI, but the studies furnishing the referred data were underpowered to determine superiority. The final recommendation was that BDI-I and HAMD (Level B) and MADRS (Level C) should be considered for screening of depression in PD.

Also in 2006, Mondolo and coworkers published the validity results of the Geriatric Depression Scale (GDS) and the Depression subscale of the Hospital Anxiety and Depression Scale (HADS-D) for screening of depression in PD. Both scales showed good validity for screening of mood at a cut-off 10/11, although the results were better for the HADS as a whole (Table 1).
The topic was also reviewed in 2007 by the MDS Task Force on Depression Rating Scales in PD and the conclusion of this review was that 6 of the 10 reviewed scales were valid for screening purposes of depression in PD. The HAMD, MADRS, and GDS, but also BDI and HADS were judged to be valid for screening of depression in PD. Those studies meeting the inclusion criteria for the present review are shown in the Table 1.

The GDS-15 showed higher sensitivity and similar positive predictive value compared to the Patient Health Questionnaire (PHQ-9) for major and minor depression in a study on 214 PD patients, but cut-off points were not displayed. Recently, Williams et al. 2012, compared nine scales for screening of depression in PD: BDI-II, Center for Epidemiologic Studies Depression Rating Scale-Revised (CESD-R), GDS-30, Inventory of Depressive Symptoms–Patient (IDS-SR), PHQ-9, Unified Parkinson’s Disease Rating Scale-Depression (UPDRS-Depression), HAMD-17, Inventory of Depressive Symptoms-Clinician (IDS-C), and MADRS. The conclusion of their analysis was that the GDS-30 has several advantages on the other scales for depression screening in PD, although all of the studied scales were valid to this purpose, with exception of the UPDRS-Depression. Cut-off score and parameters of screening performance for each scale and comparison with previous studies are shown in the Table 1 of this article.

To summarize, the high prevalence of depression in PD and the strong impact of this mood disorder on patients’ disability and quality of life have made it objective target for research and treatment. As a consequence, a number of studies have been carried out to investigate the instruments for detection and evaluation of depression. A considerable proportion of them show satisfactory properties for screening of this disorder in PD patients showing high sensitivity (≥85%) combined with relatively high specificity (≥ 75%) and, therefore, are deemed suitable for screening use in PD. These scales are: GDS (15, 20, and 30 items), HADS-D, HAMD-17, IDS-C, MADRS, UK National Institute for Health and Clinical Excellence (NICE) screening questions, PHQ-9, WHO-5, and Zung’s SDS (Table 1). Not many studies have directly compared the clinimetric properties of depression scales with the intention to assess whether one specific scale can be preferred over others. The largest of the few studies in this field was the one by Williams et al. who compared nine depression rating scales for their
screening performance. They reported that all of the included scales had good sensitivity and specificity, and that their performance did not differ much. As such, it is not possible to recommend the use of a specific scale or scales when screening for depression in PD patients, and the choice of instrument will depend largely on matters of convenience, time and cost. In clinical and research settings, the decision to routinely use an instrument is probably more important than the choice of instrument.

2. Anxiety

Anxiety disorders in PD have received less attention than depressive disorders, even though there is evidence that they are equally or more prevalent than depressive disorders, affecting up to 40% of PD patients. Anxiety is associated with increased subjective motor symptoms, more severe gait problems and dyskinesias, freezing and on/off fluctuations and a reduced quality of life. In addition, there is a large overlap with depression: 36 to 65% of PD patients that suffer from an anxiety disorder also suffer from a depressive disorder.

The properties of the instruments applied for assessment of anxiety in PD were reviewed by another MDS Task Force. Six scales were reviewed: the Beck Anxiety Inventory (BAI), the Hospital Anxiety and Depression Scale and its Anxiety subscale (HADS-A), the Zung self-rating Anxiety Scale (SAS) and the Anxiety Status Inventory (ASI, which is the observer rated version of the SAS), the Spielberger State Trait Anxiety Inventory (STAI), the Hamilton Anxiety Rating Scale (HARS), and section 5 (anxiety) of the Neuropsychiatric Inventory (NPI).

None of these scales could be recommended and were classified as “suggested” scales, following the assessment criteria of the MDS, since at the time of the review these scales had not been specifically validated in PD. Therefore, the MDS Task Force recommended supplementary studies with these scales to obtain information about their validity in PD patients. Subsequently, a validation study was performed that included the most commonly used scales for screening for anxiety: the BAI, HADS and HARS. It was demonstrated that whereas some clinimetric properties of these scales were acceptable, such as inter-rater and test-retest reliability, known groups validity and score
distributions, other properties were not satisfactory, such as convergent validity, factorial structure and notably construct validity (Table 2). Moreover, none of the scales had a good fit to the Rasch model in item response theory (IRT) analyses. Consequently, none of these scales were considered appropriate screening instruments for anxiety in PD.

Later, the same group demonstrated the particular profile of anxiety disorders in PD and the questionable validity of the DSM-IV criteria in capturing this profile. This fact challenged the content and construct validity of the existing anxiety rating scales in PD patients. Therefore, the development of a new specific scale for assessment of the anxiety disorders in PD was proposed and developed, funded by the Michael J. Fox Foundation for Parkinson’s Research. The Parkinson Anxiety Scale (PAS) includes subscales for persistent and episodic anxiety as well as for avoidance behaviour, and exists in both a patient- and a clinician-rated version. Although not specifically designed for screening or diagnosis, its capacity to this purpose was tested against the Mini International Neuropsychiatric Inventory (MINI) section for anxiety (Table 3). This scale appeared to have superior sensitivity and specificity to use as screener or diagnostic scale for anxiety in PD, as well as a good fit to the Rasch model. Tables 2 and 3 list the clinimetric parameters at the optimal cut-off (highest sum of sensitivity and specificity), but the cut-off may be adjusted to better suit the purpose of evaluation, screening or diagnosis.

Another scale that was subsequently validated in PD patients is the Geriatric Anxiety Inventory (GAI). The authors report good test-retest reliability and known group validity. The optimal cut-off of for diagnosing any anxiety disorder was 6/7, with a sensitivity of 0.88 and a specificity of 0.86 (Youden index 1.74).

Anxiety and anxiety rating scales in PD have only recently received proper attention. The PAS (Table 3) is the only instrument specifically designed for use in PD patients, showing satisfactory clinimetric properties and good divergent validity with depression rating scales. Subscales are used to assess persistent (generalized and social) anxiety, episodic anxiety (panic) and avoidance behavior. In addition to the PAS, one generic anxiety rating scale, the GAI, has recently been validated and found to have acceptable screening properties. Other validated scales, the HADS, BAI and HARS, showed less
favourable attributes for validity and fit to the Rasch model (Table 2). These commonly used generic scales focus on different symptoms, with the HARS focusing more on symptoms of generalized anxiety and the BAI more on symptoms of panic. Therefore, they identify different patient groups and have low convergent validity. Also, they are less well able to separate depression from anxiety.

Contrary to depression rating scales, not many validation studies of anxiety rating scales have been performed in PD patients, and conclusions in this review are often based on one single study, without replication studies supporting the findings. Additional research on the clinimetric properties of the reviewed scales is needed.

In summary, the PAS and the GAI are considered useful instruments for screening of anxiety in Parkinson’s disease.

3. Apathy

Apathy is defined as a lack of motivation, characterized by reduced goal directed behaviour, reduced goal directed cognitive activity and reduced emotional responsivity (or anhedonia). It is frequently present in depression and dementia, but has also been recognized as an independent syndrome.\(^{56}\) Another MDS task force reviewed the scales applied for assessment of apathy in PD.\(^{57}\) Six scales were reviewed: the Apathy Evaluation Scale (AES),\(^{58}\) the Apathy Scale (AS, an adapted version of the AES),\(^{59}\) the Lille Apathy Rating Scale (LARS),\(^{60}\) section 7 (apathy) of the NPI,\(^{49}\) and the Apathy Inventory (AI).\(^{61}\)

For three scales (AES, AS, and LARS), cut-off scores have been recommended to screen for “clinically relevant apathetic symptoms” (37/38; 13/14; and 16/17, respectively). Although (proposed) consensus diagnostic criteria for apathy have recently been formulated,\(^{56}\) only two studies have assessed the validity of apathy rating scales against these criteria.\(^{62}\) In a study\(^{62}\) on 130 PD patients, of which 42% with apathy, the LARS showed a sensitivity 80% with specificity 90% for a cut-off -14/-13 to detect apathy as diagnosed with the Starkstein and Leentjens’ criteria.\(^{63}\) The other scales lack evidence on screening properties in PD patients. Good concurrent validity of
the LARS with proposed diagnostic criteria for apathy was also reported by Drijgers et al., who found a 81% agreement between the classification on the basis of the proposed criteria and a cut-off of -16/-17 on the LARS. Other apathy rating scales including the AI, AS and AES, have not been validated against the proposed diagnostic criteria and hence cannot be recommended as screening tool since information on sensitivity and specificity is lacking.

While it is generally recognized that apathy is an important non-motor symptom with a potentially great impact on daily functioning and quality of life, it is still an evolving concept that has no firm base in any classification system. In the DSM5, apathy is only listed as a subtype of 'personality change due to a general medical condition', as well as a symptom of several unrelated disorders (APA, 2013). This uncertain status of apathy as symptom or syndrome hampers research. For the same reasons as in the assessment of depressive or anxiety syndromes, the authors recommend the use of the (proposed) diagnostic criteria alongside a rating scale for evaluation of apathy in research settings. The LARS was specifically developed for PD patients and is to date the best validated scale. It is worthwhile studying the clinimetric properties of other apathy scales since the LARS is extensive and time consuming, and therefore less suited for clinical use.

4. Psychosis
Psychosis is a major challenge in Parkinson’s disease and a key non-motor symptom. These include hallucinations which are usually a result of spontaneous aberrant perceptions, misinterpretations of real perceptual stimuli usually called illusions, and delusions. In clinical practice, visual hallucinations appear to be the most frequent psychotic manifestation, with cross-sectional studies reporting visual hallucinations in approximately 25% of chronically dopaminergic drug treated PD subjects. Development of intrusive and complex visual hallucinations constitute one primary risk factor for nursing home placement of the patient. Nonetheless, clinical studies indicate psychosis is often not declared in clinical practice and as such there needs to be a high degree of awareness. Assessment of psychosis is complex and needs considerable clinical expertise. Furthermore, patients with severe psychosis are often in hospital or institutions and could have considerable comorbid problems (e.g., cognitive
impairment) and as such application of specific psychosis based tools could also be problematic.

There are two relevant reviews on this topic by the Quality Standards Subcommittee of the AAN\textsuperscript{22} and the MDS task force,\textsuperscript{71} respectively. The AAN Quality Standards Subcommittee stated that a gold standard for diagnosis of psychosis in PD was not available and, in addition, the only study testing a specific scale, the Parkinson Psychosis Rating Scale (PPRS),\textsuperscript{72} was a Class IV study. Therefore, evidence was insufficient to conclude the quality of this scale as a screening tool for psychosis in PD, and no recommendation was made.

In the MDS task force review, the authors identified three scales that include specific probes to test the presence of a psychotic disorder. When that probe is positive, items for rating the severity of the psychotic disturbances are applied, but when the response to the probe is negative the corresponding items are skipped and the next probe is tested. The three identified scales were the Parkinson Psychosis Questionnaire (PPQ),\textsuperscript{73} the Rush Hallucination Inventory,\textsuperscript{74} and the Neuropsychiatric Inventory (NPI).\textsuperscript{49}

The PPQ\textsuperscript{73} assesses drug-induced psychotic symptoms, hallucinations and illusions, delusions, and orientation. Compared with DSM-IV criteria and considered as “positive/negative” test for the presence of psychosis (n=50), showed a sensitivity 100% with a specificity 92%. The PPQ has been recently used in several studies on PD.\textsuperscript{75-77}

The Rush Hallucination Inventory\textsuperscript{74} and the NPI\textsuperscript{49} have not been formally tested in regard to its accuracy for identifying neuropsychiatric disorders in PD.

As per the MDS task force, content validity of the reviewed scales was found to be insufficient, as none of them properly includes the complete variety of psychosis manifestations that can be present in PD. Consequently, recommendations and conclusions related to their use were similar to those of the ANA Quality Standards Subcommittee\textsuperscript{17} and none of these scales could be recommended for screening of psychotic disorders in PD.
To summarize, for psychosis, none of the existent scales covers the variety of psychotic disorders potentially affecting PD patients or has been formally tested with regard to its screening properties in this setting. Therefore, the development of a new, specific scale including all relevant aspects related with this construct and to be applied for screening and severity assessment has been proposed.\textsuperscript{71}

5. Impulse control disorders (ICDs) and related behaviors

Impulse control disorders (ICDs) and related behaviors (e.g., dopamine dysregulation syndrome, punding and hobbyism) are relatively frequent in PD. They can be very disruptive and have serious consequences for patients and their relatives, but also can remain undeclared to doctors until the impact on the psychosocial aspects has reached an extreme level. Some screening instruments and rating scales are available, both for a wide range of ICDs or focused on a specific disorder. An important issue to consider when assessing ICD symptoms is that there may be discordance between patient and informant reporting.\textsuperscript{78}

The Questionnaire for Impulsive-Compulsive Disorders in Parkinson’s Disease (QUIP)\textsuperscript{79} is the only validated specific instrument for screening of ICBs and other compulsive behaviors in PD. It is a self-administered, 12-item questionnaire with three sections: ICDs, compulsive behaviors (hobbyism, punding, walkabout), and compulsive medication use. Each section includes at least an introductory question with definition of the screened disorders and examples. Existing criteria\textsuperscript{80-84} were used for testing the screening capacity of the respective sections. The recommended cut-off scores for each questionnaire component appear in the Table 4.

A rating scale for ICD severity (frequency) assessment was derived from the QUIP, the Questionnaire for Impulsive-Compulsive Disorders in Parkinson’s Disease-Rating Scale (QUIP-RS).\textsuperscript{85} This scale has also four main questions (about thoughts, urges/desires, and behaviors) related to compulsive gambling, buying, eating, and sexual behavior, and another three focused on related disorders (medication use, punding, and hobbyism). For each disorder the total score runs from 0 to 16; the total for the four ICDs, from 0 to 64; and the total QUIP-RS ranges from 0 to 112. As for the QUIP, the
diagnostic validation was based on a semi-structured interview including existing or proposed criteria for each disorder.\textsuperscript{80-84} Results concerning diagnostic ability of the QUIP-RS are shown in the Table 4.

Other instruments for the comprehensive screening of ICDs and related behaviors in PD have been not validated in PD patients specifically (e.g., the Minnesota Impulsive Disorders Interview)\textsuperscript{86} or their diagnostic attributes have not been specifically tested.

Some of the tools potentially useful for screening of a particular ICD and used at least in a study of PD patients are considered next. The South Oaks Gambling Screen (SOGS)\textsuperscript{87} has been extensively used in non-PD and PD populations, but to our knowledge there are no detailed studies on its sensitivity and specificity in PD. The McElroy criteria for compulsive buying,\textsuperscript{88} used as the gold standard for diagnosis of buying disorder and widely applied in studies on PD populations, as well as the Compulsive Buying Scale\textsuperscript{89} lack specific testing as screening tools in PD patients.

The Shorter version of the Sexual Addiction Screening Test for PD (PD-SAST)\textsuperscript{90} is an adaptation (short form) of the Sexual Addiction Screening Test\textsuperscript{91,92} to PD patients. It comprises 5 items answered Yes/No by patients. The cut-off point for a person to be at risk of hypersexual disorder was 2/3, with a sensitivity 100% and specificity 93% for a diagnosis based on DSM-IV-TR clinical criteria.\textsuperscript{90}

The Clinician Punding Criteria and Rating Scale\textsuperscript{93} has been used in several studies on punding in PD and has been compared with the psychiatric diagnosis\textsuperscript{94} with good results. However, the screening properties of the scale are not sufficiently determined. This is also the case for the Punding Rating Scale,\textsuperscript{95} a modified version of the Clinician Punding Criteria and Rating Scale\textsuperscript{93} for use in cocaine addicts. The Saving Inventory-Revised\textsuperscript{96} is applied to assess hoarding disorder, but has been used only occasionally in PD patients and has not been explored as a screening tool in this setting.

A large number of studies have been conducted and instruments applied to ascertain the frequency and severity of ICDs and symptoms in different PD cohorts. Two instruments, the QUIP and QUIP-RS,\textsuperscript{79,85} have been found to be useful for a comprehensive screening of ICDs and related behaviors in PD populations. It can be advantageous to screen for multiple ICDs simultaneously, as these disorders are
frequently co-morbid and having co-morbid ICDs is associated with increased depression and worse quality of life. As a screening instrument the QUIP has limitations, including that nearly 40% of patients without an ICD diagnosis screen positive for an ICD, although it is possible that some of these “false positives” are experiencing subsyndromal ICD symptoms and warrant monitoring. For screening instruments assessing only a single ICD, the PD-SAST (for hypersexuality) has satisfactory psychometric properties but requires further testing.

In conclusion, only the QUIP and QUIP-RS for a comprehensive screening of ICDs and the PD-SAST for hypersexuality have been sufficiently demonstrated to be valid screening instruments in the setting of PD.

Discussion
This review outlined the psychometric properties of instruments used for screening of the most important neuropsychiatric disorders in PD. An additional important consideration with assessment of NPS, particularly with ICD symptoms or psychosis, is whether the information should be obtained from the patient or a knowledgeable informant (KI), and the responses obtained will depend on the amount of insight that the patient has into his/her behaviors, as well as the amount of information that the patient shares with a KI. There are limited data comparing patient to informed other ratings in PD. Also, when using instruments and diagnostic criteria developed for the general population and subsequently applying them to PD patients, it is important to note that the phenomenology of NPS in PD may differ in important ways from those in the general population. Clearly, there is a need for further research to improve screening instruments and rating scales, particularly for ICDs in PD, whether that be reliability and validity testing of existing instruments, testing of further psychometric properties or the development of new ones specifically for use in this population.

It is also important to emphasize that the objective of a screening procedure is to identify potential (candidate) cases of a disorder, but not to make a definitive diagnosis. Efficient screening requires high sensitivity to avoid loss of cases with the disorder and this can be achieved by decreasing the proportion of false negatives (c/a+c) (Supplementary material) by lowering of the cut-off value. Nonetheless, a balance with
an acceptable specificity (a relatively low proportion of false positives, b/b+d) must be kept. A final diagnosis can however only be made based on clinical examination and diagnostic criteria.

**Limitations.** For this review, the selection of papers was based on a literature search in only one database (PubMed), restricted to a decade (2005-2014) and to articles providing empiric data about screening of NPS other than cognitive, sleep and sexual disorders in PD. These are important limitations, but the paper’s purpose is to outline the state of the current literature on basic properties, accessible to professionals with activity in that setting. However, we have not attempted to review and assess the content validity and psychometric validity of the scales evaluated in this paper which is exclusively focused on their screening properties.

In conclusion, there are empirical data of screening validity for instruments assessing depression, anxiety, apathy, and impulse control disorders (comprehensive and hypersexuality) in PD populations, but not for psychosis. There is room for future studies in the field of NPS screening for testing their performance compared to diagnostic criteria or to develop specific instruments with appropriate validity for detecting NPS in PD patients.

**Authors’ roles**
1) Research project: A. Conception, B. Organization, C. Execution;
3) Manuscript: A. Writing of the first draft, B. Review and Critique.

PMM – 1A, 1B, 1C; 2A, 2C; 3A
AFGL – 1C; 2C; 3B
JPC – 1C; 2C; 3B
KRC – 1C; 2C; 3B
AES – 1C; 2C; 3B
DW – 1C; 2C; 3B
Financial disclosures

P. Martinez-Martin has received honorarium for lecturing or participation in symposia from Movement Disorder Society, Complutense University, and Editorial Viguera, as well as honorarium for consultancies from AbbVie (for an epidemiological study) and Mundipharma (Steering committee of a clinical trial). Research support from Parkinson’s UK and Parkinson’s Disease Non-Motor Group.

A.F.G. Leentjens has received honoraria from Elsevier Inc (as Editor-in-Chief of the Psychsom Res) and royalties from Reed Elsevier, De Tijdstroom, and Van Gorcum publishers. Grants from Michael J Fox Foundation and Stichting Parkinsonfonds.

J. de Pedro-Cuesta is member of the EU Advisory Board for the Joint Programme Research on Neurodegenerative Disorders. Has received honoraria from Fundación Ramon Areces and grants from Biogen Spain.

K. Ray Chaudhuri has received academic/educational funding from Parkinson’s UK, NIHR, the Movement Disorder Society, Parkinson’s UK, European Union and UCB, and has received honoraria for sponsored symposiums from UCB, AbbVie, Britannia, US Worldmeds, Otsuka, Medtronic, Zambon and academic funding from the Movement Disorder Society, Parkinson’s UK, European Union, NIHR and acted as a consultant for Abbvie, UCB, Britannia, Medtronic, and Mundipharma.

A.E. Schrag declares stock ownership (Astra Zeneca since 7/2014), consultancies with Medtronic, and royalties with Oxford University Press (Rating Scales in PD, 1/2013). She has received grants from EU Commission, Parkinson’s UK, GE Healthcare, and ESRC.

D. Weintraub has received research funding or support from Michael J. Fox Foundation for Parkinson’s Research, National Institutes of Health, Novartis Pharmaceuticals, Department of Veterans Affairs, Avid Radiopharmaceuticals, and Alzheimer’s Disease Cooperative Study; honoraria from AbbVie, Biotie, Teva Pharmaceuticals, Otsuka, UCB, Clintrex LLC, and the CHDI Foundation; license fee payments from the University of Pennsylvania for the QUIP and QUIP-RS; royalties from Wolters Kluweland; and fees for legal consultation for lawsuit related to antipsychotic prescribing in a patient with Parkinson’s disease.
References


13 - Gallagher DA, Lees AJ, Schrag A. What are the most important nonmotor symptoms in patients with Parkinson's disease and are we missing them?. Mov Disord. 2010;25(15):2493-500.


63 - Starkstein SE, Leentjens AF. The nosological position of apathy in clinical


76 - Sawada H, Oeda T, Umemura A, Tomita S, Hayashi R, Kohsaka M, et al. Subclinical elevation of plasma C-reactive protein and illusions/hallucinations in


92 - Carnes P. Contrary to love: helping the sexual addict. Minneapolis: Hazelden; 1989.


