Published Ahead of Print on January 13, 2016 as 10.1212/WNL.0000000000002350

Accuracy of clinical diagnosis of Parkinson disease

A systematic review and meta-analysis

Giovanni Rizzo, MD
Massimiliano Copetti, PhD
Simona Arcuti, PhD
Davide Martino, MD
Andrea Fontana, MSc
Giancarlo Logroscino, MD

Correspondence to Dr. Logroscino: giancarlo.logroscino@uniba.it

ABSTRACT

Objective: To evaluate the diagnostic accuracy of clinical diagnosis of Parkinson disease (PD) reported in the last 25 years by a systematic review and meta-analysis.

Methods: We searched for articles published between 1988 and August 2014. Studies were included if reporting diagnostic parameters regarding clinical diagnosis of PD or crude data. The selected studies were subclassified based on different study setting, type of test diagnosis, and gold standard. Bayesian meta-analyses of available data were performed.

Results: We selected 20 studies, including 11 using pathologic examination as gold standard. Considering only these 11 studies, the pooled diagnostic accuracy was 80.6% (95% credible interval [CrI] 75.2%–85.3%). Accuracy was 73.8% (95% CrI 67.8%–79.6%) for clinical diagnosis performed mainly by nonexperts. Accuracy of clinical diagnosis performed by movement disorders experts rose from 79.6% (95% CrI 46%–95.1%) of initial assessment to 83.9% (95% CrI 69.7%–92.6%) of refined diagnosis after follow-up. Using UK Parkinson's Disease Society Brain Bank Research Center criteria, the pooled diagnostic accuracy was 82.7% (95% CrI 62.6%–93%).

Conclusion: The overall validity of clinical diagnosis of PD is not satisfying. The accuracy did not significantly improve in the last 25 years, particularly in the early stages of disease, where response to dopaminergic treatment is less defined and hallmarks of alternative diagnoses such as atypical parkinsonism may not have emerged. Misclassification rate should be considered to calculate the sample size both in observational studies and randomized controlled trials. Imaging and biomarkers are urgently needed to improve the accuracy of clinical diagnosis in vivo.
Neurology® 2016;86:1–11

GLOSSARY

AD = Alzheimer disease; **DLB** = dementia with Lewy bodies; **ET** = essential tremor; **MSA** = multiple system atrophy; **NPV** = negative predictive value; **PD** = Parkinson disease; **PDD** = Parkinson disease dementia; **PPV** = positive predictive value; **PSP** = progressive supranuclear palsy; **UKPDSBRC** = United Kingdom PD Society Brain Research Center; **VaE** = vascular encephalopathy.

The correct diagnosis of Parkinson disease (PD) is important for prognostic and therapeutic reasons but also for clinical, pharmacologic, and epidemiologic studies. Despite advances in neuroimaging and genetics, the diagnosis of PD remains primarily clinical. Definite diagnosis can be obtained only pathologically.[1] Misdiagnosis is common. The diagnosis can be changed in many patients with parkinsonism after a follow-up of a few years.[2]

Accordingly, PD diagnosis in the early stages remains relatively inadequate. All current diagnostic criteria consider signs and symptoms emerging in the later stage of disease. The difficulties increase when elderly patients are evaluated. Indeed, the false-positive cases reported in the literature include patients with conditions being more frequent in older age such as essential tremor (ET), progressive supranuclear palsy (PSP), multiple system atrophy (MSA), Alzheimer disease (AD), and vascular encephalopathy (VaE). Furthermore, diagnosis in elderly people is complicated by the increasing evidence of mixed pathology.[3,4]

Supplemental data at Neurology.org

From the Department of Clinical Research in Neurology (G.R., S.A., G.L.), University of Bari, Tricase; Department of Biomedical and Neuromotor Sciences (G.R.), University of Bologna; Unit of Biostatistics (M.C., A.F.), IRCCS "Casa Sollievo della Sofferenza," San Giovanni Rotondo, Italy; Department of Neurology (D.M.), King's College NHS Foundation Trust; Department of Neurology (D.M.), Queen Elizabeth Hospital, Lewisham and Greenwich NHS Trust, London, UK; and Department of Basic Medical Science (G.L.), Neuroscience and Sense Organs, University of Bari, Italy.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

A number of studies have evaluated the accuracy of clinical diagnosis in PD, which were variably mentioned in previous systematic reviews reporting descriptive conclusions about suboptimal diagnostic accuracy.[5,6] A meta-analysis of their results has not been made. Our aim was to perform an up-to-date systematic review of these studies and to meta-analyze, in a Bayesian framework, the available data, overall and separately according to variables affecting the diagnosis, i.e., setting of recruitment, expertise of physicians, and diagnostic gold standard.

METHODS

Systematic review. We followed Preferred Reporting Items for Systematic reviews and Meta-Analyses 2009 guidelines for systematic review and meta-analysis.[7] We performed electronic searches of Medline and Embase databases using a combination of a number of medical subject heading and free-text terms (e.g., Parkinson's disease, parkinsonism, diagnostic accuracy, sensitivity, specificity, positive predictive value, negative predictive value) from 1988, date of publication of the United Kingdom PD Society Brain Research Center (UKPDSBRC) diagnostic clinical criteria,[8] to August 25, 2014. Duplicates were eliminated and all relevant articles were retrieved. No restriction was placed on language. We excluded abstracts and book chapters. We carefully reviewed the reference list of articles for additional articles missed in the research. We included articles if they reported any diagnostic parameters or raw data, specifically regarding the clinical diagnosis of PD. We excluded articles providing diagnostic accuracy fully or partially based on advanced imaging or neurophysiologic techniques (i.e., not those routinely used in clinical practice), except for those studies also reporting usable data about the diagnosis without the same techniques. We excluded articles with not properly defined diagnostic categorization in the test diagnosis, e.g., parkinsonian syndrome instead of PD. Our first aim was to perform the main analysis only on those studies that used pathologic examination as gold standard. However, we conducted additional analyses including other studies with different gold standards, to provide data as informative as possible, e.g., from community-based studies. Achievable diagnostic parameters should fully or partially include sensitivity (proportion of patients with PD who had initial diagnosis of PD); specificity (proportion of patients without PD who had no initial diagnosis of PD); positive predictive value (PPV) (proportion of patients with initial diagnosis of PD who actually had the disease); negative predictive value (NPV) (proportion of people with initial diagnosis of non-PD who did not actually have the disease); and diagnostic accuracy (proportion of all correct diagnoses).

Two authors (G.R. and D.M.) independently performed the literature search, selected all potentially relevant articles, screened the full texts, and extracted data from the eligible studies. Disagreements were resolved by asking the opinion of a third reviewer (G.L.). When relevant information was missing, we contacted study authors by e-mail.

Data preparation. We analytically read the methods of the studies in order to evaluate the difference in the study setting, type of test diagnosis, and gold standard in the attempt to define different categories of comparable studies. As for the setting, we identified 3 subtypes: clinic-based, community-based, and clinic-based, on selected patients with uncertain diagnosis (essentially on tremulous patients). As for the type of test diagnosis, we identified 4 subtypes: a clinical diagnosis performed in all or most cases by nonexperts in movement disorders, such as general neurologists, geriatricians, or general practitioners, at a variable time of the disease course and here called clinical diagnosis mainly by nonexperts (we also included in this category those studies reporting a mixed evaluation, by experts and nonexperts, but without providing the exact proportion); an initial (first visit) diagnosis performed by movement disorders experts in all or the most cases and here called initial clinical diagnosis by experts; a refined clinical diagnosis, i.e., by experts after follow-up (second or subsequent visit) and using all available clinical data and here called refined clinical diagnosis by expert; and a diagnosis based on the UKPDSBRC diagnostic clinical criteria, here called UKPDSBRC clinical criteria. As for the type of gold standard diagnosis used, we identified 3 subtypes: pathologic diagnosis, refined clinical diagnosis by expert, and UKPDSBRC clinical criteria. We performed Bayesian meta-analyses for all studies and separately for each study subtype according to setting, test diagnosis, and gold standard diagnosis. Because pathologic diagnosis is the most robust gold standard for PD diagnosis, we first performed meta-analyses specifically focusing on studies that used this gold standard. This should be considered the main analysis of this work.

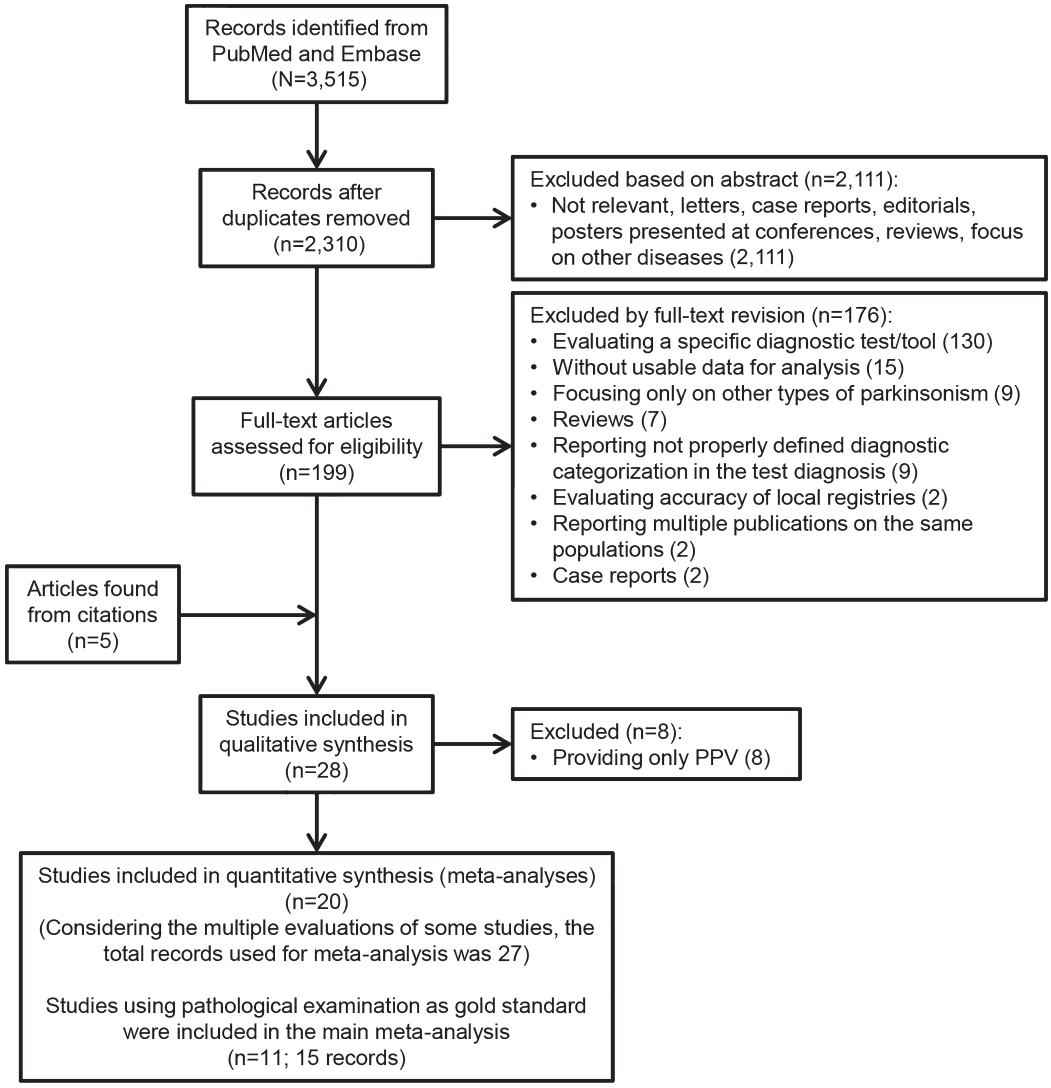
Some studies reported accuracy in different populations or for different types of diagnosis; therefore, these studies were included in the meta-analysis with more than one record. Given that PPV and NPV are substantially conditioned by the different proportion of patients with PD or atypical parkinsonism evaluated in each specific setting, and therefore less generalizable, we only meta-analyzed accuracy, sensitivity, and specificity values.

Bayesian meta-analysis. Bayesian methods offer flexibility, which allows the approach to be extended to consider complex likelihood functions other than normal. Bayesian methods might also perform better and provide robust credible intervals in applications with a relatively small number of studies. Details on the statistical analysis are reported in the supplemental data on the *Neurology*® Web site at Neurology.org.[9,10]

RESULTS

Selected studies. Out of 3,515 studies that were identified, we included 28 studies in this systematic review, from 30 populations and including 7,662 patients (see figure 1 for details). Details of the studies are provided in table e-1. Among these 28 studies, 16 were clinic-based (25 records),[3,11–25] 2 clinic-based on selected patients with uncertain diagnosis (3 records),[26,27] and 10 community-based (12 records).[2,28–36] Thirteen studies used a pathologic examination as gold standard (22 records),[3,11–14,16,17,19,21–23,25,30] all clinic-based except one[30]; 11 studies used a refined clinical diagnosis by experts (12 records)[2,15,18,20,24,26,27,31,32,35,36]; and 5 studies, all community-based, used the UKPDSBRC clinical criteria (6 records).[2,28,29,33,34] The test diagnosis was a clinical diagnosis formulated mainly by nonexperts in 10 studies (12 records),[3,11,12,14,16,21–23,28,29] including diagnoses performed at a variable time of the disease

Figure 1 Flowchart of electronic search



Selection of included studies. PPV = positive predictive value.

course, mostly after several years of illness (pathologic examination as gold standard in 8 studies).[3,11,12,14,16,21–23] An initial clinical diagnosis by experts was reported in 10 studies (12 records),[2,14,15,24–27,31,32,35] 2 using pathologic examination as gold standard.[14,25] A refined clinical diagnosis by experts after follow-up was reported in 9 studies (9 records),[2,11,13,14,17,19,21,25,30] 8 of which used pathologic examination as gold standard[11,13,14,17,19,21,25,30]; mean disease duration at last evaluation, available for 6[11,13,17,19,25,30] out of these 8 studies, was 10.2 years, with range from 3.6 to 13.8 years. A diagnosis made according the UKPDSBRC clinical criteria was reported in 7 studies (7 records),[12,17,18,20,33,34,36] 2 using pathologic examination as gold standard (criteria were applied using clinical information of patients with disease duration of 11.9[12] and 13.8 years[17]).

We extracted complete information about true-positive and true-negative rates and false-positive and false-negative rates for 26 records from 19 studies,[2,11–14,17,19,22,23,25–27,29–35] and partial information for 14 records from 11 studies (PPV for 13 records and sensitivity for one).[3,12,15–18,20,21,24,28,36] The studies reporting only PPV were excluded from quantitative analysis.[12,15–18,20,21,24,28,36] Accordingly, meta-analysis was performed on 27 records from 20 studies (figure 1 and table 1).[2,3,11–14,17,19,22,23,25–27,29–35]

Bayesian meta-analysis. Pooled diagnostic parameters (diagnostic accuracy, sensitivity, specificity) are reported in figures 2–4, overall and separately according to the different studies' categories.

Qualitative review on the misdiagnosis issue. Specific information about the correct diagnosis of false-positives or false-negatives among the selected studies

Table 1 Records from the studies included in the meta-analyses

Studies	Sample, n	Time of recruitment	Setting and population	In vivo diagnosis	Gold standard	Sensitivity, %	Specificity, %	Accuracy, %
Rajput et al.,[11] *Can J Neurol Sci* 1991	59	1968–1990	Clinic-based	Clinical diagnosis mainly by nonexperts	Pathologic diagnosis	90.3	46.4	69.5
Rajput et al.,[11] *Can J Neurol Sci* 1991	59	1968–1990	Clinic-based	Refined clinical diagnosis by experts	Pathologic diagnosis	100	64.3	83
Hughes et al.,[12] *J Neurol Neurosurg Psychiatry* 1992	100	1987–1990	Clinic-based	UKPDSBRC clinical criteria	Pathologic diagnosis	96	33.3	81
Hughes et al.,[13] *Neurology* 1992	100	1987–1990	Clinic-based	Refined clinical diagnosis by experts	Pathologic diagnosis	68.4	83.3	72
Hughes et al.,[3] *Arch Neurol* 1993	100	1986–1990	Clinic-based	Clinical diagnosis mainly by nonexperts	Pathologic diagnosis	90	—	—
Litvan et al.,[14] *Arch Neurol* 1998	105	NA	Clinic-based	Initial clinical diagnosis by experts	Pathologic diagnosis	73.3	85.6	83.8
Litvan et al.,[14] *Arch Neurol* 1998	105	NA	Clinic-based	Clinical diagnosis mainly by nonexperts	Pathologic diagnosis	93.3	76.7	79
Litvan et al.,[14] *Arch Neurol* 1998	105	NA	Clinic-based	Refined clinical diagnosis by experts	Pathologic diagnosis	80	92.2	90.5
Hughes et al.,[17] *Neurology* 2001	100	1996- 1998	Clinic-based	UKPDSBRC clinical criteria	Pathologic diagnosis	90	30	84
Hughes et al.,[19] *Brain* 2002	143	1990–1999	Clinic-based	Refined clinical diagnosis by experts	Pathologic diagnosis	91.1	98.5	94.4
Horvath et al.,[22] *Brain Pathol* 2013	261	1914–2010	Clinic-based	Clinical diagnosis mainly by nonexperts	Pathologic diagnosis	90.8	37.8	70.9
Joutsa et al.,[23] *Parkinsonism Relat Disord* 2014	110	2000–2012	Clinic-based	Clinical diagnosis mainly by nonexperts	Pathologic diagnosis	89.2	57.8	76.4
Adler et al.,[25] *Neurology* 2014	232	1997–2013	Clinic-based	Initial clinical diagnosis by experts	Pathologic diagnosis	88.1	67.9	76.7
Adler et al.,[25] *Neurology* 2014	263	1997–2013	Clinic-based	Refined clinical diagnosis by experts	Pathologic diagnosis	88.7	82.2	84.8
Marshall et al.,[26] *Mov Disord* 2009	99	1999–2005	Clinic-based, on selected patients with uncertain diagnosis	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	93	46.4	79.8
Bajaj et al.,[27] *J Neurol Neurosurg Psychiatry* 2010 (reviewer 1)	38	NA	Clinic-based, on selected patients with uncertain diagnosis	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	53.3	82.6	71
Bajaj et al.,[27] *J Neurol Neurosurg Psychiatry* 2010 (reviewer 2)	38	NA	Clinic-based, on selected patients with uncertain diagnosis	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	80	73.9	76.3
Schrag et al.,[29] *J Neurol Neurosurg Psychiatry* 2002	200	1997	Community-based	Clinical diagnosis mainly by nonexperts	UKPDSBRC clinical criteria	88.1	73	82.5
Bower et al.,[30] *Mov Disord* 2002	39	1976- 1990	Community-based	Refined clinical diagnosis by experts	Pathologic diagnosis	43.7	73.9	61.5
Caslake et al.,[2] *J Neurol Neurosurg Psychiatry* 2008	66	2002–2004	Community-based	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	89.2	55.2	74.2
Caslake et al.,[2] *J Neurol Neurosurg Psychiatry* 2008	66	2002–2004	Community-based	Initial clinical diagnosis by experts	UKPDSBRC clinical criteria	84.8	45.4	65.2
Caslake et al.,[2] *J Neurol Neurosurg Psychiatry* 2008	66	2002–2004	Community-based	Refined clinical diagnosis by experts	UKPDSBRC clinical criteria	97	84.9	90.9

Continued

Table 1 Continued

Studies	Sample, n	Time of recruitment	Setting and population	In vivo diagnosis	Gold standard	Sensitivity, %	Specificity, %	Accuracy, %
Alves et al.,[32] *J Neurol Neurosurg Psychiatry* 2009	362	2004–2006	Community-based	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	100	70.9	91.7
Linder et al.,[33] *Mov Disord* 2010	139	2004–2007	Community-based	UKPDSBRC clinical criteria	UKPDSBRC clinical criteria	100	77.8	95.7
Winter et al.,[34] *Mov Disord* 2010	341	2006–2008	Community-based	UKPDSBRC clinical criteria	UKPDSBRC clinical criteria	100	24.2	92.7
Simpson and Clarke,[31] *Parkinsonism Relat Disord* 2013	289	1999–2009	Community-based	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	97.3	78.3	92.7
Caslake et al.,[35] *Parkinsonism Relat Disord* 2013	377	2002–2006	Community-based	Initial clinical diagnosis by experts	Refined clinical diagnosis by experts	81.5	80.7	81.2

Abbreviation: UKPDSBRC = United Kingdom PD Society Brain Research Center.

was not available for all studies and, when available, data were too heterogeneous to allow a meta-analysis. Accordingly, we performed a qualitative review on the misdiagnosis issue (table e-2). Among clinic-based studies,[3,11–25] MSA (range 0%–9.8%), PSP (range 0%–6.5%), dementia with Lewy bodies (DLB) (range 0%–8%), AD (range 0%–7.2%), and VaE (range 0%–7.7%) were the most frequent final diagnoses associated with false-positives. False-negatives mostly included diagnoses of undetermined parkinsonism (range 0%–17.3%), MSA (range 0%–5.7%), VaE (range 0%–5.8%) and PSP (range 0%–3.5%).

Among community-based studies,[2,28–36] the most frequent false-positives had a final diagnosis of ET (range 0%–26.5%), PSP (range 0%–25%), VaE (range 0%–19.6%), MSA (range 0%–12.5%), DLB (range 0%–10.8%), and drug-induced parkinsonism (range 0%–6.3%). False-negatives mostly included diagnoses of parkinsonism in dementia (range 0%–21.7%), nonparkinsonian tremor (range 0%–15.9%), drug-induced parkinsonism (range 0%–13%), MSA (range 0%–4.3%), and DLB (range 0%–3.5%).

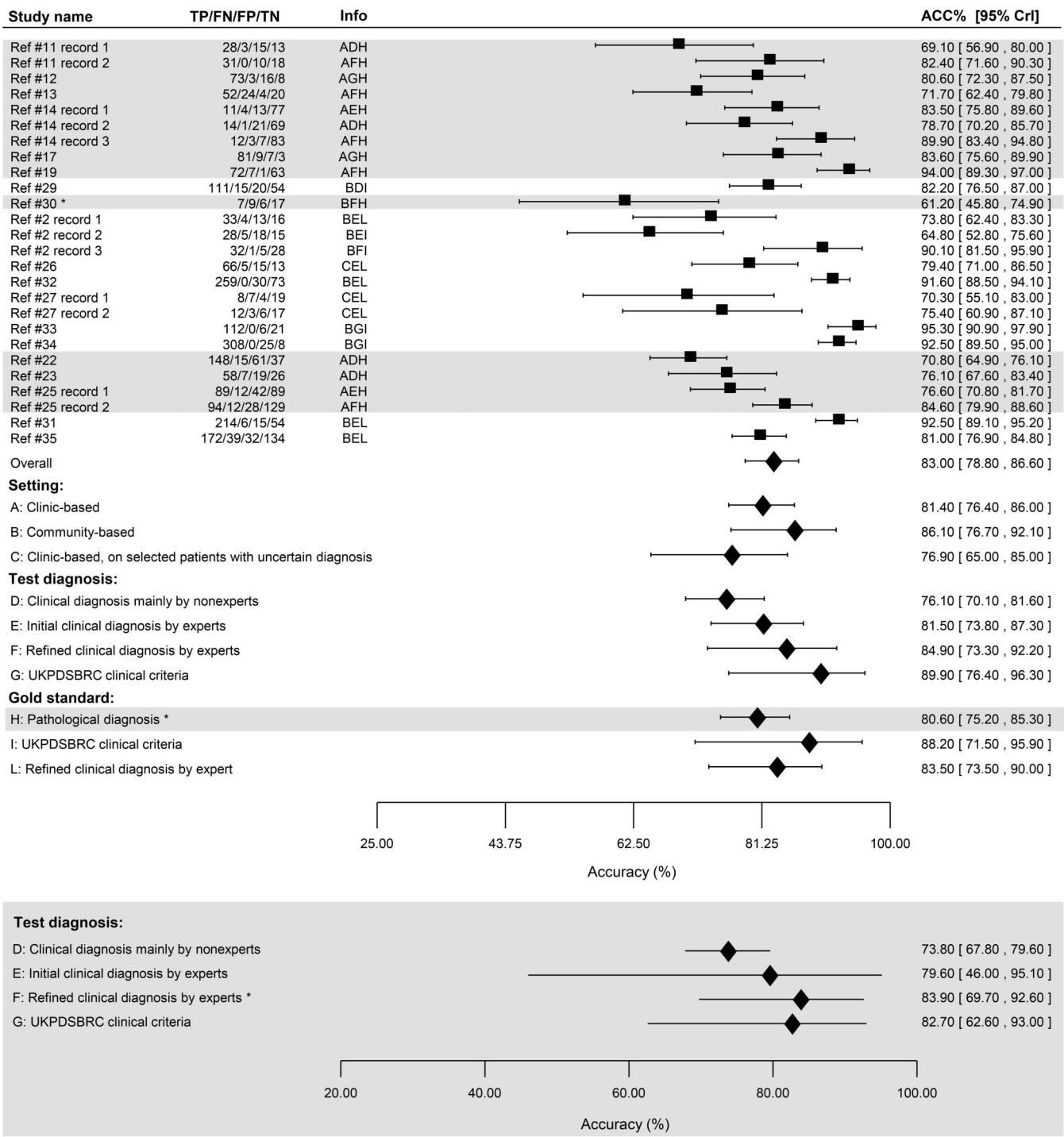
In the 2 clinic-based studies focused on tremulous patients with uncertain diagnosis,[26,27] the false-positives and false-negatives consisted of nonparkinsonian tremors (ET or dystonic tremor).

DISCUSSION Our meta-analysis showed that the accuracy of the clinical diagnosis of PD in the last 25 years has remained suboptimal, without substantial improvement in recent studies. Only 8 out of 10 patients with parkinsonism were diagnosed correctly. The diagnostic accuracy was mildly affected by the expertise of the doctors, being a little higher when the test diagnosis was provided by experts in movement disorders. The pooled data from the studies using the pathologic examination as gold standard provided an accuracy of the diagnosis by experts using all available clinical data after an adequate follow-up of 83.9%, with a sensitivity of 81.3% and a specificity of 83.5%. Compared with that, the diagnosis by nonexperts was slightly more sensitive (89.7%) but much less specific (49.2%), with a pooled accuracy of 73.8%.

Refined diagnosis by experts was also slightly better than the diagnosis performed using as test diagnosis the UKPDSBRC clinical criteria, which had a pooled diagnostic accuracy of 82.7%. UKPDSBRC clinical criteria were more sensitive (90.8% vs 81.3%), but much less specific (34%) compared to the expert clinical diagnosis (83.5%). This suggests that movement disorders experts may use a method of pattern recognition for diagnosis that goes beyond the diagnostic criteria. However, it should be noted that only 2 studies using pathology as gold standard evaluated the accuracy of UKPDSBRC clinical criteria,[12,17] and only retrospectively, leaving the possibility that applying them prospectively the diagnostic accuracy could be different. On the other hand, these criteria do not claim to detect classical idiopathic PD in the very early stages, as they are based on a 3-step procedure (step 1: definition of parkinsonism; step 2: exclusion criteria, which are retrospective and prospective; step 3: supportive prospective criteria). Accordingly, using UKPDSBRC criteria at baseline assessment, only step 1 and part of step 2 could be applied, leading to a diagnosis at most compatible with PD but that will be changed in some cases after follow-up.[32–34,37]

PD diagnosis in the early phase of disease is challenging, even for movement disorders experts, as response to dopaminergic treatment is less defined and clinical hallmarks of alternative diagnoses such as atypical parkinsonism may not have emerged. Even the initial diagnosis by experts was changed in a number of patients at follow-up. A disease duration ranging from 3.6 to 13.8 years (mean 10.2 years) is reported as useful to increase the diagnostic accuracy,[11,13,17,19,25,30] without

Figure 2 **Forest plot: Pooled accuracy of studies**



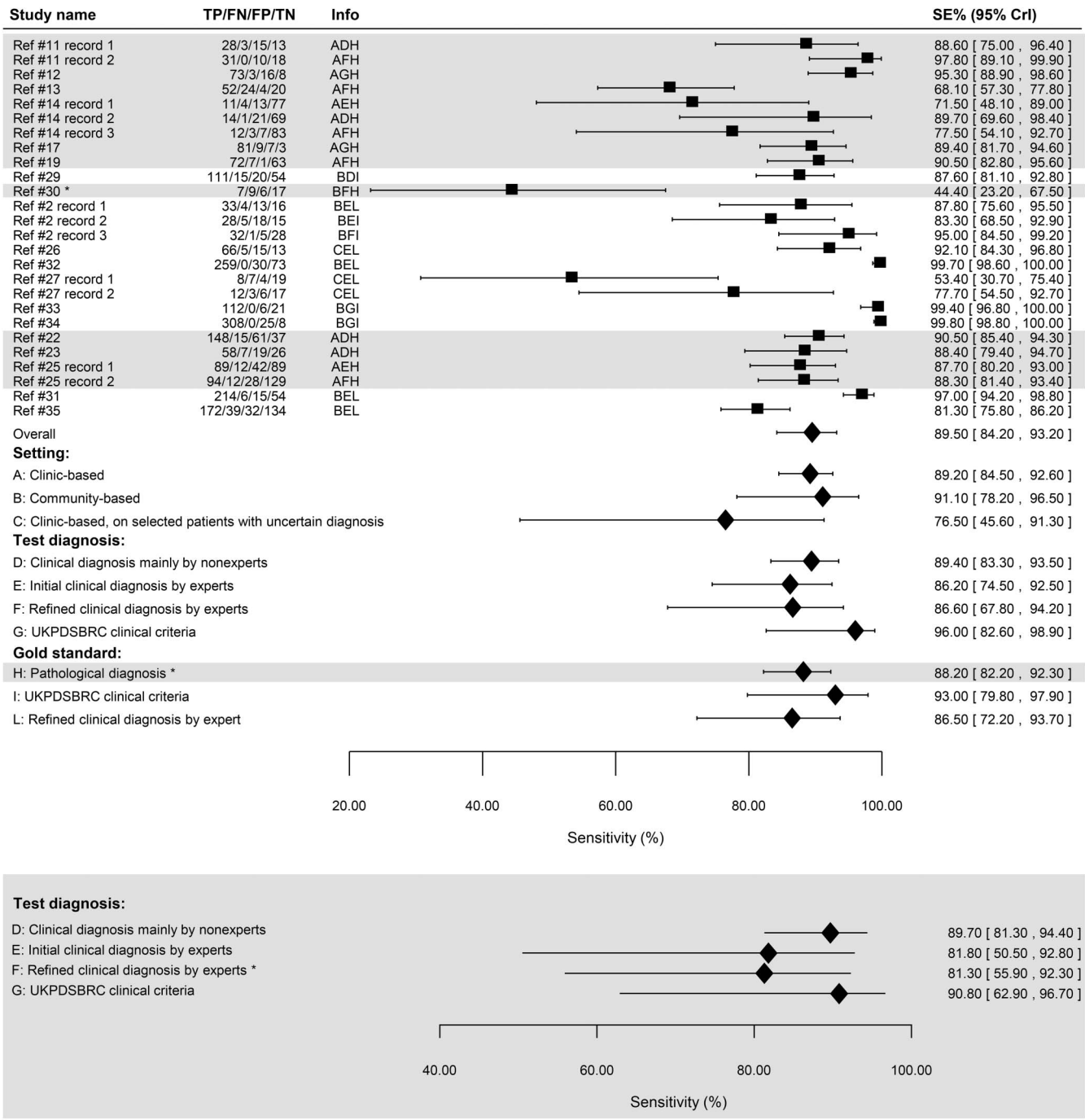
The data from the studies using pathologic examination as gold standard are highlighted in gray. *Excluding the single community-based study, Bower et al.,[30] pooled accuracy for the clinic-based studies using pathologic examination as gold standard was 81.4% (95% credible interval [CrI] 76.5%–85.7%), with 86.4% (95% CrI 76.5%–85.7%) for refined clinical diagnosis by experts. UKPDSBRC = United Kingdom PD Society Brain Research Center.

a clear trend of improvement of the accuracy over time. This can represent a major hurdle for clinical trials especially when they are designed to recruit patients with early PD. Indeed, the pooled accuracy of the initial diagnosis by experts for studies with pathologic examination was 79.6%, although reflecting the data from only 2 studies.[14,25]

When the initial PD diagnosis was performed in the earliest stages, the diagnostic error was very high, with a PPV that reached 53% in patients with <5 years of disease duration and even 26% in patients with <3 years of disease duration in one study.[25]

Other studies evaluated the accuracy of the initial diagnosis of PD by experts, reporting higher diagnostic

Figure 3 Forest plot: Pooled sensitivity of studies (except Hughes et al.[3]: a univariate model also including this study resulted in a sensitivity of 92.3% [95% CrI 86.6%–95.9%])

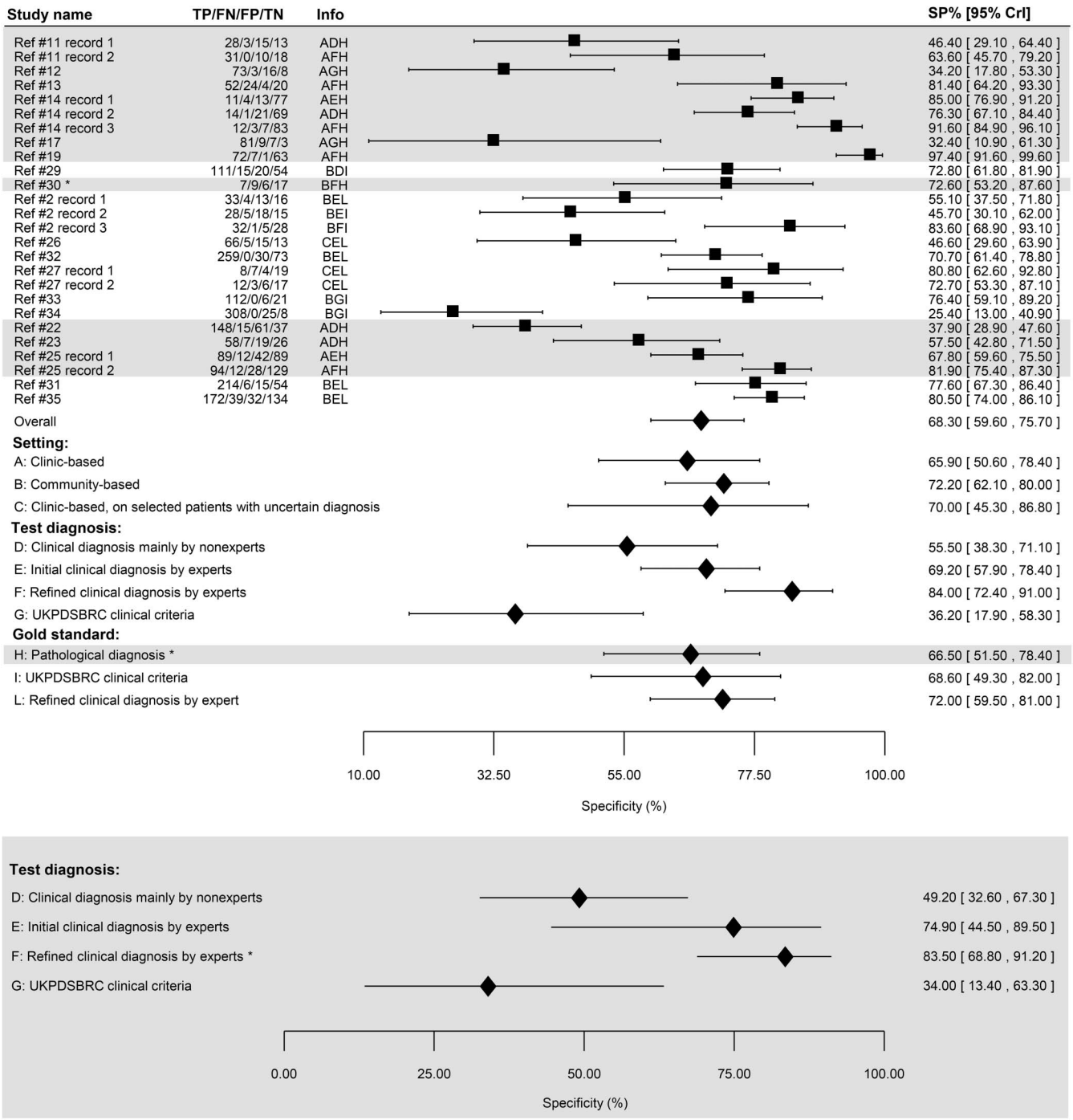


The data from the studies using pathologic examination as gold standard are highlighted in gray. *Excluding the single community-based study, Bower et al.,[30] pooled sensitivity for the clinic-based studies using pathologic examination as gold standard was 89.3% (95% credible interval [CrI] 84.5%–82.6%), with 86.6% (95% CrI 69.2%–94.1%) for refined clinical diagnosis by experts. UKPDSBRC = United Kingdom PD Society Brain Research Center.

parameters, but calculated using gold standard different from pathologic examination, specifically a refined diagnosis by experts after follow-up, which in turn resulted in suboptimal accuracy. The same problem also affected those studies focused on the accuracy of PD diagnosis in patients with uncertain diagnosis.[26,27]

We evaluated the effect of the different setting on the accuracy of PD diagnosis. Considering all studies, accuracy in the community-based studies (86.1%) was slightly higher than that in clinic-based studies (81.4%). However, it should be again considered that all but one[30] of the community-based studies did not use as gold standard the pathologic examination (used by 12 out 15 of the clinic-based studies). The use of clinical gold standard may have determined a possible inaccurate estimation of the diagnostic accuracy.

Figure 4 Forest plot: Pooled specificity of studies



The data from the studies using pathologic examination as gold standard are highlighted in gray. *Excluding the single community-based study, Bower et al.,[30] pooled specificity for the clinic-based studies using pathologic examination as gold standard was 65.8% (95% credible interval [CrI] 50.8%–78.2%), with 84.7% (95% CrI 68.7%–92.6%) for refined clinical diagnosis by experts.

Indeed, the only community-based study with a pathologic examination as gold standard reported a diagnostic accuracy of 61.5%, although on a very small sample of 39 patients.[30]

On the other hand, the pathologic examination might suffer from selection bias, since the presence of atypical signs and symptoms, male sex, younger age at death, and location at death (home or hospital) are associated with a higher autopsy rate.[38] In these series with autopsy, atypical parkinsonism has a higher prevalence. Furthermore, some forms of monogenic parkinsonism may not fulfil neuropathologic diagnostic criteria of PD.[39]

The most frequent misdiagnoses in clinic-based studies involve atypical parkinsonism, particularly MSA and PSP, while nonparkinsonian tremors, specifically ET, represent the most frequent misdiagnosis

in community-based studies (see tables e-1 and e-2). Nonparkinsonian tremors also become a possible misdiagnosis in the early stages of disease in clinic-based studies. This is the case of the so-called scans without evidence of dopaminergic deficit patients reported by a number of clinical trials,[24] who are mostly patients with ET or dystonic tremor. In our review, we selected 2 studies that focused on the diagnostic accuracy of this specific group of patients with uncertain diagnosis, observing a pooled value of diagnostic accuracy of 76.9%. However, neither study had pathologic examination as gold standard.[26,27]

A relevant percentage of misdiagnosis was observed for patients with different forms of dementia, i.e., AD, DLB, and VaD. Among these, the differential diagnosis between DLB and PD represents a known problem in the scientific community. DLB has clinical and pathologic characteristics that overlap with PD and PD dementia (PDD). The clinical criteria for DLB diagnosis were last revised in 2005 following the 3rd DLB International Workshop,[40] and an arbitrary 1-year rule continues to be recommended for the distinction between DLB and PDD, although applying this rule in clinical practice may prove challenging.[40–42] Another hurdle in the differential diagnosis of parkinsonism with or without dementia is the increasing evidence of mixed pathology in the brain of these patients. This is also true for patients with PD with pathologic evidence of mixed pathology at autopsy, such as vascular lesions, striatal plaques, AD, AD-type lesions, and diffuse Lewy body disease.[3] The issue of mixed pathology is currently gaining more interest considering the increase in the average age of the world population. Indeed, it has been increasingly recognized that co-occurrence of neurodegenerative proteinopathies (involving amyloid-β, tau, α-synuclein, and TDP-43) and other pathologies including cerebrovascular disorders is a frequent event in the aging brain.[4,43] Accordingly, the diagnosis of PD with a late onset of disease should be more complex compared to the early-onset patients. We did not find articles focused on diagnostic accuracy in late-onset PD and we did not have enough data available on the age at onset of the patients evaluated in the studies selected in this review to perform a regression between diagnostic parameters and the age at onset. Although this concept is plausible and a trend of a lower accuracy in the oldest patients appears evident by a qualitative evaluation of some studies in this review,[2,12,15,21,25] future studies should specifically evaluate the accuracy of the clinical diagnosis in late-onset PD.

A further critical point in the diagnosis of PD is the recent approach to prodromal PD, in the attempt to identify patients at risk of PD in the premotor phases of the disease. A number of nonmotor symptoms or imaging findings have been evaluated as early signs or risk factors of PD, including hyposmia, constipation, REM sleep behavior disorder, and substantia nigra hyperechogenicity.[44–48] The problems related to the sensitivity and specificity of a possible diagnosis of a prodromal PD are obvious as it actually represents in most cases a risk condition, the extent of which must be better evaluated. On the other hand, some of the same signs or symptoms could help to anticipate the diagnosis of motor PD.[37]

The heterogeneity of the studies included in our meta-analysis could be considered as a limitation. This problem was balanced by using random effects within the Bayesian framework. However, subgroup meta-analyses performed for those categories of studies including only a small number of records cannot provide definite conclusions. Study heterogeneity also limited a more detailed assessment of the relationship between final diagnoses and misdiagnosis rate.

Our review and meta-analysis demonstrated that the overall quality of clinical diagnosis of PD is inadequate, even in tertiary centers with movement disorder experts. Only 8 out of 10 patients with parkinsonism have a valid diagnosis. Low diagnostic accuracy is particularly relevant in the early stages of disease and presumably in older patients. These results emphasize the need for easily accessible biomarkers of disease to support clinical diagnosis in vivo and for new diagnostic criteria, possibly by a consensus, more appropriate for the early stages.

The rate of misdiagnosis should be taken into account when calculating sample sizes for clinical trials in patients with a diagnosis of PD, particularly close to the clinical onset of disease.

AUTHOR CONTRIBUTIONS

Dr. Giovanni Rizzo: study concept and design, acquisition of data, analysis and interpretation of data, drafting the manuscript. Dr. Massimiliano Copetti: analysis and interpretation of data, drafting the manuscript. Dr. Simona Arcuti: analysis and interpretation of data, drafting the manuscript. Dr. Davide Martino: acquisition of data, revising the manuscript. Dr. Andrea Fontana: analysis and interpretation of data. Dr. Giancarlo Logroscino: study concept and design, analysis and interpretation of data, revising the manuscript for content.

ACKNOWLEDGMENT

The authors thank the authors of articles selected for this review who helped recover missing data.

STUDY FUNDING

No targeted funding reported.

DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

Received March 5, 2015. Accepted in final form October 16, 2015.

REFERENCES

1. Braak H, Del Tredici K, Rüb U, et al. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 2003;24:197–211.

2. Caslake R, Moore JN, Gordon JC, et al. Changes in diagnosis with follow-up in an incident cohort of patients with parkinsonism. J Neurol Neurosurg Psychiatry 2008;79:1202–1207.
3. Hughes AJ, Daniel SE, Blankson S, Lees AJ. A clinicopathologic study of 100 cases of Parkinson's disease. Arch Neurol 1993;50:140–148.
4. Jellinger KA, Attems J. Challenges of multimorbidity of the aging brain: a critical update. J Neural Transm 2015;122:505–521.
5. National Collaborating Centre for Chronic Conditions (UK). Parkinson's Disease: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care. London: Royal College of Physicians; 2006.
6. Berardelli A, Wenning GK, Antonini A, et al. EFNS/MDS-ES/ENS recommendations for the diagnosis of Parkinson's disease. Eur J Neurol 2013;20:16–34.
7. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
8. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988;51:745–752.
9. Verde PE. Meta-analysis of diagnostic test data: a bivariate Bayesian modeling approach. Stat Med 2010;29:3088–3102.
10. Spiegelhalter DJ, Thomas A, Best N. WinBUGS, Version 1.4, Upgraded to 1.4.1, User Manual. Cambridge: MRC Biostatistics Unit; 2004.
11. Rajput AH, Rozdilsky B, Rajput A. Accuracy of clinical diagnosis in parkinsonism: a prospective study. Can J Neurol Sci 1991;18:275–278.
12. 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.
13. Hughes AJ, Ben-Shlomo Y, Daniel SE, Lees AJ. What features improve the accuracy of clinical diagnosis in Parkinson's disease: a clinicopathologic study. Neurology 1992;42:1142–1146.
14. Litvan I, MacIntyre A, Goetz CG, et al. Accuracy of the clinical diagnoses of Lewy body disease, Parkinson disease, and dementia with Lewy bodies: a clinicopathologic study. Arch Neurol 1998;55:969–978.
15. Jankovic J, Rajput AH, McDermott MP, Perl DP. The evolution of diagnosis in early Parkinson disease: Parkinson Study Group. Arch Neurol 2000;57:369–372.
16. Jellinger KA. The pathology of Parkinson's disease. Adv Neurol 2001;86:55–72.
17. Hughes AJ, Daniel SE, Lees AJ. Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. Neurology 2001;57:1497–1499.
18. Lees AJ, Katzenschlager R, Head J, Ben-Shlomo Y. Ten-year follow-up of three different initial treatments in de-novo PD: a randomized trial. Neurology 2001;57:1687–1694.
19. Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. Brain 2002;125:861–870.
20. Stoffers D, Booij J, Bosscher L, Winogrodzka A, Wolters EC, Berendse HW. Early-stage [123I]beta-CIT SPECT and long-term clinical follow-up in patients with an initial diagnosis of Parkinson's disease. Eur J Nucl Med Mol Imaging 2005;32:689–695.
21. Jellinger KA. Morphological substrates of parkinsonism with and without dementia: a retrospective clinicopathological study. J Neural Transm Suppl 2007:91–104.
22. Horvath J, Burkhard PR, Bouras C, Kövari E. Etiologies of parkinsonism in a century-long autopsy-based cohort. Brain Pathol 2013;23:28–33.
23. Joutsa J, Gardberg M, Röyttä M, Kaasinen V. Diagnostic accuracy of parkinsonism syndromes by general neurologists. Parkinsonism Relat Disord 2014;20:840–844.
24. Marek K, Seibyl J, Eberly S, et al. Longitudinal follow-up of SWEDD subjects in the PRECEPT Study. Neurology 2014;82:1791–1797.
25. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease: clinicopathologic study. Neurology 2014;83:406–412.
26. Marshall VL, Reininger CB, Marquardt M, et al. Parkinson's disease is overdiagnosed clinically at baseline in diagnostically uncertain cases: a 3-year European multicenter study with repeat [123I]FP-CIT SPECT. Mov Disord 2009;24:500–508.
27. Bajaj NP, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study. J Neurol Neurosurg Psychiatry 2010;81:1223–1228.
28. Meara J, Bhowmick BK, Hobson P. Accuracy of diagnosis in patients with presumed Parkinson's disease. Age Ageing 1999;28:99–102.
29. Schrag A, Ben-Shlomo Y, Quinn N. How valid is the clinical diagnosis of Parkinson's disease in the community? J Neurol Neurosurg Psychiatry 2002;73:529–534.
30. Bower JH, Dickson DW, Taylor L, et al. Clinical correlates of the pathology underlying parkinsonism: a population perspective. Mov Disord 2002;17:910–916.
31. Simpson BS, Clarke CE. Retrospective evaluation of the diagnostic accuracy of Parkinsonism in a UK community based movement disorders clinic. Parkinsonism Relat Disord 2013;19:461–462.
32. Alves G, Müller B, Herlofson K, et al. Incidence of Parkinson's disease in Norway: the Norwegian ParkWest study. J Neurol Neurosurg Psychiatry 2009;80:851–857.
33. Linder J, Stenlund H, Forsgren L. Incidence of Parkinson's disease and parkinsonism in northern Sweden: a population-based study. Mov Disord 2010;25:341–348.
34. Winter Y, Bezdolnyy Y, Katunina E, et al. Incidence of Parkinson's disease and atypical parkinsonism: Russian population-based study. Mov Disord 2010;25:349–356.
35. Caslake R, Taylor K, Scott N, et al. Age-, gender-, and socioeconomic status-specific incidence of Parkinson's disease and parkinsonism in northeast Scotland: the PINE study. Parkinsonism Relat Disord 2013;19:515–521.
36. Duncan GW, Khoo TK, Coleman SY, et al. The incidence of Parkinson's disease in the North-East of England. Age Ageing 2014;43:257–263.
37. Gaenslen A, Unmuth B, Godau J, et al. The specificity and sensitivity of transcranial ultrasound in the differential diagnosis of Parkinson's disease: a prospective blinded study. Lancet Neurol 2008;7:417–424.
38. Maraganore DM, Anderson DW, Bower JH, et al. Autopsy patterns for Parkinson's disease and related disorders in Olmsted County, Minnesota. Neurology 1999;53:1342–1344.

39. Uitti RJ, Calne DB, Dickson DW, Wszolek ZK. Is the neuropathological "gold standard" diagnosis dead? Implications of clinicopathological findings in an autosomal dominant neurodegenerative disorder. Parkinsonism Relat Disord 2004;10:461–463.
40. McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology 2005;65:1863–1872.
41. McKeith I. Dementia with Lewy bodies and Parkinson's disease with dementia: where two worlds collide. Pract Neurol 2007;7:374–382.
42. Huang Y, Halliday G. Can we clinically diagnose dementia with Lewy bodies yet? Transl Neurodegener 2013;2:4.
43. Kovacs GG, Milenkovic I, Wöhrer A, et al. Non-Alzheimer neurodegenerative pathologies and their combinations are more frequent than commonly believed in the elderly brain: a community-based autopsy series. Acta Neuropathol 2013;126:365–384.
44. Abbott RD, Petrovitch H, White LR, et al. Frequency of bowel movements and the future risk of Parkinson's disease. Neurology 2001;57:456–462.
45. Iranzo A, Lomeña F, Stockner H, et al. Decreased striatal dopamine transporter uptake and substantia nigra hyperechogenicity as risk markers of synucleinopathy in patients with idiopathic rapid-eye-movement sleep behaviour disorder: a prospective study. Lancet Neurol 2010;9:1070–1077.
46. Mahlknecht P, Iranzo A, Högl B, et al. Olfactory dysfunction predicts early transition to a Lewy body disease in idiopathic RBD. Neurology 2015;84:654–658.
47. Berg D, Behnke S, Seppi K, et al. Enlarged hyperechogenic substantia nigra as a risk marker for Parkinson's disease. Mov Disord 2013;28:216–219.
48. Jennings D, Siderowf A, Stern M, et al. Imaging prodromal parkinson disease: the Parkinson associated risk syndrome study. Neurology 2014;83:1739–1746.

Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis

Giovanni Rizzo, Massimiliano Copetti, Simona Arcuti, et al.

Neurology published online January 13, 2016
DOI 10.1212/WNL.0000000000002350