ORIGINAL ARTICLE

Effects of Antiparkinson Medication on Cognition in Parkinson's Disease: A Systematic Review

Marc-André Roy, Maxime Doiron, Jessica Talon-Croteau, Nicolas Dupré, Martine Simard

ABSTRACT: ***Objective:*** This study aimed to systematically review the effects of currently prescribed antiparkinson medication on cognition in patients with mild-to-moderate Parkinson's disease (PD) who were either cognitively intact or mildly impaired. ***Methods:*** English- and French-language studies published between 1969 and 2017 were accessed via MedLine, PsychNET, EMBASE and EBSCO databases. Methodological quality (MQ) was evaluated with the quality assessment instrument of the *Cochrane Collaboration Depression, Anxiety and Neurosis Review* (scores from 0% to 44% indicate very low quality; scores from 45% to 64% indicate low quality; scores from 65% to 84% indicate medium quality; and scores from 85% to 100% indicate high quality). Hedges' *g* and Student's *t*-test were performed on all cognitive outcome measures reported. ***Results:*** In total, 14 studies assessed the cognitive effects of levodopa (L-D), pramipexole (PRX), selegiline (SEL) and rasagiline (RAS) in mild-to-moderate non-demented PD patients. The MQ was overall low, with an average score of 49.1%. Results for L-D showed deleterious effects on a test of cognitive inhibition, as well as benefits on tests of attention/processing speed/working memory, executive functions and episodic memory. Pramipexole was associated with a worsening of episodic memory and impulse control. Results on SEL indicated a deterioration of global cognition over time and of concept formation. Rasagiline had some benefits on working memory and verbal fluency. ***Conclusion:*** Antiparkinson medications can have deleterious (L-D; PRX; SEL) and beneficial (L-D; RAS) effects on cognition. However, randomized double-blind placebo-controlled trials with larger sample sizes are required to better elucidate this issue.

RÉSUMÉ: ***Objectif:*** Cette étude vise à recenser systématiquement les effets cognitifs de certains antiparkinsoniens chez des patients aux stades léger à modéré de maladie de Parkinson (MP) avec ou sans atteintes cognitives légères. ***Méthode:*** Les publications anglophones et francophones de 1969 à 2017 ont été recensées via MedLine, PsychNET, EMBASE et EBSCO. La qualité méthodologique (QM) a été évaluée avec un instrument de la *Cochrane Collaboration Depression, Anxiety and Neurosis Review* (0% à 44% signifie une très faible qualité; 45% à 64% une faible qualité; 65% à 84% une qualité moyenne; 85% à 100% une haute qualité). Le *g* de Hedges et le *t* de Student ont été calculés pour tous les résultats cognitifs. ***Résultats:*** Quatorze études évaluaient les effets cognitifs de levodopa, pramipexole, selegiline et rasagiline chez des patients aux stades léger à modéré de MP sans démence. Globalement, la QM était faible (moyenne de 49.1%). Levodopa est associé à une altération des capacités d'inhibition et des bénéfices aux tests d'attention/vitesse de traitement/mémoire de travail, de fonctions exécutives et de mémoire épisodique. Pramipexole est associé à une détérioration en mémoire épisodique et en contrôle des impulsions. Selegiline est associé à une détérioration de la cognition globale et de la formation de concepts. Rasagiline est associé à des bénéfices en mémoire de travail et en fluidité verbale. ***Conclusion:*** Les antiparkinsoniens peuvent avoir des effets cognitifs délétères (levodopa; pramipexole; selegiline) et bénéfiques (rasagiline; levodopa). Davantage d'essais randomisés en double-aveugle et contrôle placebo sont nécessaires pour détailler ces effets.

Keywords: Parkinson's disease, Cognition, Levodopa, Pramipexole, Rasagiline

doi:10.1017/cjn.2018.21 Can J Neurol Sci. 2018; 45: 375-404

Introduction

Diagnostic criteria for mild cognitive impairment (MCI) in Parkinson's disease (PD) were developed[1] because PD-MCI is common; it is present in 10% to 64%, that is approximately 40%, of non-demented PD patients (although exact estimates differ from study to study).[2-6] This condition is characterized by the presence of cognitive impairment that does not significantly alter daily living. In PD-MCI, attentional/executive, visuo-spatial and memory functions are usually the most impaired, although other domains are also often altered, especially as the disease progresses.[1,6-9] Identifying a way to slow down or prevent the cognitive decline in PD-MCI is crucial, as longitudinal studies suggest that around 11% of PD-MCI patients annually convert to

From the School of Psychology, Laval University, Quebec City, Quebec, Canada; Cervo Brain Research Center of Institut Universitaire en Santé Mentale de Québec, Quebec City, Quebec, Canada (M-AR, MD, JT-C, MS); Faculty of Medicine, Laval University, Quebec City, Quebec, Canada (MD, ND); Centre Hospitalier Universitaire de Québec, Université Laval, Quebec City, Quebec, Canada (ND).

Received June 6, 2017. Final Revisions Submitted February 11, 2018. Date of Acceptance February 15, 2018.

Correspondence to: Marc-André Roy, School of Psychology, Laval University, Pav. F.-A. Savard, 2325 rue des Bibliothèques, Quebec City, QC, Canada, G1V 0A6. Email: marc-andre.roy.11@ulaval.ca

Parkinson's disease with dementia (PDD)[6] and that most new PD-MCI cases could progress to PDD within as little as 5 years.[10]

The progressive degeneration of dopaminergic neurons in the nigrostriatal and mesocorticolimbic pathways and the presence of Lewy bodies in midbrain neurons are the neuropathological substrates of early clinical symptoms in PD and PD-MCI. Changes in the cholinergic, noradrenergic and serotoninergic systems also appear during the evolution of PD,[11-15] and contribute to motor and cognitive symptoms. As for dopamine (DA), nigrostriatal damage results in DA depletion in the caudate nucleus, which progresses from its dorso-lateral to ventro-medial segments. These areas are involved in frontostriatal loops. This phenomenon partly reflects the usual evolution of cognitive impairment in PD: an early presence of a dysexecutive syndrome (dorso-lateral loop) and the later development of difficulties with reward-based control of behavior (orbitofrontal loop).[16,17] In parallel, neurodegeneration also occurs in the DA mesocorticolimbic system, which innervates many structures involved in cognitive functioning, such as the prefrontal cortex and the hippocampal formation.[18,19]

As the earliest and most important pathological mechanism of PD affects DA transmission, prescribed medications are principally dopaminergic agents. Although DA depletion does not explain all cognitive impairments, it nevertheless alters cognition.[20,21] Therefore, dopaminergic agents, by compensating DA depletion, are likely to affect cognition in PD. The effects of antiparkinson medications on cognition have been investigated by several scientists since the advent of levodopa (L-D). However, the results of this work have never been reviewed in a systematic manner. In 2013, the first review tackling this question[22] was more a narrative than a systematic review. Although this work provided some insight into the subject, few details were reported about the selection process and the characteristics of the studies. In addition, neither methodological quality of the articles nor effect sizes of the results were reported.

Therefore, the objective of the present article was to study the impact of clinically relevant dopaminergic antiparkinson medications on cognitive functioning in mild-to-moderate non-demented PD patients with compromised or intact cognition, by systematically reviewing the literature and providing quantitative data to assess the significance of the reported effects. Among all the antiparkinson drugs, this review covers 71% of treatments that are currently approved for clinical use. More precisely, antiparkinsonians recommended by the Canadian Neurological Sciences Federation (CNSF) Guidelines,[23] namely L-D (with or without concomitant dopa decarboxylase inhibitor [DDCI] and/or catechol o-methyltransferase [COMT] inhibitor), pramipexole (PRX), ropinirole, selegiline (SEL) and rasagiline (RAS), were chosen to be reviewed.

Method

Search Strategy and Selection Criteria

A systematic search of the English- and French-language literature published between January 1969 and November 2017 was undertaken using the electronic databases MEDLINE, PsychNET, EMBASE and EBSCO (see Table 1 for keywords and search strategy). A manual search in the references of the selected articles was also made.

Eligible studies had to be randomized trials or have a non-randomized design (e.g., pre-post or off-on studies) with a control group or within-group comparisons, with or without placebo (Pb). Designs that allowed the isolation of the effect of one selected medication on at least one cognitive measure were accepted. Participants had to be diagnosed with PD according to the Queen Square Brain Bank (QSBB)[24] or the National Institute of Neurological Disorders and Stroke[25] criteria. Studies published before the QSBB criteria had to specify that a neurologist made the diagnosis and had to minimally use the Hoehn and Yahr (H&Y) stages[26,27] to document the evolution of the disease. Parkinson's disease severity had to be mild to moderate, defined by H&Y stages ≤ 3 and a PD duration ≤ 10 years at baseline. Medication dosage had to be indicated and results had to be reported on at least one validated cognitive measure before and after intervention. Exclusion criteria were as follows: cohort and case-control studies; case reports and animal model studies; trials including participants presenting with major psychiatric or neurologic disorder besides PD; and PDD per the DSM[28,29] criteria or an Mini-Mental State Examination (MMSE)[30] score <26, as recommended by the Movement Disorder Society (MDS).[31]

Article Selection and Data Extraction

Figure 1 illustrates the three phases of article selection. Phases two and three were conducted by two independent reviewers; all discrepancies were solved by consensus. The methodological quality analysis was made using the *Cochrane Collaboration Depression, Anxiety and Neurosis Review* (CCDAN) quality assessment scale,[32] which is a 23-criteria scale, with each criterion worth up to 2 points, for a maximum score of 46. This score was converted to percentage to obtain the methodological quality (MQ) score, and it was interpreted as follows: 0%-44% = very low quality; 45%-64% = low quality; 65%-84% = medium, acceptable quality with some risk of bias; and 85%-100% = high quality with little or no risk of bias. As the authors of this tool did not provide any interpretation method for their scores, we devised these benchmarks based on common sense. It was done in order to facilitate the interpretation of data, to help comparisons between articles and to provide rough estimates of the quality of the articles for each medication. As these interpretation benchmarks are not validated, the results section contains the main limitations for the group of articles associated with each medication in an effort to clarify the reasons why they were classified as they were.

Statistical Analysis

It was impossible to conduct a meta-analysis of our data. Data on several cognitive outcome measures were needed to achieve our objective of assessing the effects on all cognitive domains, making it hard to pool the data in a relevant way. The second best option was thus chosen, namely a systematic review of the literature with unpooled statistical analyses. Effect size calculations were conducted whenever available data allowed the comparison of cognitive performance before and after treatment. Hedges' g[33] was used, as it provides a correction to Cohen's d, which overestimates the effect size when small samples are involved. For matched-sample comparisons, the formula took into consideration the covariance of the patients' performance. Whenever there were insufficient data to estimate the covariance, it was computed to be of medium size ($r = 0.3$).[34,35] Cohen's d benchmarks were used for interpretation, as Hedges'g and Cohen's d can be interpreted the same way. Values of $g = 0.2$, 0.5 and 0.8 were, respectively,

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 1: Search strategy and keywords used in the MEDLINE, EMBASE, PsychNET and EBSCO databases

Database	
MEDLINE via OVID	
Date	November 22, 2017
Strategy	(#1 AND #2 AND #3) NOT #4
#1	exp Parkinson Disease/OR Parkinson*.af.
#2	exp Mild Cognitive Impairment/OR exp Neuropsychology/OR exp cognition/OR exp executive function/OR exp memory/OR exp spatial learning/OR exp Attention/OR exp Neuropsychological Tests/OR exp Space Perception/OR (cognit* OR neuropsycholog* OR executive function* OR visuospatial OR attention* OR memory OR prefrontal function*).af.
#3	(dopamine agonist* OR pramipexole OR ropinirole OR MAO-B inhibitor* OR monoamine oxidase inhibitor* OR rasagiline OR selegiline OR antiparkinson medication* OR L-DOPA OR L-3,4-dihydroxyphenylalanine OR levodopa OR levodopa-carbidopa OR levodopa-benserazide OR COMT inhibitor* OR catechol o-methyltransferase inhibitor* OR entacapone).af. OR antiparkinson agents/OR benserazide/OR carbidopa/OR levodopa/OR selegiline/OR aromatic amino acid decarboxylase inhibitors/OR catechol o-methyltransferase inhibitors/OR Dopamine Agonists/OR Monoamine Oxidase Inhibitors/
#4	exp Genes/OR exp models, animal/OR exp primates/OR exp rodentia/OR exp mice/OR exp rats/OR gene*.af. OR (animal* or rat* or mouse or mice or macaque*).af.
Results	1175
EMBASE	
Date	November 22, 2017
Strategy	(#1 AND #2 AND #3) NOT #4
#1	('parkinson disease'/exp OR parkinson*))
#2	((('cognition'/exp OR 'attention'/exp OR 'executive function'/exp OR 'memory'/exp) OR ('neuropsychology'/exp OR 'mild cognitive impairment'/exp) OR (cognit* OR neuropsycholog* OR executive next/1 function* OR visuospatial OR attention* OR memory OR prefrontal next/1 function*))
#3	(('dopamine receptor stimulating agent'/de OR 'pramipexole'/exp OR 'ropinirole'/exp OR 'antiparkinson agent'/de OR 'benserazide plus levodopa'/exp OR 'carbidopa plus entacapone plus levodopa'/exp OR 'dopa decarboxylase inhibitor'/de OR 'carbidopa plus levodopa'/exp OR 'levodopa'/exp OR 'monoamine oxidase inhibitor'/de OR 'selegiline'/exp OR 'rasagiline'/exp OR 'catechol methyltransferase inhibitor'/de OR 'entacapone'/exp) OR (dopamine next/1 agonist* OR pramipexole OR ropinirole OR 'mao b' next/1 inhibitor* OR monoamine next/1 oxidase next/1 inhibitor* OR rasagiline OR selegiline OR antiparkinson medication* OR 'l dopa' OR 'l 3,4 dihydroxyphenylalanine' OR levodopa OR 'levodopa carbidopa' OR 'levodopa benserazide' OR comt next/1 inhibitor* OR (catechol next/1 'o methyltransferase' and inhibitor*) OR entacapone))
#4	((('animal'/exp OR 'rat'/exp OR 'mouse'/exp OR 'macaca'/exp) OR (macaque* OR rat* OR mouse OR mice OR animal*)) OR ('gene'/exp OR gene*))
Results	222
PsychNET	
Date	November 22, 2017
Strategy	(#1 AND #2 AND #3) NOT #4
#1	(AnyField:(parkinson*) OR (IndexTermsFilt:("Parkinson's Disease")))
#2	(((AnyField:(Dopamine agonist*) OR AnyField:(Pramipexole) OR AnyField:(ropinirole) OR AnyField:(MAO-B inhibitor*) OR AnyField:(monoamine oxidase inhibitor*) OR AnyField:(rasagiline) OR AnyField:(selegiline) OR AnyField:(antiparkinson medication*) OR AnyField:(L-DOPA OR levodopa OR "levodopa-carbidopa" OR "L-3,4-dihydroxyphenylalanine" OR "levodopa-benserazide") OR AnyField:(COMT inhibitor) OR AnyField:("catechol o-methyltransferase inhibitor*") OR AnyField:(entacapone)) OR ((IndexTermsFilt:("Antitremor Drugs") OR IndexTermsFilt:("Carbidopa") OR IndexTermsFilt:("Decarboxylase Inhibitors") OR IndexTermsFilt:("Dopamine Agonists") OR IndexTermsFilt:("Levodopa") OR IndexTermsFilt:("Monoamine Oxidase Inhibitors"))))
#3	((AnyField:(cognit*) OR AnyField:(neuropsycholog*) OR AnyField:(executive function*) OR AnyField:(prefrontal function*) OR AnyField:(visuospatial) OR AnyField:(attention) OR AnyField:(memory)) OR ((AnyFieldFilt:("Attention") OR AnyFieldFilt:("Attentional Capture") OR AnyFieldFilt:("Autobiographical Memory") OR AnyFieldFilt:("Cognition") OR AnyFieldFilt:("Cognitive Control") OR AnyFieldFilt:("Divided Attention") OR AnyFieldFilt:("Early Memories") OR AnyFieldFilt:("Eidetic Imagery") OR AnyFieldFilt:("Episodic Memory") OR AnyFieldFilt:("Executive Function") OR AnyFieldFilt:("Explicit Memory") OR AnyFieldFilt:("False Memory") OR AnyFieldFilt:("Implicit Memory") OR AnyFieldFilt:("Long Term Memory") OR AnyFieldFilt:("Memory") OR AnyFieldFilt:("Memory Consolidation") OR AnyFieldFilt:("Memory Decay") OR AnyFieldFilt:("Memory Trace") OR AnyFieldFilt:("Monitoring") OR AnyFieldFilt:("Neuropsychological Assessment") OR AnyFieldFilt:("Neuropsychology") OR AnyFieldFilt:("Prospective Memory") OR AnyFieldFilt:("Reminiscence") OR AnyFieldFilt:("Repressed Memory") OR AnyFieldFilt:("Retrospective Memory") OR AnyFieldFilt:("Selective Attention") OR AnyFieldFilt:("Set Shifting") OR AnyFieldFilt:("Short Term Memory") OR AnyFieldFilt:("Spatial Memory") OR AnyFieldFilt:("Spontaneous Recovery (Learning)") OR AnyFieldFilt:("Sustained Attention") OR AnyFieldFilt:("Task Switching") OR AnyFieldFilt:("Verbal Memory") OR AnyFieldFilt:("Vigilance") OR AnyFieldFilt:("Visual Attention") OR AnyFieldFilt:("Visual Memory") OR AnyFieldFilt:("Visuospatial Ability") OR AnyFieldFilt:("Visuospatial Memory")))))
#4	((AnyField:(gene*)) OR ((IndexTermsFilt:("Alleles") OR IndexTermsFilt:("CLOCK Gene") OR IndexTermsFilt:("Genes") OR IndexTermsFilt:("Immediate Early Genes") OR IndexTermsFilt:("Quantitative Trait Loci"))))OR ((AnyField:(animal*) OR AnyField:(rat*) OR AnyField:(mouse) OR AnyField:(mice) OR AnyField:(macaque*)) OR ((IndexTermsFilt:("Animal Models") OR IndexTermsFilt:("Baboons") OR IndexTermsFilt:("Bonobos") OR IndexTermsFilt:("Chimpanzees") OR IndexTermsFilt:("Gorillas") OR IndexTermsFilt:("Mice") OR IndexTermsFilt:("Monkeys") OR IndexTermsFilt:("Norway Rats") OR IndexTermsFilt:("Primates (Nonhuman)") OR IndexTermsFilt:("Rats") OR IndexTermsFilt:("Rodents"))))
Results	372

Table 1. *Continued*

EBSCO via AgeLine	
Date	November 22, 2017
Strategy	(#1 AND #2 AND #3) NOT #4
#1	(((DE + %26quot%3bParkinsons + Disease%26quot%3b) + OR + Parkinson*)
#2	((((dopamine + AND + agonist*)) + OR + (pramipexole) + OR + (ropinirole) + OR + ((MAO-B + AND + inhibitor*)) + OR + ((monoamine + AND + oxidase + AND + inhibitor*)) + OR + (rasagiline) + OR + (selegiline) + OR + (antiparkinson medication*) + OR + (L-DOPA) + OR + (levodopa) + OR + (L-3%2c4-dihydroxyphenylalanine) + OR + (levodopa-carbidopa)) + OR + ((levodopa-benserazide) + OR + ((COMT + AND + inhibitor*)) + OR + ((catechol + AND + o-methyltransferase + AND + inhibitor*)) + OR + (entacapone))))
#3	((DE + %26quot%3bCognitive + Impairment%26quot%3b + OR + DE + %26quot%3bMemory + Impairment%26quot%3b + OR + DE + %26quot%3bCognition%26quot%3b + OR + DE + %26quot%3bSocial + Cognition%26quot%3b + OR + DE + %26quot%3bCognitive + Impairment%26quot%3b + OR + DE + %26quot%3bDivided + Attention%26quot%3b + OR + DE + %26quot%3bIntelligence%26quot% 3b + OR + DE + %26quot%3bMemory%26quot%3b + OR + DE + %26quot%3bSelective + Attention%26quot%3b + OR + DE + %26quot%3bSpatial + Abilities%26quot%3b) + OR + ((cognit*) + OR + (neuropsycholog*) + OR + ((executive + AND + function*)) + OR + (visuospatial) + OR + (attention*) + OR + (memory) + OR + ((prefrontal + AND + function*))))
#4	(((animal*) + OR + (rat*) + OR + (mouse) + OR + (mice) + OR + (macaque*)) + OR + ((DE + %26quot%3bGenetics%26quot%3b) + OR + gene*))
Results	15
Total	
Before removal of duplicates	1784
After removal of duplicates	1553

interpreted as small, medium and large effect sizes. Student's *t*-tests were also performed. Results were considered statistically significant with $p < 0.05$.

Results

Figure 1 presents the flowchart. Fourteen articles meeting the selection criteria studied L-D with and without DDCI, PRX, SEL and RAS. No article was found regarding the effects of ropinirole or COMT inhibitors on cognition. No study used an extensive neuropsychological assessment to differentiate patients with intact versus compromised cognition (except for one[36]), nor did they use MCI or PD-MCI clinical criteria to categorize cognitively altered patients. The global scores of the MMSE/Montreal Cognitive Assessment (MoCA) were used to exclude PDD patients. Tables 2A-2D show the studies' characteristics. Table 3 describes the cognitive outcome measures administered in the studies.

Of the 14 studies that were finally included in this review, three (21%) did not report any statistically significant result on cognitive variables ("negative" studies): two were on PRX and one was on SEL. In total, 11 studies (79%) reported significant cognitive results ("positive" studies), together with some non-significant results.

Levodopa

Table 2A presents the characteristics of the seven studies of very low to medium MQ on L-D. A small sample size, the lack of blinding for the participants, as well as for the assessors, the absence of a power calculation and the lack of concealment of subject allocation were the most common limitations of the low- and very low-MQ studies on L-D. The main limitations of the medium MQ studies were the absence of power calculation, of concealment of allocation and of blinding of assessors. The L-D studies described the effects of 358-540 mg/day of L-D and of 200-281 mg/day of L-D and concomitant dopa decarboxylase inhibitor (L-D + DDCI). One study was a randomized controlled trial (RCT),[39] three were randomized cross-over trials[40,41,43] and three were quasi-experimental studies.[37,42,44] All of the quasi-experimental studies used a pretest-posttest design and within-group comparisons, with two of them having a healthy control group.[42,44] The trials lasted between <5 days and 24 months.

The number of participants enrolled in the patient groups varied from 10[39] to 387,[37] with 5/7 studies having sample sizes ≤20. Mean age and disease duration ranged, respectively, from 56.0 to 67.3 and from 1.2 to 3.9 years. A study selected only *de novo* patients.[39] The means of total scores on the MMSE/MoCA at baseline (from 28.4 to 28.9/30) suggest that the majority of PD patients had either mildly impaired or intact cognitive functioning. Table 4A presents the results from the L-D studies.

Every study presented some statistically significant results on their cognitive outcome measures. Of a total of 185 results, 32 (17.3%) indicated a statistically significant effect; 12.5% of these effects were deleterious and 87.5% were beneficial. Most significant effects were found on measures of executive functions (53.1%) and episodic memory (28.1%). Half of the significant effects were reported by the two studies with medium MQ, whereas the other half were reported by the five studies with low to very low MQ. Within each cognitive domain,

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

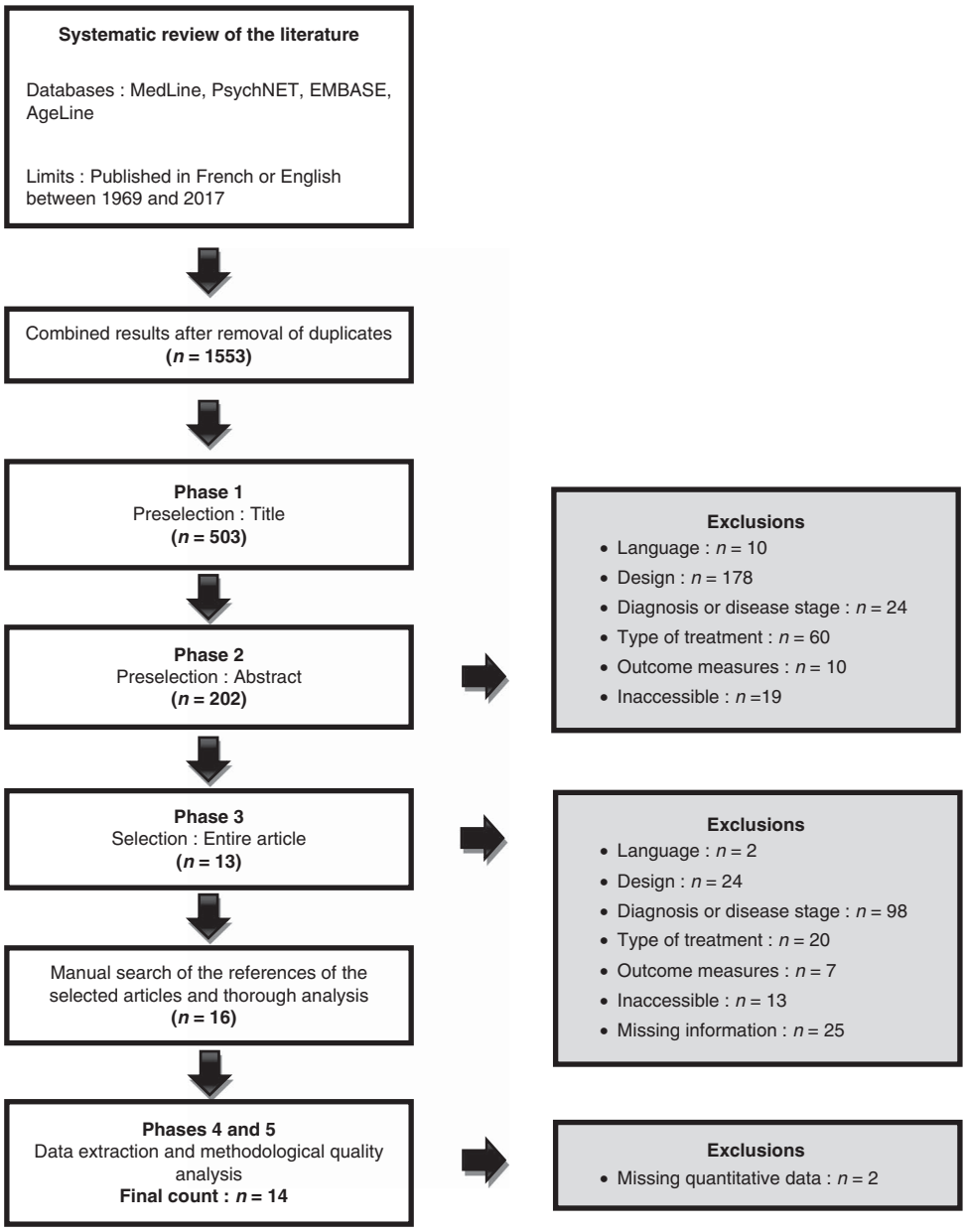

Figure 1: *Flowchart.*

results from studies with lower and higher MQ scores were generally convergent.

Global Cognition

Only one study[44] reported the L-D effects on global cognition using the MMSE and the MoCA. The comparison between the "On" (L-D-200 mg/carbidopa-100 mg) and "Off" treatment states of the 17 patients showed no significant difference on the two measures.

Attention/Processing Speed/Working Memory

A study ($n = 10$)[39] reported a deterioration of large effect size on the Stroop-Word condition (information processing speed; $p = 0.028$; $g = 0.754$) at 12 months of treatment. The results for months 3, 6, 18 and 24 were in the same direction and had medium effects sizes, but did not reach statistical significance (p from 0.063 to 0.116; g from 0.503 to 0.614). There were no significant differences reported on the SRT and the Stroop Color condition.

This study,[39] along with most other studies that administered the Digit Span test ($n = 20$;[40] 40;[41] 387[37]), did not report any significant effects for this task. However, another study ($n = 20$)[43] reported an improvement of medium size ($p = 0.027$; $g = 0.515$) on the Digit Span Backward after 2 months of treatment with 450 mg/day of L-D.

No significant change was observed after treatment on most tests of visuo-spatial attention and working memory (Trail Making Test (TMT)-A, New Dot, Corsi and Spinnler Matrices tests).[37,39,40,43] However, a small benefit of L-D ($p < 0.001$; $g = 0.174$) was obtained on the Symbol Digit Modalities Test (SDMT) after 6 months of treatment.[37]

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 2A: Characteristics of participants and medications in studies on the effects of levodopa on cognition

		Patient characteristics								Medication			Study characteristics			
Refs.	**Study group**	***n***	**% ♂**	**Age (years)**	**Educ. (years)**	**H&Y**	**Dur. (years)**	**MMSE**	**UPDRS**	**Rx**	**mg/day**	**%Max**	**Design**	**Blindness**	**Duration**	**MQ (%)**
[37,38]	Gr. 1 (iPD)	387	68	63.7 (9.7)	14.2 (3.3)	2.2 (0.5)	3.9 (1.6)	nil	nil	L-D + DDCI	286	14.3	PP-WG	n/a	6 mth.	67.4
[39]	Gr. 1 (L-D)*	10	30	67.3 (7.5)	5.2 (4.1)	1.8 (0.5)	1.2 (0.5)	nil	24.8 (15.1)	L-D (0-6 mth.) L-D (6-24 mth.)	435 540	5.4 6.8	RCT	Single	24 mth.	65.2
	Gr. 2 (PER)*	10	40	63.7 (10.5)	6.7 (6.4)	1.5 (0.6)	1.2 (0.6)	nil	22.9 (13.8)	PER (0-24 mth.) L-D (6-24 mth.)	2.8 380	56.0 4.8				
[40]	Gr. 1 (iPD)	20	65	57.0 (9.3)	10.6 (3.5)	≤ 2.5	2.5 (1.3)	nil	29.3 (3.6)**,***	L-D or PRX	458 3.9	5.7 86.7	RCO	None	18 wks.	47.8
[41]	Gr. 1 (rotigotine)	20	55	56.0 (5.6)	10.2 (2.7)	≤ 2.5	2.3 (1.4)	≥ 27	29.0 (3.1)**,***	L-D or Rotigotine	358 8	4.5 50.0	RCO	None	6 mth.	45.7
	Gr. 2 (cabergoline)	20		57.0 (2.1)	10.5 (3.4)	≤ 2.5	3.1 (0.5)	≥ 27	28.1 (4.2)**,***	L-D or Cabergoline	358 6	4.5 52.2				
[42]	Gr. 1 (iPD)	12	50	59.9 (2.5)	nil	2.3 (0.2)	3.5 (0.3)	28.4 (0.5)***	31.0 (4.4)	L-D + DDCI	250	12.5	PP-WG-C	None	2 wks.	43.5
	Gr. 2 (HCS)	10	50	55.2 (3.1)	nil	n/a	n/a	29.9 (0.1)	n/a	n/a	n/a	n/a				
[43]	Gr. 1 (iPD)	20	65	58.0 (7.8)	10.1 (2.9)	≤ 2.5	2.6 (1.8)	nil	30.0 (4.1)**,***	L-D or PER	450 3.2	5.6 64.0	RCO	None	18 wks.	43.5
[44]	Gr. 1 (iPD)	17	88	63.2 (8)	nil	2.1 (0.3)	3.7 (3.0)	28.9 (1.0)	17.2 (8)	L-D + DDCI or Pb	200 n/a	10.0 n/a	PP-WG-PC	Single	< 5 d.	34.8
[44]	Gr. 2 (HCS)	21	76	63.3 (7)	nil	n/a	n/a	29 (1.0)	n/a	n/a	n/a	n/a				

% ♂ = mean percentage of male participants in the group; %Max = percentage of maximal daily dose recommended; Age = mean age of participants in years; d. = days; Dur. = mean disease duration in years; Educ. = mean duration of formal education expressed in years; Gr. = group; H&Y = mean stage of Parkinson's disease progression on the Hoehn & Yahr scale; HCS = healthy control subjects; iPD = pharmacologically treated participants with idiopathic Parkinson's disease; L-D = levodopa; L-D + DDCI = levodopa + dopa decarboxylase inhibitor; mg/day = mean daily dose expressed in milligrams per day; MMSE = mean total score on the *Mini-Mental State Examination*; mth. = months; MQ = Methodological Qualiy score; *n* = number of participants; n/a = not applicable; nil = information not available; Pb = placebo; PER = pergolide; PP-WG = pretest-posttest within-group design; PP-WG-C = pretest-posttest within-group controlled design; PP-WG-PC = pretest-posttest within-group placebo-controlled design; PRX = pramipexole; RCT = randomized controlled trial; RCO = randomized cross-over trial; Ref. = reference; Rx = medication; UPDRS = Unified Parkinson's Disease Rating scale; wks. = weeks.

All numerical data are expressed in mean (SD) or mean [range], except for *n* and study duration.

**de novo* patients at baseline.

***Authors specify that the score was obtained on the UPDRS section 3 (UPDRS-III) only.

Table 2B: Characteristics of participants and medications in studies on the effects of pramipexole on cognition

		Patient characteristics								Medication			Study characteristics			
Ref.	**Study group**	***n***	**% ♂**	**Age (years)**	**Educ. (years)**	**H&Y**	**Dur. (years)**	**MMSE or MoCA**	**UPDRS**	**Rx**	**mg/ day**	**% Max**	**Design**	**Blindness**	**Duration**	**MQ (%)**
[40]	Gr. 1 (iPD)	20	65	57.0 (9.3)	10.6 (3.5)	≤2.5	2.5 (1.3)	nil	29.3 (3.6)*,**	L-D or PRX	458 3.9	5.7 86.7	RCO	None	18 wks.	47.8
[45]	Gr. 1 (PRX)***	10	60	60.1 (10.1)	10.0 (3.9)	2.3 (0.4)	3.1 (1.9)	28.8 (1.3)	16.7 (6.7)	PRX	2.1	46.7	RCT	None	3 mth.	45.7
	Gr. 2 (PER)***	9	56	57.9 (9.4)	10.2 (3.9)	2.2 (0.5)	3.5 (1.9)	28.2 (1.3)	13.4 (7.8)	PER	3	60.0				
	Gr. 3 (HCS)	13	85	60.3 (8.2)	11.2 (4.2)	n/a	n/a	28.9 (1.2)	nil	n/a	n/a	n/a				
[46]	Gr. 1 (PRX)	30	57	61.7 (14.2)	12.3 (4.4)	2.0 (0.8)	4.6 (4.8)	28.5 (2.2)	nil	PRX and L-D	3 nil	66.7 n/a	RCT	None	6 mth.	34.8
	Gr. 2 (L-D)	25	60	63.0 (12.7)	11.9 (4.3)	2.3 (1.0)	5.4 (3.1)	28.2 (2.1)	nil	L-D	nil	n/a				
	Gr. 3 (HCS)	25	44	62.5 (15.0)	12.5 (2.8)	n/a	n/a	28.9 (1.3)	nil	n/a	n/a	n/a				
[47]	Gr. 1 (iPD)	7	86	58.6 (6.0)	nil	nil	6.8 (2.9)	28.1 (1.3)	28.6 (6.8)**	PRX	1	22.2	PP-WG	n/a	<1 wks.	34.8

% ♂ = mean percentage of male participants in the group; %Max = percentage of maximal daily dose recommended; Age = mean age of participants in years; Dur. = mean disease duration in years; Educ. = mean duration of formal education expressed in years; Gr. = group; H&Y = mean stage of Parkinson's disease progression on the Hoehn & Yahr scale; HCS = healthy control subjects; iPD = pharmacologically treated participants with idiopathic Parkinson's disease; L-D = levodopa; mg/day = mean daily dose expressed in milligrams per day; MMSE = mean score on the *Mini-Mental State Examination*; MoCA = Montreal Cognitive Assessment; mth. = months; MQ = Methodological Qualiy score; *n* = number of participants; n/a = not applicable; nil = information not available; PER = pergolide; PP-WG = pretest-posttest within-group design; PRX = pramipexole; RCT = randomized controlled trial; RCO = randomized cross-over trial; Ref. = reference; Rx = medication; UPDRS = Unified Parkinson's Disease Rating Scale; wks. = weeks.

All numerical data are expressed in mean (SD) or mean (range), except for *n* and study duration.

*Authors specify that the score was obtained on the UPDRS section 3 (UPDRS-III) only.

**Authors specify the patients were measured in "*Off*" state.

****de novo* patients at baseline.

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 2C: Characteristics of participants and medications in studies on the effects of selegiline on cognition

		Patient characteristics								Medication			Study characteristics			
Ref.	**Study group**	***n***	**% ♂**	**Age (years)**	**Educ. (years)**	**H&Y**	**Dur. (years)**	**MMSE**	**UPDRS**	**Rx**	**mg/ day**	**% Max**	**Design**	**Blindness**	**Duration**	**MQ (%)**
[38,48]	Gr. 1 (Pb)*	174	66	61.1 (9.5)	14.3 (3.4)	≤ 2.0	≤ 5.0	28.8 (1.4)	nil	Pb	n/a	n/a	RPC	Double	24 mth.	84.8
	Gr. 2 (tocopherol)*	174								Tocopherol	2000**	133.3				
	Gr. 3 (SEL)*	187								SEL	10	100.0				
	Gr. 4 (SEL + tocopherol)*	175								SEL and Tocopherol	10 2000**	100.0 133.3				
[49]	Gr. 1 (Pb-first)	11	50	66.9 (5.4)	10.6 (2.6)	1.0 (1-2)	1.3 (1.8)	nil	10.0 (4-20)***	Pb (0-2 mth.) SEL (2-4 mth.)	n/a 10	n/a 100.0	RCO-P	Double (0-2 mth.)	4 mth.	63.0
	Gr. 2 (SEL-only)	9		64.3 (12.6)	12.1 (3.3)	2.0 (1-3)	2.2 (1.4)	nil	13.0 (8-17)***	SEL (0-2 mth.)	10	100.0				
	Gr. 3 (HCS)	11	54	61.6 (11.9)	11.5 (1.9)	n/a	n/a	nil	n/a	n/a	n/a	n/a				
[50]	Gr. 1 (SEL)	17	88	55.4 (5.8)	nil	nil	3.7 (2.4)	nil	24.1 (14.1)	SEL	10	100.0	PP-C	None	3 mth.	23.9
	Gr. 2 (Ø Rx)	15	80	56.7 (6.3)	nil	nil	3.9 (3.1)	nil	40.5 (18.5)	n/a	n/a	n/a				

% ♂ = mean percentage of male participants in the group; %Max = percentage of maximal daily dose recommended; Ø Rx = untreated participants with idiopathic Parkinson's disease; Age = mean age of participants in years; Dur. = mean disease duration in years; Educ. = mean duration of formal education expressed in years; Gr. = group; H&Y = mean stage of Parkinson's disease progression on the Hoehn & Yahr scale; HCS = healthy control subjects; mg/day = mean daily dose expressed in milligrams per day; MMSE = mean score on the *Mini-Mental State Examination*; mth. = months; MQ = Methodological Qualiy score; *n* = number of participants in study group; n/a = not applicable; nil = information not available; Pb = placebo; PP-C = pretest-posttest controlled design; RCO-P = randomized cross-over placebo-controlled trial; Ref. = reference; RPC = randomized placebo-controlled trial; Rx = medication; SEL = selegiline; UPDRS = Unified Parkinson's Disease Rating Scale; wks. = weeks. All numerical data are expressed in mean (SD) or mean (range), except for *n* and study duration.

**de novo* patients at baseline.

**Data presented in IU.

***Authors specify that the score was obtained on the UPDRS section 3 (UPDRS-III) only.

Table 2D: Characteristics of participants and tested medications in studies on the effects of rasagiline on cognition

		Patient characteristics							Medication			Study characteristics			
Ref.	Study Group	*n*	% ♂	Age (years)	Educ. (years)	H&Y	Dur. (years)	MMSE	UPDRS	Rx	mg/day	%Max	Design	Blindness	Duration	MQ (%)
[36]	Gr. 1 (RAS)	23	74	65.2 (9.5)	8.1 (4.2)	2.0 (0.7)	4.1 (2.5)	nil	23.1 (10.0)**	RAS	1	100.0%	RPC	Double	3 mth.	52.2
	Gr. 2 (Pb)	25	64	67.6 (10.1)	8.4 (3.9)	1.6 (0.6)	4.0 (2.3)	nil	18.8 (8.6)**	Pb	n/a	n/a				

% ♂ = mean percentage of male participants in the group; %Max = percentage of maximal daily dose recommended; Age = mean age of participants in years; Dur. = mean disease duration in years; Educ. = mean duration of formal education expressed in years; Gr. = group; H&Y = mean stage of Parkinson's Disease progression on the Hoehn & Yahr scale; mg/day = mean daily dose expressed in milligrams per day; MMSE = mean score on the *Mini-Mental State Examination*; mth. = months; MQ = Methodological Quality score; *n* = number of participants in study group; n/a = not applicable; nil = information not available; Pb = placebo; RAS = rasagiline; RPC = randomized placebo-controlled trial; Ref. = Reference; Rx = medication; UPDRS = Unified Parkinson's Disease Rating scale; wks. = weeks.
All numerical data are expressed in mean (SD) or mean [range], except for *n* and study duration.
**Authors specify that the score was obtained on the UPDRS section 3 (UPDRS-III) only.

Episodic Memory

Three different verbal episodic memory tests were administered in seven L-D studies: the Buschke Selective Reminding Test (BSRT),[37] the California Verbal Learning Test (CVLT)[40,43] and the Rey Auditory Verbal Learning Test (RAVLT);[39,41] however, significant differences were only observed on the RAVLT. Improvements on the RAVLT-trials 1-5 (verbal encoding and learning) were reported after 3,[41] 12 and 24 months[39] of treatment with medium ($p = 0.016$; $g = 0.568$ at 3 months) to large effect sizes ($p = 0.016$ and $g = 0.859$ at 12 months; $p = 0.020$ and $g = 0.819$ at 24 months). On the delayed recall trial (retention over time), no significant effect was reported.

Regarding visual episodic memory, L-D did not have any effect on the BVRT scores, but the performance on the Rey-Osterieth Complex Figure Test (ROCFT) improved. Improvements of large effect sizes (p from 0.006 to 0.016; g from 0.852 to 1.029) were recorded at 6, 12 and 24 months post baseline evaluations on the ROCFT delayed recall.[39]

Construction and Motor Performance

The trend toward beneficial effects of treatment increased at each post-baseline evaluation on the FTT (left hand), reaching statistical significance and large effect sizes at months 18 and 24 (from $p = 0.024$ and $g = 0.783$ at month 18 to $p = 0.009$ and $g = 0.961$ at month 24).[39] No change was reported on the ROCFT-Copy and on the Grooved Pegboard tests.[37,39]

Concept Formation and Reasoning

A small beneficial effect was obtained on the first of two administrations of the odd man out (OMO) test ($p = 0.002$; $g = 0.165$) after 6 months of treatment.[37] L-D treatment yielded no significant effect on Raven's Progressive Matrices (RPM), Weschler Adult Intelligence Scale (WAIS)-Arithmetics and modified Wisconsin card sorting test (mWCST) performances.[40,41,43]

Executive Functions

Three studies[40,41,43] yielded deleterious effects of medium to large effect size (p from <0.001 to 0.002 and g from 0.779 to 1.736) exclusively on the Stroop Color/Word (inhibition) condition, with the administration of 358-458 mg/day of L-D during 18 to 24 weeks. A divergent result came from one study ($n = 12$)[42] reporting an improvement on the inhibition condition of a modified version of the Stroop ($p < 0.03$) using an "On-/Off-treatment" paradigm with 250 mg/day of L-D + DDCI.

Treatment was most frequently associated with beneficial effects in four studies using the verbal fluency test.[37,40,41,43] Two studies[41,43] showed beneficial effects of medium size on the Lexical Fluency task ($p = 0.031$; $g = 0.501$ at 2 months[43] and $p = 0.046$ and $g = 0.457$ at 3 months[41]). The performance on the Category Fluency task also improved with small,[37] medium[43] and large[40] effect sizes (p from 0.003 to 0.046 and g from 0.140 to 0.738 after 2 or 6 months of treatment).

On the Luria tasks, large beneficial effects of treatment were reported on the motor tasks from months 3 to 24 (p from <0.001 to 0.005; g from 1.071 to 1.636). There was no significant effect of treatment for the rhythm reproduction task.

On the Tower of London test (procedural learning; mental flexibility), patients' performance improved with medium effect

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 3: Cognitive measures of efficacy

Measures categorized by main cognitive domain assessed	Description	Scoring*	14**
Global cognition			
MMSE[30]	Test commonly used in clinical practice to globally screen the cognitive functions of a patient. It briefly assesses spatio-temporal orientation, verbal learning and recall, attention and calculation, language and visuo-construction. Scoring: It includes 30 questions, each worth 1 point, for a maximum score of 30 points. Depending on the normative data, the total score may be adjusted for socio-economical status, educational level and duration of administration.	Positive	2
MoCA[51]	Global cognition test designed to screen for mild cognitive impairment (MCI) and dementia. It briefly assesses spatio-temporal orientation, sustained attention, serial subtraction, auditory working memory, sentence repetition, verbal fluency, abstract thinking, visuospatial/executive functions, verbal learning and recall. Scoring: it is scored on a 30-point scale, with each question worth 1 point	Positive	1
WAIS,[52] translated for Indian participants	The Indian translation of the third edition of the Weschler Intelligence Scales for adults (16 years or older). This cognitive test battery includes tests that tap into a wide range of cognitive functions, including abstract perceptual and verbal reasoning, auditory working memory, verbal and visual attention, semantic memory and processing speed, among others. Scoring: each test score is integrated into a composite score and composite scores are transformed into a global IQ score	Positive	1
Attention/processing speed/working memory			
Corsi Test[53]	Visuo-spatial working memory task consisting of nine identical cubes placed in a random order on a board. The examiner taps the blocks in sequences and the examinee must reproduce the same tapping patterns immediately after presentation. Scoring: number of tapping sequences correctly reproduced	Positive	2
Digit Span[54-56]; DOT-A[57]	Auditory attention span tests. Patients must repeat a series of numbers presented orally, in the same order as it was dictated (Digit span Forward), in reverse order (Digit Span Backwards) or in ascending order (DOT-A). Scoring: number of correctly repeated sequences Psychometric characteristics: cited by the MDS task force guidelines[1] as good examples of attention and working memory tests for PD patients; High internal consistency, excellent inter-rater reliability and good construct validity[58]; Particularly sensitive for patients with PD, sufficient reliability, good validity[57]	Positive	8
SDMT[59]	Task consisting of copying numbers associated with specific symbols according to a visually presented key, as fast as possible, in 90 seconds. It is also possible to ask for an oral response in addition or in replacement of the written one. Scoring: number of digits correctly copied. Psychometric characteristics: good sensitivity and validity to assess information processing speed[60]. The oral-format SDMT demonstrated excellent test-retest reliability[61]; Reasonable test-retest reliability, and subjects who completed the same form revealed significant practice effects, which were almost non-existent in those filling different forms[62]	Positive	2
Spinnler Matrices[63]	Test measuring selective attention in a visual-search condition. The patient must search for target numbers hidden in a matrix that also presents distractor numbers. Scoring: number of correct items found within the 45-second delay	Positive	3
Spot-the-dot task; New dot test[64,65]	Visuo-spatial working memory test, where the patient has to remember where he saw a specific dot on a computer screen, in order to later locate this target dot among distractor dots. Another version (New Dot test) also exists, in which the patient must identify the new dot among other dots previously shown on the screen. He thus must learn the spatial location of the other dots to achieve this goal. Scoring: number of dots correctly located	Positive	2
SRT[66]	Task consisting of pushing a button or a keyboard key as fast as possible every time a stimulus is presented. The stimuli may be in the visual or auditory modalities. Scoring: reaction time in milliseconds (the shorter the better)	Negative	2
Stroop[67]	Condition 1/color: the stimuli sheet presents differently colored rectangles; the patient must name the color); Condition 2/word: The stimuli sheet presents the names of a three colors printed in black ink; the patient must read the name of the colors out loud. Each sheet contains several lines of stimuli, which must be named/read entirely, as fast as possible, from left to right, line after line. Scoring: completion time or time per 100 stimuli (depending on the version), in seconds (the shorter the better), and number of errors/condition. Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of cognitive tests to use with PD patients; good temporal reliability[68]	Negative	6
Modified Stroop task[42]	This is a computerized task based on the original Stroop test, designed to be easily administered to participants while inside an fMRI scanner and without using verbal responses. There are three conditions; the first two assess attentional functions and processing speed: "Sensorimotor" (control task; identify the "foreign" target among four possible choices), and "Neutral" (cognitive task; select one of four color names correctly representing the color of a presented square). The third condition is named "Incongruent" and is described in the "Executive functions" section of this table. Scoring: Accuracy (% of correct responses relative to errors), Interference effect (["Incongruent" RT]—["Neutral" RT]), Differences in RT/ΔRT (["Sensorimotor" RT]—["Neutral" RT] OR ["Incongruent" RT])	% = positive RT = negative	1

TMT-A[69]	The stimuli sheet presents several circles, randomly arranged across the page, each containing a number. The patient has to connect the circles in ascending order as fast as possible. Scoring: completion time in seconds (the shorter the better). Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of cognitive test to use with PD patients; Excellent test-retest reliability[70]	Negative	6
Perception			
BFR[71]	This task consists of matching the picture of a face with one or more of the six response choices representing different angles or lighting of the same target person or pictures of faces other than the target one. There are three conditions for representation of the faces on the multiple-choice response cards: a front view (identical to the view on the picture the patient has), a front with side view or a front view taken under different lighting conditions. Scoring: number of correctly matched pictures Psychometric characteristics: a reliability correlation of 0.60 (questionable) was obtained after a 1-year retesting of elderly control subjects[72]. Practice effects appear to be mostly negligible[73]	Positive	1
BJLO[71]	This test includes 30 items in which participants must match two angled lines on top of the booklet to a set of 11 angled lines that are arranged in a semicircle at the bottom of the booklet. Scoring: number of items on which judgement for both lines are correct, for a possible maximum score of 30 points. Psychometric characteristics: acceptable construct validity and criterion-related validity[74]	Positive	1
Memory			
BSRT[75]	Test procedure aiming at differentiating the memory processes of encoding, storage, retention and retrieval. It usually involves exposing the patient to a list of words (either verbally or in written/image form) for a number of immediate free and cued recall trials. From the second trial onward, patients are only told/shown the words they omitted on the previous trial. The number of trials depends on the version of the test procedure used. A delayed free and cued recall trial (usually after 30 minutes) is then administered, and is sometimes followed by a four-choice recognition trial. Scoring: number of correctly recalled words/trial	Positive	2
BVRT[76]	Visual recall task that consists of reproducing the figures presented on a card, either immediately or after a 5-15-second delay, depending on the administration procedure used. Most of the cards contain three figures arranged on a horizontal plane, two large and one small, with the small figure always to one side or the other. Scoring: number of correctly recalled designs and number of errors (omissions, distortions, perseverations, rotations, misplacements)	Positive and negative	1
CVLT[77,78]	The test consists of a five-trial list-learning task, where the patient has to learn a list of 16 orally presented words (list A), each belonging to one of four semantic categories. This learning phase is followed by an immediate recall of a different 16-word list (interference list—list B), and a free and cued recall of list A. This is followed by a 20-minute delayed free and cued recall of list A. The task ends with a recognition trial of list A. Scoring: number of correctly recalled or recognized words/trial. Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of episodic memory tests for PD patients; good test-retest reliability and good internal consistency[79]	Positive	2
Explicit motor sequence learning task[44]	Computerized task designed to be easily administered inside an fMRI scanner. Patients must press one of four buttons as fast as possible when an "X" appears on the screen. Each button corresponds to a stimulus box on the screen, in which the "X" appears in a determined order or random six-element sequences. Each sequence is followed by a 20-second pause, during which the patients are informed whether the next sequence is going to appear in a random order or not. Scoring: learning magnitude measured in error or in RT (milliseconds, the shorter the better).	Negative	1
RAVLT[80]	The test consists of a five-trial list-learning task, where the patient must learn a list of 15 orally presented words, none of which are semantically related. This learning phase is followed by an immediate recall of a different 15-word list (interference list) and a free recall of the original list. This is followed by a 30-minute delayed free recall and a recognition trial of the original list. Scoring: number of correctly recalled or recognized words/trial. Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of episodic memory tests for PD patients; high test-retest reliability[81]	Positive	2
RBMT[82]	Battery of ecological tests measuring memory functions mandatory in daily living. It is composed of 10 tasks: "Remembering a name" (associated to a photograph), "Remembering a hidden belonging" (the examiner hides an object that belongs to the patient and instructs him to remember where it is hidden and to ask for it after a specific cue is given), "Remembering an appointment" (the patient must ask about it after a 20-minute delay indicated by an alarm), "Remembering a newspaper article" or "Story recall" (once after hearing the story and once after a 20-minute delay), "Face recognition" (five pictures seen a few minutes earlier have to be identified among a group of 10), "Remembering a new route" (that the examiner traces among predetermined points in the examination room, immediate and 10-minute delayed recall), "Delivering a message" (during the route-recall; the message is given before setting out on the route), "Orientation" (time and place), "Date", "Picture recognition" (10 pictures are shown to the patient, who is then asked a few moments later to identify them after they have been mixed with 10 distractor pictures). Scoring: "Total Score": Sum of the test scores that make up a test profile; "Total Screening Score": Combination of the screening scores of all tasks (except for "Delivering a message"), which are based on a pass/fail criteria. Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of episodic memory tests for PD patients; high validity and inter-rater reliability[83]	Positive	1

Table 3. *Continued*

Measures categorized by main cognitive domain assessed	Description	Scoring*	14**
ROCFT[80]	Reproduction of a complex figure composed of 18 elements, first in a copy condition and then in immediate (3 minutes after copy) and delayed (30 minutes after copy) recall conditions. A recognition trial is then administered. The copy condition is used to assess construction praxis. A "programming" trial is available if the copy is failed, in which a construction strategy is given to the patient, in order to minimize the executive demands of the task and better isolate its visuo-spatial component. Scoring: Many scoring systems exist for this task. The original and most used scoring system grants a score of 0, 1 or 2 for each of the 18 elements, depending on the quality of its reproduction, with a maximum score of 36. Psychometric characteristics: Reliable tool for patients with memory disease[84]	Positive	1
Verbal Paired Associates (WMS-R)[85]	The examiner orally presents a series of eight word pairs. After the last pair is given, the first word of each pair is presented again and the patient has to recall the word that was paired with it. There are four learning trials, in which the pairs and recall prompts are presented in a different order. A 20-30-minute delayed recall trial is then conducted, followed by a recognition trial. Scoring: number of correctly recalled words Psychometric characteristics: WMS-R has acceptable levels of reliability[86]	Positive	1
Visual Reproduction (WMS-R)[85]	Patients must draw what they remember of geometric designs they were previously exposed to for a 10-second period. A 30-minute delayed recall is administered, followed by a 48-item recognition trial. Scoring: the scoring system divides every design in a variable number of parts, each of which may be given a score of 0, 1 or 2, depending on the quality of the reproduction. Psychometric characteristics: WMS-R has acceptable levels of reliability, although the Visual Memory index is susceptible to a number of influences, which raises concerns about its construct validity[86]	Positive	1
Verbal function and language skills			
BNT[87]	Task measuring confrontational word naming, consisting of naming each presented image within a 20-second delay. It involves an ascending difficulty, where images less frequently seen in everyday life are presented increasingly as the test evolves. Scoring: number of correctly named images. It includes a maximum of 60 images, each worth 1 point if named correctly. Validated short 15- and 30-item versions are also available. Psychometric characteristics: its qualities are controversial. It is cited by the MDS task force guidelines[1] as a good example of language tests for PD patients, but at the same time a recent review[88] questions whether it is still fit for its purpose, reporting criticisms such as "poor psychometric properties", "inadequately standardized" and "inadequate norms". Notably, they report a ceiling effect and a non-normal score distribution that renders it less adequate to identify subtle deficits or gradual declines in performance, as seen in conditions such as mild cognitive impairment	Positive	1
Construction and motor performance			
CDT[89]	The patient must draw a clock, including all its parts (frame, numbers, hands), that indicates a given hour (11:10). A copy task is also possible if the free drawing trial has failed. The test allows for qualitative and quantitative scores (frame integrity, number placement and sequencing, presence and placement of the hands). CDT screens several cognitive functions, including visuo-spatial processing and executive functions. Scoring: many scoring systems are available. An example of a 7-point scoring system verifies whether (1) the numbers from 1 to 12 are present; (2) the numbers are properly sequenced; (3) the numbers are positioned well; (4) the 2 clock hands are drawn; (5) the clock hand indicating the hours is positioned correctly; (6) the clock hand indicating the minutes is positioned correctly; (7) the size of the different clock hands are respected. The scoring system used in the original articles of the present review is not always specified	Positive	1
Grooved Pegboard test[90]	Manual dexterity task consisting of a metal board containing a matrix of 25 holes with randomly positioned slots. All of the pegs have a ridge along one side and have to be rotated in a position matching the hole in order to be inserted. The patient must insert the pegs in the holes in sequence, as fast as they can, until all the pegs have been placed. Two trials are conducted, first with the dominant and then the non-dominant hand. Scoring: time (in seconds) required to complete the task for each trial (the faster the better)	Negative	1
ROCFT(copy)[91,92]	This task notably involves memory in addition to construction praxis. Please refer to the "Memory" section of this table for a complete description.	Positive	1
ROT[93]	The task uses two identical apparatus, one for the patient to use and one used as a model to replicate. These apparatus are vertical rods that can rotate on a 360° axis, to the top of which are attached another rod that can be moved up and down in a sagittal plane. The patients must match the displacement of their hinged rod, so that it is identical to the model apparatus. Patients are only allowed to use their preferred hand, and to move their head and their eyes, but not their trunk. There are two conditions to the test, depending on whether the patient is blindfolded (Tactile condition) or not (Visual condition). Vertical and horizontal errors are measured using a goniometer. Scoring: number of errors	Negative	1
FTT[94]	This task measures psychomotor speed and manual dexterity. It consists of repeatedly pressing a keyboard key or the button of an adapted tapper and counter, using the index finger of the dominant hand, as many times as possible, within a given time delay (10-30 seconds), depending on the version administered. A second set of trials is conducted using the non-dominant hand. Scoring: total number of button presses in a set of trials.	Positive	1

Concept formation and reasoning			
APM[95]; RPM[96]	The RPM is a non-verbal problem-solving task, in which the patient must identify the element that correctly completes the presented series among six response choices. Other versions of the test exist, including the APM, which was developed for patients with above average intellectual ability. Scoring: number of correctly solved problems, with a maximum score of 60 points.	Positive	2
Arithmetics[54]	Arithmetic problems are presented orally in a story format, arranged according to the level of difficulty. A time limit, varying with the difficulty of the problem, is imposed to the patient. The patient cannot use visual aids. Scoring: number of correctly solved problems within the time limit	Positive	1
OMO[97]	The patient must designate which set of letters or geometrical shapes differs from the others on two decks of cards, using two rules of classification alternately on two successive trials. Each deck includes eight cards/trial (total of two trials = 16 cards). Each card displays three letters or shapes. Of those three items, two are similar and the third is different. The classification rule can be the size or shape of the item. On the first trial, the rule is decided by the patient, and on the second trial the patient is asked to use another rule, without stating explicitly what those rules are. This rule alternation continues up to item eight, and then the rules are told to the patient, if he has not already guessed them. Scoring: number of correct choices on each trial; number of errors made	Positive and negative	2
WCST[98]; mWCST[99]	This task consists of two decks of 64 cards each that the patient has to sort in one of four categories indicated by four stimulus cards, according to some undisclosed principle. The patient must deduce the sorting principle from the ambiguous feedback of the examiner (who only tells whether the placement of each card is "right" or "wrong"). On every card appears one to four symbols (triangle, star, circle or cross) printed in either red, green, yellow or blue, which are disposed in a variable pattern, making every card different from one another. The sorting principle is either by color, form or number, and it must be shifted by the examiner after a run of 10 correct placements, without informing the patient of this sudden change. The test stops after 6 runs of 10 correct placements or after all the cards have been placed. Several versions of this test are available, including the mWCST, in which the number of cards has been reduced to 2 decks of 24 cards, the number of trials before switching principle is lowered to six and the patient is informed of the category switching. Scoring: number of cards correctly sorted (total correct), number of cards incorrectly sorted (total errors), number of correct runs from 10 (or 6) sorts (categories), number of cards sorted when the patient persists in responding according to an incorrect and/or previously successful principle (perseverative errors/responses), number of errors that are not classified as perseverative (non-perseverative errors), number of trials to complete the first category, achieved by giving 10 (or 6) correct responses in a row (trials to first category) Psychometric characteristics: cited by the MDS task force guidelines[1] as good examples of cognitive tests to use with PD patients; controversial test-retest reliability, with some qualifying it as poor [mWCST],[100] and others reporting that normal controls only modestly improved after repeated testing, with results not pointing toward a strong practice effect[101]	Positive and negative	3
Executive functions			
Delay Discounting Task[102]	In this computerized task, patients in a PET scanner must choose between a small immediate reward (fictional money) and a larger delayed reward. For each trial, the length of the delay varies (× weeks/months/years). Thus, this task is a measure of impulsivity level. Scoring: it uses a *k*-value derived from an hyperbolic function that takes into account the value of the reward and the length of the delay. Psychometric characteristics: modest levels of reliability. Some findings raised questions about the commonly assumed relationship between discounting and the construct of impulsivity[103]	Negative	1
Go/No-go Task[104]	In this computerized task, patients in a PET scanner must press a button as fast as they can when the target stimulus is displayed (a white circle) and to withhold their response when the distractor stimulus is displayed (a white X). No other stimuli are used during the test sequence. Scoring: RT in milliseconds (the shorter the better) and errors (anticipation, commission)	Negative	1
Luria tasks[105,106]	"Motor Test": section of the *Dynamic Organization* part of the evaluation of the motor functions of the hands from the "Luria's neuropsychological investigation" battery.[106] The patient is asked to rest his hands on a table and place them in three successive positions, namely "fist", "edge" and "palm". In the "fist" position, a fist is formed with the hand and is placed on the table such that the back of the hand is visible. In the "edge" position, the edge of the hand is the only point of contact with the table, placing the hand in a vertical position. In the "palm" position, the hand is placed horizontally on the table, in a way that allows the palm to be in contact with it. Before starting his sequence, the patient has to tell the examiner what he intends to do, in order to ensure that failed attempts are not due to a memory defect. "Rhythm reproduction": This refers to the *Motor performance of rhythmic groups* segment of the evaluation of perception and reproduction of rhythmic structures from the "Luria's neuropsychological investigation" battery.[106] First, the patient has to use his fingers to replicate the rhythmic tapping sequences demonstrated by the examiner. Then, the examiner gives verbal direction for the production of rhythmic sequences, which the patient must reproduce via finger tapping. Before performing its reproductions, the patient must tell the examiner what he intends to do, in order to ensure that failed attempts are not due to a memory defect. Scoring: for the motor test, points are given according to the number of sequences correctly performed, and for the rhythm task the points are derived from the number of errors committed	Positive and negative	1
Stroop[67]	Condition 3/Color-word: The stimuli sheet presents the name of three colors, printed in a colored ink, but the written word never matches the color of the ink it is printed in; the patient has to name the color of the ink without reading the written word. It is a measure of cognitive inhibition. The sheet contains several lines of stimuli, which must be named entirely, as fast as possible, from left to right, line after line. Scoring: completion time or time per 100 stimuli (depending on the version), in seconds (the shorter the better), and number of errors/condition. Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of cognitive tests to use with PD patients; test-retest reliability is high with children, as well as with adults[107]	Negative	6

Table 3. *Continued*

Measures categorized by main cognitive domain assessed	Description	Scoring*	14**
Modified Stroop task[42,67]	This is a computerized task based on the original Stroop test, designed to be easily administered to participants while inside a fMRI scanner without using verbal responses. There are three conditions, and only the third response assesses executive functions, namely cognitive inhibition: "Incongruent" (cognitive task; select one of the four color names correctly representing the ink in which is written another color name that is incongruent to the meaning of the presented name). Scoring: accuracy (% of correct responses relative to errors), interference effect (["Incongruent" RT]—["Neutral" RT]), Differences in RT/ΔRT (["Sensorimotor" RT]—["Neutral" RT] OR ["Incongruent" RT])	Positive and negative	1
Tower of London[108]	The test uses one board with three pegs, beads of different colors and a booklet presenting images of the problem-solving tasks. To reach the best possible solution, the participant must think ahead in order to determine the optimal order of moves necessary to replicate the bead pattern displayed on the image in the booklet. Two main rules must be respected. First, the patient is only allowed to move one bead at a time, and only from one peg to another. Second, each peg can hold up a different maximum number of beads. The patient has a maximum of three trials to solve each problem. Scoring: there are 12 problems, each worth a maximum of 3 points, for a maximum possible score of 36 points. The number of points awarded for each problem depends on the number of trials necessary to solve it (1 trial = 3 points; 2 trials = 2 points; 3 trials = 1 point). Psychometric characteristics: despite poor to questionable internal consistency for this scoring system, the discriminative validity and clinical utility for assessing planning deficits in PD-MCI is high[109]	Positive	1
TMT-B[69]	The stimuli sheet presents several circles, randomly arranged across the page, each containing either a letter or a number. The patient must connect the circles as fast as possible, alternating between ascending numbers and letters in alphabetical order. It measures mental flexibility. Scoring: completion time in seconds (the shorter the better) Psychometric characteristics: cited by the MDS task force guidelines[1] as a good example of cognitive tests to use with PD patients; good sensitivity to brain damage. Excellent test-retest reliability[70]	Negative	5
Verbal Fluency tests[110,111]; COWAT[112]	These tests measure the initiation capacity. It consists of either producing as many words as possible that begin with a given letter (e.g., F, A, S) or producing as many words as possible belonging to a given semantic category (e.g., animals), within a 60-second delay/trial. Scoring: number of correct words/trial. Psychometric characteristics: cited by the MDS task force guidelines[1] as good examples of executive functions tests for PD patients; The COWAT has adequate test-retest reliability[113]	Positive	7

APM = Advanced Progressive Matrices; BFR = Benton Facial Recognition; BJLO = Benton Judgment of Line Orientation; BNT = Boston Naming Test; BSRT = Buschke Selective Reminding Test; BVRT = Benton's Visual Retention Test; CDT = Clock Drawing Test; COWAT = Controlled Oral Word Association Test; CVLT = California Verbal Learning Test; DOT-A = Adaptive Digit Ordering Test; FTT = Finger Tapping Test; MDS = Movement Disorders Society; MMSE = Mini-Mental State Examination; mWCST = Modified Wisconsin Card Sorting Test; MoCA = Montréal Cognitive Assessment; OMO = Odd Man Out; RAVLT = Rey Auditory Verbal Learning Test; RBMT = Rivermead Behavioral Memory Test; ROCFT = Rey-Osterrieth Complex Figure Test; ROT = Rod Orientation Test; RPM = Raven's Progressive Matrices; RT = Reaction time; SDMT = Symbol Digit Modalities Test; SRT = Simple Reaction Time; TMT = Trail Making Test; WAIS = Weschler Adult Intelligence Scale; WCST = Wisconsin Card Sorting Test; WMS-R = Wechsler Memory Scale Revised.

Cognitive domain classification is based on Lezak, Howieson et al.[114] Some descriptions are based on Strauss et al[115] and Lezak et al.[114]

*Positive scoring implies that the higher the score is, the better the performance is, whereas negative scoring implies that the higher the score is, the worse the performance is.

**Number of selected original articles using the described cognitive task as an outcome measure.

Table 4A: Effects of levodopa on cognitive outcome measures

Ref.	Measures of efficacy	Results	
[37]	Digit span	–	
	Forward	BL versus 06 mth: ns, $g = 0.034$ (−0.071; 0.140)	
	Backward	BL versus 06 mth: ns, $g = 0.000$ (−0.105; 0.105)	
	New Dot Test	BL versus 06 mth: ns, $g = 0.053$ (−0.049; 0.154)	
	SDMT	BL versus 06 mth: ↑, $p < 0.001$, $g = 0.174$ (0.071; 0.276)	
	BSRT	–	
	Total recall	BL versus 06 mth: ns, $g = -0.030$ (−0.132; 0.071)	
	Delayed recall	BL versus 06 mth: ns, $g = 0.000$ (−0.102; 0.102)	
	OMO	–	
	Form 1	BL versus 06 mth: ↑, $p = 0.002$, $g = 0.165$ (0.059; 0.271)	
	Form 2	BL versus 06 mth: ns, $g = 0.097$ (−0.009; 0.202)	
	Verbal fluency	BL versus 06 mth: ↑, $p = 0.010$, $g = 0.140$ (0.034; 0.246)	
[39]	Digit Span	BL versus 03 mth: ns, $g = -0.055$ (−0.676; 0.565) BL versus 06 mth: ns, $g = 0.051$ (−0.569; 0.672) BL versus 12 mth: ns, $g = 0.190$ (−0.435; 0.816)	BL versus 18 mth: ns, $g = 0.306$ (−0.328; 0.940) BL versus 24 mth: $p = 0.110$, $g = -0.513$ (−1.172; 0.146)
	SRT	–	–
	Visual RT (ms)	BL versus 03 mth: ns, $g = -0.375$ (−1.016; 0.266) BL versus 06 mth: $p = 0.075$, $g = -0.582$ (−1.252; 0.089) BL versus 12 mth: ns, $g = -0.479$ (−1.134; 0.175)	BL versus 18 mth: $p = 0.103$, $g = -0.525$ (−1.186; 0.136) BL versus 24 mth: ns, $g = 0.000$ (−0.620; 0.620)
	Aleatory visual RT (ms)	BL versus 03 mth: ns, $g = -0.048$ (−0.668; 0.572) BL versus 06 mth: ns, $g = -0.272$ (−0.903; 0.360) BL versus 12 mth: ns, $g = -0.207$ (−0.833; 0.420)	BL versus 18 mth: ns, $g = -0.288$ (−0.921; 0.344) BL versus 24 mth: ns, $g = 0.220$ (−0.407; 0.848)
	Auditory RT (ms)	BL versus 03 mth: ns, $g = -0.037$ (−0.657; 0.583) BL versus 06 mth: $p = 0.094$, $g = -0.542$ (−1.205; 0.122) BL versus 12 mth: ns, $g = -0.240$ (−0.869; 0.389)	BL versus 18 mth: ns, $g = -0.310$ (−0.944; 0.325) BL versus 24 mth: ns, $g = -0.312$ (−0.947; 0.323)
	Stroop (computerized)	–	–
	Word (RT, ms)	BL versus 03 mth: $p = 0.063$, $g = 0.614$ (−0.061; 1.290) BL versus 06 mth: $p = 0.116$, $g = 0.503$ (−0.155; 1.161) BL versus 12 mth: ↓, $p = 0.028$, $g = 0.754$ (0.052; 1.457)	BL versus 18 mth: $p = 0.079$, $g = 0.573$ (−0.096; 1.241) BL versus 24 mth: $p = 0.072$, $g = 0.584$ (−0.087; 1.255)
	Color (RT, ms)	BL versus 03 mth: ns, $g = 0.256$ (−0.374; 0.885) BL versus 06 mth: ns, $g = -0.099$ (−0.720; 0.522) BL versus 12 mth: ns, $g = -0.202$ (−0.828; 0.424)	BL versus 18 mth: $p = 0.082$, $g = -0.565$ (−1.233; 0.102) BL versus 24 mth: ns, $g = -0.036$ (−0.656; 0.584)
	Color-word	BL versus 03 mth: ns, $g = 0.275$ (−0.356; −0.906) BL versus 06 mth: ns, $g = 0.470$ (−0.183; 1.123) BL versus 12 mth: ns, $g = 0.179$ (−0.446; 0.803)	BL versus 18 mth: ns, $g = -0.208$ (−0.835; 0.418) BL versus 24 mth: ns, $g = 0.037$ (−0.583; 0.657)
	TMT—A	BL versus 03 mth: ns, $g = -0.231$ (−0.859; 0.397) BL versus 06 mth: ns, $g = -0.224$ (−0.851; 0.404) BL versus 12 mth: $p = 0.114$, $g = -0.506$ (1.164; 0.153)	BL versus 18 mth: ns, $g = 0.182$ (−0.443; 0.807) BL versus 24 mth: ns, $g = 0.074$ (−0.546; 0.695)
	BVRT (errors)	BL versus 03 mth: ns, $g = -0.220$ (−0.848; 0.407) BL versus 06 mth: ns, $g = -0.087$ (−0.708; 0.534) BL versus 12 mth: ns, $g = -0.031$ (−0.651; 0.589)	BL versus 18 mth: ns, $g = -0.122$ (−0.744; 0.500) BL versus 24 mth: ns, $g = -0.105$ (−0.727; 0.516)
	RAVLT	–	–
	Trials 1-5	BL versus 03 mth: ns, $g = 0.249$ (−0.380; 0.879) BL versus 06 mth: $p = 0.068$, $g = 0.599$ (−0.074; 1.272) BL versus 12 mth: ↑, $p = 0.016$, $g = 0.859$ (0.134; 1.584)	BL versus 18 mth: ns, $g = 0.488$ (−0.168; 1.144) BL versus 24 mth: ↑, $p = 0.020$, $g = 0.819$ (0.103; 1.535)
	Trial 6	BL versus 03 mth: ns, $g = 0.340$ (−0.297; 0.978) BL versus 06 mth: ns, $g = 0.388$ (−0.255; 1.030) BL versus 12 mth: $p = 0.060$, $g = 0.621$ (−0.056; 1.298)	BL versus 18 mth: ns, $g = 0.270$ (−0.361; 0.901) BL versus 24 mth: ns, $g = 0.269$ (−0.362; 0.900)
	ROCFT	–	–
	Copy	BL versus 03 mth: ns, $g = 0.298$ (−0.336; 0.931) BL versus 06 mth: ns, $g = 0.431$ (−0.217; 1.079) BL versus 12 mth: ns, $g = 0.251$ (−0.378; 0.881	BL versus 18 mth: ns, $g = -0.092$ (−0.713; 0.529) BL versus 24 mth: ns, $g = 0.019$ (−0.601; 0.639)
	Delayed recall	BL versus 03 mth: $p = 0.053$, $g = 0.645$ (−0.037; 1.326) BL versus 06 mth: ↑, $p = 0.012$, $g = 0.906$ (0.170; 1.642) BL versus 12 mth: ↑, $p = 0.006$, $g = 1.029$ (0.262; 1.795)	BL versus 18 mth: $p = 0.056$, $g = 0.633$ (−0.046; 1.312) BL versus 24 mth: ↑, $p = 0.016$, $g = 0.852$ (0.129; 1.576)

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4A. *Continued*

Ref.	Measures of efficacy	Results	
	FTT	–	–
	Right hand	BL versus 03 mth: ns, $g = -0.272$ (−0.903; 0.359) BL versus 06 mth: ns, $g = -0.086$ (−0.707; 0.535) BL versus 12 mth: ns, $g = 0.116$ (−0.506; 0.737)	BL versus 18 mth: ns, $g = -0.008$ (−0.627; 0.612) BL versus 24 mth: ns, $g = 0.137$ (−0.486; 0.760)
	Left hand	BL versus 03 mth: ns, $g = 0.165$ (−0.459; 0.7889) BL versus 06 mth: $p = 0.116$, $g = 0.503$ (−0.155; 1.160) BL versus 12 mth: $p = 0.092$, $g = 0.545$ (−0.119; 1.209)	BL versus 18 mth: ↑, $p = 0.024$, $g = 0.783$ (0.075; 1.492) BL versus 24 mth: ↑, $p = 0.009$, $g = 0.961$ (0.211; 1.710)
	Arithmetic (WAIS)	BL versus 03 mth: ns, $g = 0.457$ (−0.195; 1.108) BL versus 06 mth: ns, $g = 0.477$ (−0.177; 1.131) BL versus 12 mth: ns, $g = 0.242$ (−0.387; 0.870)	BL versus 18 mth: ns, $g = 0.478$ (−0.176; 1.132) BL versus 24 mth: $p = 0.070$, $g = 0.595$ (−0.078; 1.267)
	Luria tasks	–	–
	Rythm reproduction	BL versus 03 mth: ns, $g = -0.467$ (−1.120; 0.186) BL versus 06 mth: ns, $g = -0.413$ (−1.058; 0.233) BL versus 12 mth: $p = 0.118$, $g = -0.500$ (−1.158; 0.157)	BL versus 18 mth: ns, $g = -0.222$ (−0.850; 0.405) BL versus 24 mth: ns, $g = -0.147$ (−0.771; 0.476)
	Motor test	BL versus 03 mth: ↑, $p = 0.005$, $g = 1.071$ (0.294; 1.849) BL versus 06 mth: ↑, $p < 0.001$, $g = 1.413$ (0.537; 2.290) BL versus 12 mth: ↑, $p = 0.002$, $g = 1.266$ (0.434; 2.098)	BL versus 18 mth: ↑, $p < 0.001$, $g = 1.636$ (0.688; 2.583) BL versus 24 mth: ↑, $p = 0.002$, $g = 1.243$ (0.417; 2.068)
	Verbal fluency	–	–
	Letter	BL versus 03 mth: ns, $g = 0.203$ (−0.424; 0.829) BL versus 06 mth: $p = 0.092$, $g = 0.545$ (−0.120; 1.209) BL versus 12 mth: ns, $g = 0.401$ (−0.243; 1.045)	BL versus 18 mth: ns, $g = 0.430$ (−0.218; 1.078) BL versus 24 mth: ns, $g = 0.110$ (−0.512; 0.731)
	Category	BL versus 03 mth: ns, $g = 0.121$ (−0.501; 0.743) BL versus 06 mth: $p = 0.098$, $g = 0.534$ (−0.129; 1.196) BL versus 12 mth: $p = 0.089$, $g = 0.551$ (−0.114; 1.216)	BL versus 18 mth: ns, $g = 0.201$ (−0.425; 0.827) BL versus 24 mth: ns, $g = -0.050$ (−0.671; 0.570)
[40]	Corsi test	–	
	Forward	BL versus 02 mth: ns, $g = -0.082$ (−0.521; 0.357)	
	Backward	BL versus 02 mth: ns, $g = 0.007$ (−0.432; 0.445)	
	Digit span	–	
	Forward	BL versus 02 mth: ns, $g = 0.080$ (−0.359; 0.519)	
	Backward	BL versus 02 mth: ns, $g = 0.296$ (−0.152; 0.743)	
	Spinnler matrices	BL versus 02 mth: ns, $g = 0.098$ (−0.342; 0.537)	
	Stroop	–	
	Word	BL versus 02 mth: ns, $g = 0.021$ (−0.418; 0.459)	
	Color	BL versus 02 mth: ns, $g = 0.138$ (−0.302; 0.578)	
	Color-word	BL versus 02 mth: ↓, $p < 0.001$, $g = 1.736$ (1.042; 2.430)	
	TMT	–	
	Condition A	BL versus 02 mth: ns, $g = 0.316$ (−0.133; 0.765)	
	Condition B	BL versus 02 mth: ns, $g = 0.258$ (−0.187; 0.704)	
	B-A (interference)	BL versus 02 mth: ns, $g = 0.122$ (−0.318; 0.562)	
	CVLT	–	
	Trials 1-5	BL versus 02 mth: ns, $g = 0.057$ (−0.382; 0.495)	
	LD-FR	BL versus 02 mth: ns, $g = 0.029$ (−0.409; 0.468)	
	mWCST	–	
	Categories	BL versus 02 mth: ns, $g = 0.339$ (−0.111; 0.790)	
	Perseverative errors	BL versus 02 mth: ns, $g = -0.107$ (−0.546; 0.333)	
	Verbal fluency	–	
	Letter	BL versus 02 mth: ns, $g = 0.085$ (−0.354; 0.524)	
	Category	BL versus 02 mth: ↑, $p = 0.003$, $g = 0.738$ (0.244; 1.232)	
[41]	Digit span	–	
	Forward	Gr. 1: BL versus 03 mth: ns, $g = 0.096$ (−0.344; 0.535) Gr. 2: BL versus 03 mth: ns, $g = 0.116$ (−0.324; 0.556)	
	Backward	Gr. 1: BL versus 03 mth: ns, $g = 0.355$ (−0.097; 0.807) Gr. 2: BL versus 03 mth: ns, $g = 0.162$ (−0.279; 0.604)	

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4A. *Continued*

Ref.	Measures of efficacy	Results
	Stroop	–
	Word	Gr. 1: BL versus 03 mth: ns, $g = 0.419$ (−0.038; 0.876) Gr. 2: BL versus 03 mth: ns, $g = 0.179$ (−0.263; 0.621)
	Color	Gr. 1: BL versus 03 mth: ns, $g = 0.266$ (−0.180; 0.712) Gr. 2: BL versus 03 mth: ns, $g = 0.148$ (−0.292; 0.589)
	Color-word	Gr. 1: BL versus 03 mth: ns, $g = 0.310$ (−0.139; 0.759) Gr. 2: BL versus 03 mth: ↓, $p = 0.002$, $g = 0.779$ (0.279; 1.280)
	TMT	–
	Condition A	Gr. 1: BL versus 03 mth: ns, $g = -0.089$ (−0.528; 0.350) Gr. 2: BL versus 03 mth: ns, $g = -0.062$ (−0.501; 0.377)
	Condition B	Gr. 1: BL versus 03 mth: ↑, $p < 0.001$, $g = 0.898$ (−1.417; −0.379) Gr. 2: BL versus 03 mth: ns, $g = -0.064$ (−0.502; 0.375)
	B-A (interference)	Gr. 1: BL versus 03 mth: ↑, $p = 0.008$, $g = -0.641$ (−1.122; −0.160) Gr. 2: BL versus 03 mth: ns, $g = -0.108$ (−0.547; 0.332)
	RAVLT	–
	Trials 1-5	Gr. 1: BL versus 03 mth: ↑, $p = 0.016$, $g = 0.568$ (0.095; 1.040) Gr. 2: BL versus 03 mth: ns, $g = 0.145$ (−0.296; 0.586)
	LD-FR	Gr. 1: BL versus 03 mth: ns, $g = 0.375$ (−0.079; 0.828) Gr. 2: BL versus 03 mth: ns, $g = 0.087$ (−0.352; 0.526)
	RPM	Gr. 1: BL versus 03 mth: ns, $g = 0.464$ (0.003; 0.925) Gr. 2: BL versus 03 mth: ns, $g = 0.440$ (−0.019; 0.899)
	Tower of London (Krikorian)	Gr. 1: BL versus 03 mth: ↑, $p = 0.024$, $g = 0.528$ (0.060; 0.995) Gr. 2: BL versus 03 mth: ns, $g = 0.019$ (−0.419; 0.457)
	Verbal fluency	–
	Letter	Gr. 1: BL versus 03 mth: ↑, $p = 0.046$, $g = 0.457$ (−0.003; 0.918) Gr. 2: BL versus 03 mth: ns, $g = 0.261$ (−0.185; 0.706)
	Category	Gr. 1: BL versus 03 mth: ns, $g = 0.074$ (−0.365; 0.512) Gr. 2: BL versus 03 mth: ns, $g = 0.151$ (−0.290; 0.591)
[42]	Modified Stroop task	–
	Accuracy (%)	–
	Sensorimotor condition	"*Off*" versus "*On*": ns, $g =$ nil; "*Off*" or "*On*" versus HCS: ↑, $p < 0.0001$, $g =$ nil
	Neutral condition	"*Off*" versus "*On*": ns, $g =$ nil; "*Off*" or "*On*" versus HCS: ↑, $p < 0.0001$, $g =$ nil
	Incongruent condition	"*Off*" versus "*On*": ↑, $p < 0.03$, $g =$ nil; "*Off*" or "*On*" versus HCS: ↑, $p < 0.0001$, $g =$ nil
	Interference effect (RT, milliseconds): "Incongruent"—"Neutral"	"*Off*" versus "*On*": ns, $g = 0.127$ (−0.441; 0.695)
	ΔRT (milliseconds)	–
	"Neutral"—"Sensorimotor"	"*Off*" versus "*On*": ns, $g = -0.498$ (−1.098; 0.102); HCS versus "*Off*": $p = 0.092$, $g = 0.728$ (−0.138; 1.595); HCS versus "*On*": ns, $g = 0.263$ (−0.580; 1.106)
	"Incongruent"—"Sensorimotor"	"*Off*" versus "*On*": ns, $g = -0.386$ (−0.973; 0.200); HCS versus "*Off*": $p = 0.174$, $g = 0.581$ (−0.275; 1.438); HCS versus "*On*": ns, $g = 0.080$ (−0.760; 0.919)
[43]	Corsi test	–
	Forward	BL versus 02 mth: ns, $g = -0.167$ (−0.609; 0.274)
	Backward	BL versus 02 mth: ns, $g = 0.000$ (−0.438; 0.438)
	Digit span	–
	Forward	BL versus 02 mth: ns, $g = 0.294$ (−0.154; 0.742)
	Backward	BL versus 02 mth: ↑, $p = 0.027$, $g = 0.515$ (0.048; 0.981)
	Spinnler matrices	BL versus 02 mth: ns, $g = 0.376$ (−0.078; 0.829)
	Stroop	–
	Word	BL versus 02 mth: ns, $g = 0.312$ (−0.128; 0.771)
	Color	BL versus 02 mth: ns, $g = 0.335$ (−0.116; 0.785)
	Color-word	BL versus 02 mth: ↓, $p < 0.001$, $g = 1.042$ (0.498; 1.587)
	TMT	–
	Condition A	BL versus 02 mth: ns, $g = -0.075$ (−0.514; 0.364)
	Condition B	BL versus 02 mth: ns, $g = -0.161$ (−0.602; 0.280)
	B-A (interference)	BL versus 02 mth: ns, $g = -0.254$ (−0.699; 0.192)

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4A. *Continued*

Ref.	Measures of efficacy	Results
	CVLT	–
	Trials 1-5	BL versus 02 mth: ns, $g = 0.304$ (−0.144; 0.753)
	Long delay free recall	BL versus 02 mth: ns, $g = 0.105$ (−0.334; 0.545)
	mWCST	–
	Categories	BL versus 02 mth: ns, $g = 0.308$ (−0.141; 0.757)
	Perseverative errors	BL versus 02 mth: ns, $g = 0.045$ (−0.394; 0.483)
	Verbal fluency	–
	Letter	BL versus 02 mth: ↑, $p = 0.031$, $g = 0.501$ (0.036; 0.966)
	Category	BL versus 02 mth: ↑, $p = 0.046$, $g = 0.458$ (−0.003; 0.919)
[44]	MMSE	"*Off*" versus "*On*": ns, $g = 0.081$ (−0.970; 0.032); HCS versus "*Off*": ns, $g = -0.196$ (−0.837; 0.445); HCS versus "*On*": ns, $g = -0.098$ (−0.738; 0.542)
	MoCA	"*Off*" versus "*On*": ns, $g = 0.000$ (−0.475; 0.475); HCS versus "*Off*": ns, $g = 0.038$ (−0.602; 0.677); HCS versus "*On*": ns, $g = 0.033$ (−0.607; 0.672)
	Explicit motor sequence learning task	–
	Learning magnitude (RT; ms)	–
	Effect of learning phase	"*Off*" versus "*On*": ↑, $p < 0.0001$, $g =$ nil; HCS versus "*Off*": ↑, $p < 0.0001$, $g =$ nil; HCS versus "*On*": ↑, $p < 0.0001$, $g =$ nil
	Effect of Rx status	"*Off*" versus "*On*": ns, $g =$ nil
	Learning effect (no. of error)	–
	Rx status × learning phase	"*Off*" versus "*On*": ↑, $p < 0.05$, $g =$ nil
	Early phase	"*Off*" versus "*On*": ↑, $p < 0.05$, $g =$ nil
	Middle phase	"*Off*" versus "*On*": ns, $g =$ nil
	Late phase	"*Off*" versus "*On*": ns, $g =$ nil
	Learning phase effect	"*Off*" state: ↑, $p < 0.05$, $g =$ nil; "*On*" state: ns, $g =$ nil; HCS: ns, $g =$ nil
	Grooved Pegboard test (most affected hand)	"*Off*" versus "*On*": ns, $g = -0.469$ (−0.970; 0.032)

↑ = Statistically significant improvement post-treatment/better performance in the experimental group; ↓ = statistically significant deterioration post-treatment/worse performance in experimental group; ΔRT = differences (Δ) in RT associated with the task conditions (incongruent and neutral) after subtraction of the sensorimotor RT; BL = baseline; BSRT = Buschke Selective Reminding Test; BVRT = Benton's Visual Retention Test; CVLT = California Verbal Learning Test; FTT = Finger Tapping Test; g = Hedges' g effect size; Gr. = group; HCS = healthy control subjects; LD-FR = long delay-free recall; MMSE = Mini-Mental State Examination; mth = months; MoCA = Montreal Cognitive Assessment; nil = information not available; ns = non-significant; OMO = Odd Man Out; RAVLT = Rey Auditory Verbal Learning Test; Ref. = reference; ROCFT = Rey-Osterrieth Complex Figure Test; RPM = Raven's Progressive Matrices; RT = reaction times; Rx = medication administered to the experimental group; SDMT = Symbol Digit Modalities Test; SRT = Simple Reaction Time; TMT = Trail Making Test; WAIS = Wechsler Adult Intelligence Scale; mWCST = modified Wisconsin Card Sorting Test.

size ($p = 0.024$; $g = 0.528$) after 3 months of treatment.[41] Other beneficial effects were reported on the TMT-B (alternance and attention/processing speed; large effect; $p < 0.001$; $g = 0.898$) and TMT B-A (alternance only; medium effect; $p = 0.008$; $g = 0.641$).[41]

Acute Versus Chronic Effects of Treatment

As the study duration for the seven L-D articles varied greatly, we compared the results of the two studies[42,44] assessing the acute effects (On/Off paradigm) and the five studies[37,39-41,43] reporting the chronic effects (longer study duration) of L-D on cognition. Significant acute effects were only reported in a test of cognitive inhibition[42] and a test of motor sequence learning,[44] whereas chronic effects were obtained on multiple tests across five cognitive domains. However, the studies on acute effects were of poor quality and only reported the results of four tests divided across four cognitive domains. Furthermore, only one test (modified Stroop task) is comparable to one of the different tests used in the studies on chronic effects, and the result is incongruent with those. This might indicate that acute and chronic treatments have opposing effects on cognitive inhibition (acute being beneficial and chronic being deleterious), but it is currently impossible to conclude anything on that matter considering the scarce data on acute effects of L-D in the context of this review.

Pramipexole

Table 2B presents the characteristics of four articles with very low to low MQ reporting the effects of 1-3.9 mg/day of PRX on cognition. The main limitations of these studies were the same as the L-D studies with low or very low MQ. Two trials were randomized controlled,[45,46] one used a randomized cross-over design[40] and the last one used a pretest-posttest within-group design.[47] Study duration ranged from <1 week to 6 months.

Sample sizes varied from 7 to 30 patients. Mean age and disease duration ranged, respectively, from 57.0 to 63.0 years old and from 2.5 to 6.8 years. A study exclusively enrolled *de novo*

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

patients.[45] The MMSE/MoCA mean scores (from 28.1 to 28.8) suggest that the majority of PD patients had either slightly altered or intact cognition before treatment. Table 4B shows the results of the PRX studies.

Two of the four PRX studies presented some statistically significant results on their cognitive outcome measures. On a total of 35 results, 2 (5.7%) indicated a statistically significant effect, both of those effects being deleterious. With these few significant results, it was not possible to appreciate the possible convergence of the effects.

Attention/Processing Speed/Working Memory

No significant effects of PRX were found on tests of visual and verbal attention/working memory (Spinnler Matrices,[40,45] Stroop test,[40,46] TMT A-B,[40,46] SRT,[45] Corsi test,[40] Digit Span[40]) after <1 to 24 weeks of treatment.

Episodic Memory

Only one study ($n = 20$)[40] reported data on the effects of PRX on memory, using the CVLT. A detrimental effect of small to medium effect size ($p = 0.037$; $g = 0.481$) was reported after 2 months of treatment with 3.9 mg/day of PRX on the CVLT-trials 1-5 score (verbal encoding and learning). The difference in performance for the CVLT-delayed recall (retrieval and retention over time) did not reach statistical significance ($p = 0.052$; $g = 0.445$).

Concept Formation and Reasoning

There was no significant difference in performance on the mWCST after 2 months of treatment.[40]

Executive Functions

The performance of the seven participants involved in an On/Off-treatment paradigm deteriorated on the large reward choices condition of a delay discounting task ($p = 0.003$), but did not change on a go/no-go task (both tests assess impulsivity/inhibition).[47] No significant effect of PRX was found on verbal fluency tasks ($n = 20$[40]; 30[46]) after 2 and 6 months of treatment.

Acute Versus Chronic Effects of Treatment

There was one study on acute effects and three studies on chronic effects for PRX. Out of the two tests administered in the acute effects study, both measuring inhibition/impulsivity, one yielded a significant deleterious effect. However, as no studies on chronic effects administered tests of impulse control or cognitive inhibition, it is impossible to compare the impact of acute and chronic administration of PRX.

Selegiline

Table 2C presents the characteristics of three studies with very low to high MQ that evaluated the effects of 10 mg/day of SEL on cognition. The main limitations of the studies with low-very low MQ were the same as the L-D studies in the same categories. As for the high-MQ study, it only lost a few points for the lack of description of side effects and for not detailing the reasons for patient withdrawals. One study was a randomized, double-blind, placebo-controlled trial,[48] another was a randomized, double-blind, cross-over, placebo-controlled trial[49] and the last one was a pretest-posttest controlled trial.[50] Study durations ranged from 3 to 24 months.

Group sample sizes varied from 9 to 187 patients. Mean age and disease duration ranged from, respectively, 55.4 to 66.9 years old and from 1.3 to 3.9 years. Only one study enrolled *de novo* patients and provided MMSE data before treatment (mean score = 28.8),[48] indicating mildly impaired or intact cognitive functioning. Table 4C shows the results of the SEL studies.

Two of the three SEL studies presented some statistically significant results on their cognitive outcome measures. On a total of 124 results, 11 (8.9%) indicated a statistically significant effect, 36.4% of these effects being deleterious and 63.6% being beneficial. Most significant effects were found on measures of concept formation and reasoning (45.5%). In all, 55% of the significant effects were reported by the high-MQ study, whereas the remaining 45% were reported by the low-MQ study. For the cognitive domains that were assessed by more than one study, results tended to be divergent between studies.

Global Cognition

Performance in the SEL group was maintained, whereas the Pb group's performance improved, leading to a small deleterious effect of SEL when comparing the two groups ($p = 0.010$; $g = 0.274$) in the annual rate of change (ARoC) on the MMSE ($n = 187$).[48] Regarding the WAIS-IQ score, no significant difference was reported between the SEL and the untreated PD groups after 3 months of treatment ($n = 17$).[50]

Attention/Processing Speed/Working Memory

There was no significant change after SEL treatment on the Digit Span, the Spot-the-Dot task and the SDMT.[48]

Episodic Memory

A robust study[48] obtained a detrimental effect of SEL treatment on the ARoC on the BSRT-delayed recall, but with a small effect size ($p = 0.018$; $g = 0.173$). However, this effect was no longer significant when compared with Pb. In another study,[50] the performance of the SEL group compared with the untreated group on the PGIMS total score was in the direction of an amelioration, but did not reach statistical significance ($p = 0.052$; $g = 0.699$). Finally, a third study[49] reported no significant difference on the Rivermead Behavioral Memory Test (RBMT) after 2 months of treatment, except for a large beneficial effect on the screening score ($p = 0.044$; $g = 0.717$) for one of the two groups ($n = 17$ and 11) when compared with baseline, which was no longer significant when compared with Pb.

Construction and Motor Performance

No significant change on the Rod Orientation Test (fine manual dexterity) was observed after 2 months of SEL treatment.[49]

Concept Formation and Reasoning

The performance of patients deteriorated when they switched from a 2-month Pb trial to a 2-month SEL trial on the WCST, with an increase of the perseverative responses ($p = 0.012$; $g = 1.129$) and errors ($p = 0.035$; $g = 0.926$).[49] On a similar test, the OMO, there was no difference compared with Pb. The OMO is generally

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4B: Effects of pramipexole on cognitive outcome measures

Ref.	Measures of efficacy	Results
[40]	Corsi test	–
	Forward	BL versus 02 mth: ns, g = −0.226 (−0.670; 0.218)
	Backward	BL versus 02 mth: ns, *g* = −0.205 (−0.648; 0.238)
	Digit span	–
	Forward	BL versus 02 mth: ns, *g* = 0.140 (−0.301; 0.580)
	Backward	BL versus 02 mth: ns, *g* = 0.247 (−0.198; 0.692)
	Spinnler matrices	BL versus 02 mth: ns, *g* = −0.419 (−0.876; 0.038)
	Stroop	–
	Word	BL versus 02 mth: ns, *g* = −0.159 (−0.600; 0.282)
	Color	BL versus 02 mth: ns, *g* = −0.235 (−0.679; 0.210)
	Color-word	BL versus 02 mth: ns, *g* = −0.087 (−0.526; 0.352)
	TMT	–
	Condition A	BL versus 02 mth: ns, *g* = 0.448 (−0.012; 0.907)
	Condition B	BL versus 02 mth: ns, *g* = 0.421 (−0.037; 0.878)
	B-A (interference)	BL versus 02 mth: ns, *g* = 0.329 (−0.121; 0.779)
	CVLT	–
	Trials 1-5	BL versus 02 mth: ↓, *p* = 0.037, *g* = −0.481 (−0.944; −0.018)
	LD-FR	BL versus 02 mth: ns, *g* = −0.445 (−0.904; 0.015)
	mWCST	–
	Categories	BL versus 02 mth: ns, *g* = 0.200 (−0.242; 0.643)
	Perseverative errors	BL versus 02 mth: ns, *g* = −0.103 (−0.542; 0.337)
	Verbal fluency	–
	Letter	BL versus 02 mth: ns, *g* = −0.169 (−0.611; 0.272)
	Category	BL versus 02 mth: ns, *g* = −0.162 (−0.603; 0.279)
[45]	Spinnler matrices	"*Off*" versus "*On*": ns, *g* = 0.247 (−0.383; 0.876)
	SRT (RT, ms)	"*Off*" versus "*On*": ns, *g* = 0.203 (−0.423; 0.829)
[46]	Stroop (Color-word)	Gr. 1: BL versus 06 mth: ns, *g* = 0.176 (−0.184; 0.537) Gr. 2: BL versus 06 mth: ns, *g* = 0.303 (−0.098; 0.704) Gr. 2 versus Gr. 1: BL versus 06 mth: ns, *g* = −0.104 (−0.636; 0.427)
	TMT	–
	Condition A	Gr. 1: BL versus 06 mth: ns, *g* = 0.080 (−0.278; 0.438) Gr. 2: BL versus 06 mth: ns, *g* = 0.071 (−0.322; 0.464) Gr. 2 versus Gr. 1: BL versus 06 mth: ns, *g* = 0.010 (−0.521; 0.541)
	Condition B	Gr. 1: BL versus 06 mth: ns, *g* = 0.105 (−0.253; 0.464) Gr. 2: BL versus 06 mth: ns, *g* = 0.033 (−0.359; 0.425) Gr. 2 versus Gr. 1: BL versus 06 mth: ns, *g* = 0.083 (−0.448; 0.614)
	Verbal fluency (letter)	Gr. 1: BL versus 06 mth: ns, *g* = −0.102 (−0.461; 0.257) Gr. 2: BL versus 06 mth: ns, *g* = −0.063 (−0.455; 0.330) Gr. 2 versus Gr. 1: BL versus 06 mth: ns, *g* = −0.056 (−0.587; 0.475)
[47]	Delay discounting task	–
	Small-reward choices	"*Off*" versus "*On*": ns, *g* = nil
	Large-reward choices	"*Off*" versus "*On*": ↓, *p* = 0.003, *g* = nil
	Go/no-go task	–
	RT (ms)	"*Off*" versus "*On*": ns, *g* = −0.275 (−1.030; 0.480)
	Errors	"*Off*" versus "*On*": ns, *g* = nil

↓ = Statistically significant deterioration post-treatment/worse performance in experimental group; BL = baseline; CVLT = California Verbal Learning Test; *g* = Hedges' *g* effect size; Gr. = group; LD-FR = long delay-free recall; mth = months; mWCST = modified Wisconsin Card Sorting Test; ns = non-significant; Ref. = reference; RT = reaction times; SRT = simple reaction time; TMT = Trail Making Test.

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4C: Effects of selegiline on cognitive outcome measures

Ref.	Measures of efficacy	Results
[48]	MMSE	Gr. 3: ARoC: ns, $g = 0.000$ (−0.143; 0.143) Gr. 1: ARoC: ↑, $p < 0.001$, $g = 0.273$ (0.122; 0.424) Gr. 1 versus Gr. 3: ARoC: ↓, $p = 0.010$, $g = -0.274$ (−0.481; −0.066)
	Digit span	–
	Forward	Gr. 3: ARoC: ns, $g = 0.000$ (−0.143; 0.143) Gr. 1: ARoC: ns, $g = -0.090$ (−0.239; 0.059) Gr. 1 versus Gr. 3: ARoC: ns, $g = 0.095$ (−0.112; 0.302)
	Backward	Gr. 3: ARoC: ns, $g = 0.059$ (−0.085; 0.202) Gr. 1: ARoC: ns, $g = -0.048$ (−0.196; 0.101) Gr. 1 versus Gr. 3: ARoC: ns, $g = 0.105$ (−0.102; 0.312)
	Spot-the-dot task	Gr. 3: ARoC: ns, $g = -0.143$ (−0.287; 0.001) Gr. 1: ARoC: ns, $g = -0.117$ (−0.266; 0.032) Gr. 1 versus Gr. 3: ARoC: ns, $g = 0.000$ (−0.206; 0.206)
	SDMT	Gr. 3: ARoC: ns, $g = 0.040$ (−0.103; 0.184) Gr. 1: ARoC: ns, $g = -0.112$ (−0.261; 0.037) Gr. 1 versus Gr. 3: ARoC: ns, $g = 0.157$ (−0.050; 0.363)
	BSRT	
	Total recall	Gr. 3: ARoC: ns, $g = 0.010$ (−0.044; 0.243) Gr. 1: ARoC: ns, $g = 0.070$ (−0.079; 0.129) Gr. 1 versus Gr. 3: ARoC: ns, $g = 0.000$ (−0.206; 0.206)
	Long-term recall	Gr. 3: ARoC: ns, $g = 0.038$ (−0.105; 0.181) Gr. 1: ARoC: ns, $g = 0.120$ (−0.029; 0.270) Gr. 1 versus Gr. 3: ARoC: ns, $g = -0.096$ (−0.302; 0.111)
	Delayed recall	Gr. 3: ARoC: ↓, $p = 0.018$, $g = -0.173$ (−0.318; −0.029) Gr. 1: ARoC: ns, $g = -0.077$ (−0.225; 0.072) Gr. 1 versus Gr. 3: ARoC: ns, $g = -0.082$ (−0.288; 0.125)
	OMO	
	Trials 1 + 3	Gr. 3: ARoC: ns, $g = 0.093$ (−0.050; 0.237) Gr. 1: ARoC: ns, $g = 0.000$ (−0.149; 0.149) Gr. 1 versus Gr. 3: ARoC: ns, $g = 0.102$ (−0.104; 0.309)
	Trials 2 + 4	Gr. 3: ARoC: ns, $g = 0.142$ (−0.002; 0.286) Gr. 1: ARoC: ↑, $p < 0.001$, $g = 0.280$ (0.128; 0.431) Gr. 1 versus Gr. 3: ARoC: ns, $g = -0.167$ (−0.373; 0.041)
	COWAT	Gr. 3: ARoC: ↑, $p < 0.001$, $g = 0.588$ (0.433; 0.743) Gr. 1: ARoC: ↑, $p < 0.001$, $g = 0.629$ (0.466; 0.792) Gr. 1 versus Gr. 3: ARoC: ns, $g = -0.051$ (−0.257; 0.156)
[49]	RBMT	–
	Profile score	Gr. 2: BL versus 02 mth: $p = 0.056$, $g = 0.673$ (−0.050; 1.397) Gr. 1: BL versus 02 mth: ns, $g = 0.088$ (−0.504; 0.680) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.346$ (−0.262; 0.955) Gr. 1 versus Gr. 2: BL versus 02 mth: $p = 0.184$, $g = 0.595$ (−0.305; 1.496) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.273$ (−0.566; 1.113)
	Screening score	Gr. 2: BL versus 02 mth: ↑, $p = 0.044$, $g = 0.717$ (−0.016; 1.450) Gr. 1: BL versus 02 mth: ns, $g = 0.179$ (−0.417; 0.774) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.128$ (−0.465; 0.722) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.416$ (−0.4474; 1.307) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = -0.053$ (−0.889; 0.783)
	Prose-immediate	Gr. 2: BL versus 02 mth: ns, $g = 0.426$ (−0.257; 1.108) Gr. 1: BL versus 02 mth: ns, $g = 0.298$ (−0.306; 0.902) Gr. 1: 02 mth versus 04 mth: ns, $g = -0.086$ (−0.678; 0.506) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.241$ (−0.643; 1.125) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = -0.384$ (−1.227; 0.460)
	Prose-delayed	Gr. 2: BL versus 02 mth: ns, $g = 0.452$ (−0.234; 1.138) Gr. 1: BL versus 02 mth: $p = 0.072$, $g = 0.560$ (−0.076; 1.195) Gr. 1: 02 mth versus 04 mth: ns, $g = -0.198$ (−0.795; 0.399) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.211$ (−1.094; 0.673) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): $p = 0.065$, $g = -0.801$ (−1.700; 0.0675)

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4C. *Continued*

Ref.	Measures of efficacy	Results
	ROT	–
	Total errors	Gr. 2: BL versus 02 mth: ns, $g = -0.482$ (−1.172; 0.209) Gr. 1: BL versus 02 mth: ns, $g = -0.171$ (−0.767; 0.424) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.096$ (−0.497; 0.688) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.108$ (−0.990; 0.774) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.282$ (−0.558; 1.122)
	Visual vertical errors	Gr. 2: BL versus 02 mth: ns, $g = 0.183$ (−0.476; 0.842) Gr. 1: BL versus 02 mth: ns, $g = 0.482$ (−0.143; 1.106) Gr. 1: 02 mth versus 04 mth: ns, $g = -0.111$ (−0.704; 0.482) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.143$ (−1.025; 0.739) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): $p = 0.227$, $g = -0.512$ (−1.361; 0.338)
	Visual horizontal errors	Gr. 2: BL versus 02 mth: ns, $g = -0.255$ (−0.919; 0.409) Gr. 1: BL versus 02 mth: ns, $g = -0.274$ (−0.876; 0.328) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.178$ (−0.418; 0.774) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.222$ (−0.661; 1.106) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.476$ (−0.372; 1.323)
	Tactile vertical errors	Gr. 2: BL versus 02 mth: ns, $g = -0.347$ (−1.020; 0.326) Gr. 1: BL versus 02 mth: ns, $g = -0.013$ (−0.604; 0.579) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.073$ (−0.519; 0.665) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.220$ (−1.104; 0.663) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.090$ (−0.746; 0.926)
	Tactile horizontal errors	Gr. 2: BL versus 02 mth: ns, $g = -0.345$ (−1.017; 0.328) Gr. 1: BL versus 02 mth: ns, $g = -0.202$ (−0.799; 0.395) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.293$ (−0.311; 0.896) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.000$ (−0.881; 0.881) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.463$ (−0.384; 1.310)
	APM	Gr. 2: BL versus 02 mth: ns, $g = -0.261$ (−0.925; 0.403) Gr. 1: BL versus 02 mth: ns, $g = -0.199$ (−0.796; 0.397) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.000$ (−0.591; 0.591) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.166$ (−1.049; 0.716) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.193$ (−0.645; 1.030)
	WCST	–
	Total correct	Gr. 2: BL versus 02 mth: ns, $g = 0.054$ (−0.600; 0.708) Gr. 1: BL versus 02 mth: ns, $g = 0.311$ (−0.294; 0.916) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.120$ (−0.473; 0.713) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.261$ (−1.145; 0.624) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = -0.203$ (−1.041; 0.635)
	Total errors	Gr. 2: BL versus 02 mth: ns, $g = -0.211$ (−0.872; 0.450) Gr. 1: BL versus 02 mth: ns, $g = -0.204$ (−0.801; 0.393) Gr. 1: 02 mth versus 04 mth: ns, $g = -0.169$ (−0.764; 0.427) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.066$ (−0.947; 0.816) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.046$ (−0.790; 0.882)
	Non-perseverative errors	Gr. 2: BL versus 02 mth: ns, $g = -0.131$ (−0.787; 0.525) Gr. 1: BL versus 02 mth: ns, $g = 0.227$ (−0.372; 0.825) Gr. 1: 02 mth versus 04 mth: ns, $g = -0.366$ (−0.977; 0.244) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.378$ (−1.267; 0.511) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): $p = 0.153$, $g = -0.610$ (−1.465; 0.245)
	Perseverative responses	Gr. 2: BL versus 02 mth: ns, $g = -0.204$ (−0.864; 0.456) Gr. 1: BL versus 02 mth: ↑, $p = 0.026$, $g = -0.726$ (−1.391; −0.062) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.329$ (−0.278; 0.935) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.297$ (−0.588; 1.183) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ↓, $p = 0.012$, $g = 1.129$ (0.229; 2.029)
	Perseverative errors	Gr. 2: BL versus 02 mth: ns, $g = -0.196$ (−0.855; 0.464) Gr. 1: BL versus 02 mth: ↑, $p = 0.025$, $g = -0.731$ (−1.396; −0.066) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.136$ (−0.458; 0.730) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.274$ (−0.611; 1.159) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ↓, $p = 0.035$, $g = 0.926$ (0.047; 1.806)
	Categories achieved	Gr. 2: BL versus 02 mth: ns, $g = -0.031$ (−0.685; 0.622) Gr. 1: BL versus 02 mth: ns, $g = 0.438$ (−0.181; 1.056) Gr. 1: 02 mth versus 04 mth: ns, $g = 0.000$ (−0.591; 0.591) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = -0.443$ (−1.335; 0.449) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = -0.474$ (−1.321; 0.374)

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4C. *Continued*

Ref.	Measures of efficacy	Results
	Trials to first category	Gr. 2: BL versus 02 mth: ns, $g = -0.011$ (−0.664; 0.643) Gr. 1: BL versus 02 mth: ns, $g = -0.299$ (−0.904; 0.305) Gr. 1: 02 mth versus 04 mth: ns, $g = -0.107$ (−0.700; 0.485) Gr. 1 versus Gr. 2: BL versus 02 mth: ns, $g = 0.276$ (−0.609; 1.162) Gr. 1: (BL versus 02 mth) versus (02 mth versus 04 mth): ns, $g = 0.275$ (−0.565; 1.114)
[50]	WAIS	Gr. 1: BL versus 03 mth: ns, $g = 0.433$ (−0.064; 0.930) Gr. 2: BL versus 03 mth: ns, $g = -0.018$ (−0.524; 0.488) Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.481$ (−0.222; 1.186)
	PGIMS	Gr. 1: BL versus 03 mth: ns, $g = 0.488$ (−0.015; 0.991) Gr. 2: BL versus 03 mth: ns, $g = -0.205$ (−0.716; 0.307) Gr. 2 versus Gr. 1: BL versus 03 mth: $p = 0.052$, $g = 0.699$ (−0.016; 1.414)

↑ = Statistically significant improvement post-treatment/better performance in the experimental group; ↓ = statistically significant deterioration post-treatment/worse performance in experimental group; APM = Advanced Progressive Matrices; ARoC = Annual Rates of Change; BL = baseline; BSRT = Buschke Selective Reminding Test; COWAT = Controlled Oral Word Association Test; *g* = Hedges' *g* effect size; Gr. = group; MMSE = Mini-Mental State Examination; mth = months; ns = non-significant; OMO-Odd Man Out; PGIMS = PGI Memory Scale; RBMT = Rivermead Behavioral Memory Test; Ref. = Reference; ROT = Rod Orientation Test; RT = reaction times; SDMT = Symbol Digit Modalities Test; WAIS = Weschler Adult Intelligence Scale; WCST = Wisconsin Card Sorting Test.

easier to perform than the WCST, because it involves fewer concepts to discover and also less trials to complete.

Executive Functions

Both the SEL-treated ($p < 0.001$; $g = 0.588$) and Pb ($p < 0.001$; $g = 0.629$) groups improved on the ARoC of the Controlled Oral Word Association Test (COWAT) (verbal fluency). However, no significant difference was found between the ARoC of the two groups. Thus, the improvement probably indicates a practice effect, rather than a real treatment effect of SEL.

Rasagiline

Table 2D presents the characteristics of the only study[36] on RAS that met our inclusion criteria. As a study with low MQ, it has the same limitations as the L-D studies with low-very low MQ, except for the blinding of assessors and subjects. In this randomized, double-blind, placebo-controlled trial, a dose of 1 mg/day of RAS was given to 23 participants for a 3-month period. Mean age and disease duration of the RAS and Pb groups were, respectively, 65.2 and 67.6 years old and 4.1 and 4.0 years. The mean MMSE score at baseline was unavailable. However, according to the authors, patients had to be cognitively impaired, but not demented, to be enrolled in the study. The authors defined cognitive impairment as having a performance 1.5 SDs below normative scores on the screening neuropsychological tests for two out of the four cognitive domains they assessed. These criteria are similar to the PD-MCI definition published later.[1] Its main methodological limitations were same as those of the SEL studies. Table 4D shows the results on the cognitive outcomes.

Of a total of 25 cognitive results, two (8%) indicated a statistically significant effect, both of them being beneficial. With these few significant results, it was not possible to appreciate the possible convergence of the effects.

Attention/Processing Speed/Working Memory

The performance remained unchanged on the Digit Span forward (attention span) after 3 months of treatment, but improved on the Digit Span backward (verbal working memory), with a medium-size effect ($p = 0.042$; $g = 0.594$). On the DOT-A, a beneficial change of medium effect size was also reported, although it did not reach statistical significance ($p = 0.051$; $g = 0.569$). No significant difference was obtained on the TMT-A (attention/processing speed). Altogether, these results suggest a positive, but selective, effect of RAS on the Central Executive System of Baddeley's Working Memory model.[116,117]

Perception, Episodic Memory, Naming and Construction Performance

No significant effects of RAS were registered on the tests of naming, visual perception and construction, verbal and visual episodic memory after 3 months of treatment.

Executive Functions

Patients treated with RAS ameliorated performances on the verbal fluency total score with medium effect size ($p = 0.038$; $g = 0.608$). However, when the results of each fluency task (category and lexical fluency) were analyzed individually, the improvement was not significant on the lexical ($p = 0.155$; $g = 0.411$) and the category ($p = 0.124$; $g = 0.445$) fluency tasks. No change occurred on performing the Stroop test and the TMT-B.

Discussion

The present systematic review aimed at studying the impact of currently relevant antiparkinson medications on cognition in mild-to-moderate PD. In total, 14 studies involving non-demented PD patients with intact or compromised cognition were analyzed. Overall, quality of evidence was poor; 11/14 studies had low or very low MQ scores. On average, SEL studies were of higher quality (mean MQ score = 57.2%), followed by RAS (52.2%), L-D (49.7%) and PRX (40.8%). Nevertheless, all studies used well-validated cognitive tests, and most studies administered an exhaustive neuropsychological battery. However, some cognitive domains such as language, praxis performances and perception were under-represented in the reviewed studies. Therefore, we lack data to clearly assess the effects of antiparkinson medication on these cognitive domains.

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Table 4D: Effects of rasagiline on cognitive outcome measures

Ref.	Measures of efficacy	Results
[36]	Digit span	–
	Forward	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.158$ (−0.409; 0.726)
	Backward	Gr. 2 versus Gr. 1: BL versus 03 mth: ↑, $p = 0.042$, $g = 0.594$ (0.016; 1.173)
	Total	Gr. 2 versus Gr. 1: BL versus 03 mth: $p = 0.053$, $g = 0.563$ (−0.015; 1.140)
	DOT-A	Gr. 2 versus Gr. 1: BL versus 03 mth: $p = 0.051$, $g = 0.569$ (−0.008; 1.147)
	Stroop	–
	Time difference	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.239$ (−0.330; 0.807)
	Error	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.052$ (−0.515; 0.618)
	Spontaneous corrections	Gr. 2 versus Gr. 1: BL versus 03 mth: $p = 0.070$, $g = 0.527$ (−0.050; 1.103)
	TMT	–
	Condition A	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.070$ (−0.497; 0.636)
	Condition B	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.008$ (−0.574; 0.558)
	B-A (interference)	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.022$ (−0.588; 0.544)
	BFR	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.059$ (−0.625; 0.508)
	BJLO	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.368$ (−0.203; 0.939)
	Verbal paired associates (WMS-R)	–
	Immediate recall	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.268$ (−0.867; 0.301)
	Delayed free recall	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.006$ (−0.572; 0.560)
	Recognition	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.163$ (−0.730; 0.404)
	Delayed recall + recognition	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = -0.235$ (−0.804; 0.333)
	Learning score	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.015$ (−0.551; 0.582)
	Visual reproduction (WMS-R)	–
	Immediate recall	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.187$ (−0.381; 0.754)
	Delayed recall	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.309$ (−0.260; 0.879)
	Recognition	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.078$ (−0.488; 0.645)
	BNT	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.104$ (−0.463; 0.671)
	CDT	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.207$ (−0.361; 0.775)
	Verbal fluency	–
	Letter	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.411$ (−0.161; 0.984)
	Category	Gr. 2 versus Gr. 1: BL versus 03 mth: ns, $g = 0.445$ (−0.128; 1.018)
	Total	Gr. 2 versus Gr. 1: BL versus 03 mth: ↑, $p = 0.038$, $g = 0.608$ (0.029; 1.187)

↑ = Statistically significant improvement post-treatment/better performance in the experimental group; ↓ = statistically significant deterioration post-treatment/worse performance in experimental group; BFR = Benton Facial Recognition; BJLO = Benton Judgment of Line Orientation; BL = baseline; BNT = Boston Naming Test; CDT = Clock Drawing Test; DOT-A = Adaptive Digit Ordering Test; g = Hedges' g effect size; Gr. = group; mth = months; ns = non-significant; Ref. = reference; TMT = Trail Making Test; WAIS = Weschler Adult Intelligence Scale; WMS-R = Weschler Memory Scale-Revised.

Regarding L-D, medium to large deleterious effect sizes were almost exclusively obtained on a well-validated measure of inhibition capacity (Stroop). On the other hand, beneficial effects were mainly reported on processes such as memory encoding and retrieval (medium to large effect sizes), planning/organization (medium effect size), flexibility (medium to large effect sizes), verbal fluency (small to large effect sizes) and concept formation (small effect size). Results were less consistent for attention/processing speed and for working memory, sometimes with a beneficial effect of L-D (small to medium effect sizes) and sometimes with deterioration or no change.

These results are mostly in accordance with the *dopamine overdose hypothesis*,[19,118] stating that in mild PD clinically effective doses of L-D could compensate for the DA depletion in the dorso-lateral frontostriatal loop, while at the same time overdosing relatively intact circuits such as the orbitofrontal loop. As suggested in the *inverted U-shaped model*[16] of dopaminergic stimulation, insufficient and excessive levels of DA in a loop's structures could alter the cognitive processes associated with these structures. The present findings, in line with these hypotheses, suggest that L-D principally improves executive processes associated with the dorso-lateral prefrontal cortex (DLPFC),

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

while also altering the inhibition capacity associated with the orbitofrontal cortex (OFC) in mild-to-moderate PD.

Yet, as mentioned earlier, some contradictory results were also reported, with some benefits on inhibition and deterioration in information processing speed. The large deterioration on the Stroop-Word was recorded by the only study that had recruited *de novo* patients[39] and used the highest mean daily dosage of L-D (540 mg/day). The combination of milder disease and higher dosage might partly explain the negative effects, possibly via overdosing the less-damaged DA loops. The improvement of the inhibitory processes was reported by the only study[42] that used an LD + DDCI treatment, a Stroop test modified for MRI administration, and an On/Off-treatment paradigm. In addition, the participants had longer disease durations but the UPDRS scores were comparable to other studies using the Stroop.[39-41,43] Thus, these patients possibly deteriorated more slowly and/or had better motor response from antiparkinson medication compared with patients from other studies. Taken together, these methodological differences might explain the contradictory results.

Cognitive results following PRX treatment suggest that PRX negatively affects verbal learning capacities (medium effect) at high dosage (3.9 mg/day) and impulse control at low dosage (1 mg/day). On a brighter note, there were no indications that PRX induced changes in other cognitive tasks (attentional/executive).

However, many factors must be taken into consideration regarding the PRX results. First, the four PRX studies were of very low to low MQ, exposing them to bias and limiting their capacity to detect or exclude important effects. Second, the two significant results were each reported solely on one condition of one test, making it impossible to verify whether the significant effects tend to converge or diverge within each cognitive domain. Thus, there is insufficient evidence to conclude that PRX affects cognition, but it is impossible to exclude that there might be some effects either. At best, these results could indicate that episodic memory and impulse control might be more susceptible to be affected by this medication. Some support for impulse control already exists, as other studies report an increased incidence of impulse-control disorders following PRX treatment in PD.[119,120]

Regarding SEL, the three studies administered the maximum recommended dose (10 mg/day) and their results suggest that SEL negatively affects concept formation and reasoning (large effect), as well as global cognition (small effect). Some divergent results were obtained on episodic memory, although no change in performance ended up being statistically significant. These conflicting results seem to be best explained by methodological differences between studies (i.e., low vs. high MQ, presence of a Pb group or not) and from differences in group characteristics within studies (i.e., disease stage, severity of symptoms, scolarity), than by an effect of SEL treatment.

The detrimental effect on global cognition was reported on the MMSE (small effect size) by a robust longitudinal study[48] (MQ = 84.8%) when SEL was compared with Pb. Considering that the MMSE is not particularly sensitive to cognitive changes in PD,[121,122] the fact that a significant deterioration was caught, albeit small, likely suggests that SEL could have some longer-term deleterious effect on global cognition.

Data on RAS came from only one study[36] of low MQ that included exclusively cognitively impaired, but not demented, patients. Three months of RAS treatment at maximum recommended doses (1 mg/day) did not alter these patients' cognitive abilities. It might even have benefited the central executive component of working memory (medium effect) and verbal fluency (medium effect).

Nevertheless, it is impossible to conclude that, overall, RAS has the potential to benefit certain cognitive functions in mild-to-moderate PD, nor is it possible to exclude this possibility. Although the study[36] had many strengths (e.g., randomized, double-blind, placebo-controlled trial, extensive cognitive battery), its results were obtained on a small sample of cognitively altered PD patients on a short period of time, which alters the validity and generalizability of the reported effects. It might nonetheless indicate that working memory and verbal fluency might be more susceptible to be affected by RAS. However, two recent studies that did not quite meet our selection criteria seem to support the absence of cognitive effects for RAS in cognitively altered or cognitively intact PD patients. The first study[123] tested the effects of 24 weeks of RAS treatment on cognition, motricity and activities of daily living (ADL) on 170 PD-MCI patients ($n = 86$ RAS; 84 Pb) using a multicenter, randomized, double-blind, placebo-controlled design. The authors report beneficial effects on motricity and ADL, but not on cognition. Although this study[123] can neither support nor invalidate the possibility that certain cognitive functions might be more susceptible to be affected by RAS in PD-MCI, given that global cognitive measures were administered instead of a detailed neuropsychological battery, it does support that RAS does not seem to worsen cognition in this population. The second study[124] was also a randomized, double-blind, placebo-controlled trial, but did not restrict its sample to PD-MCI patients and used a neuropsychological battery covering multiple domains, including attention, executive functions, language, visuo-spatial perception and memory. The effects of 26 weeks of RAS treatment were tested on cognition and motricity using a sample of 45 non-demented patients with mild-to-moderate stage PD ($n = 23$ RAS; 22 Pb). It concludes that RAS is an effective treatment for motor symptoms, but has no effect on cognition in this population.

When compared with one another, L-D had the most beneficial and deleterious effects on cognition; PRX showed indications toward a few negative effects, but did not affect cognition overall; SEL mildly deteriorated global cognition over time and altered concept formation; and RAS principally showed no change, but might induce a few benefits. In general, beneficial effects were principally reported on DLPFC-related attentional/executive functions, and deleterious effects were mostly obtained on OFC-related inhibitory processes and impulse control. The results are thus mostly in accordance with the predictions of the DA overdose hypothesis in mild-to-moderate PD, which is especially true for L-D, as data on other drugs were scarce. These findings are also in accordance with the conclusions of a narrative review on anti-parkinsonian effects in mild PD.[22] However, the present work analyzed the effects of more treatments (MAO-B inhibitors) and presented more exhaustive and quantitative data.

A more puzzling conclusion to be made is with regard to the effects of antiparkinson medications on episodic memory. The findings in this domain differed with each treatment and ultimately did not converge toward a general tendency.

Psychometric Qualities of the Cognitive Tests

In total, 42 different cognitive tests were used through the 14 selected studies. The majority of those tests were well-known measures of their respective cognitive domain, such as the Digit

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Span, the Stroop test, the TMT, the SDMT, the CVLT, the RAVLT, the ROCFT, the RBMT, the WCST or the COWAT (see Table 3 for more details). The most frequently used tests by far were the Digit Span, the Stroop test, the TMT and the verbal fluency tests. These five tests, along with the RAVLT, concurrently yielded most of the significant results. The vast majority of the above-listed tests are cited by the MDS task force guidelines[1] as good examples of tests to administer for cognitive assessment in PD. This ensures their validity to detect difficulties with attention, executive function, episodic memory and verbal fluency for this population. In addition, they have also shown good or acceptable test-retest and/or inter-rater reliability with various populations, which is an important factor in clinical trials in which the measures are administered by different examiners repetitively over a short period of time.

However, one test raised questions among the present authors regarding its psychometric qualities. The Delay Discounting Task, used in one study,[47] is reported to have modest levels of reliability and questionable construct validity.[103] Considering we already advised caution in the interpretation of the significant result obtained on this task following PRX treatment, these psychometric concerns reinforce our statement.

Limitations

Many factors must be taken into account for the interpretation of the present data. First, samples were generally small, with most studies including around 15 to 20 treated PD patients, with the exception of the two DATATOP studies ($n > 180$ treated participants). Small sample sizes result in lower statistical power, which can in turn result in a failure to detect subtle treatment effects. Moreover, given that PD patients have a heterogeneous cognitive decline, large sample sizes are mandatory to ensure external validity in studies investigating the impact of medications on cognition.

Second, methodological designs varied widely across the studies, affecting the internal validity of the results. Only five of the 14 reviewed studies were RCTs.[37,39,46,48,49] Randomized controlled trials are the most robust research designs, yet most studies used only within-group comparisons. It is thus impossible to know whether a placebo effect or the disease progression contributed to some results. Nine studies were not blinded, making them vulnerable to patient and investigator biases. Furthermore, when using an On-/Off-treatment design, studies are more vulnerable to be contaminated by a learning effect, because of short test-retest periods. Nevertheless, the utilization of tests that show low vulnerability to learning effects at retest in the reviewed studies should have minimized these effects.

Third, only two studies[39,48] lasted more than 6 months, leaving long-term effects of treatment on cognition mostly unknown. When looking closely at the results of the longest studies, one notices that the effects of some medications vary significantly over time. Hence, it is mandatory to realize more longitudinal studies to understand the possible variations of the antiparkinson medications' effects on cognition over time.

Fourth, the present review selected studies involving patients taking the relevant drugs in monotherapy or with an adjuvant medication for which the dosage was not always reported by the authors. Thus, an interaction effect cannot be excluded for some results, making the comparison of effects between studies more complex. However, only studies with a design that isolated the effect of the selected drug were included in this review. Furthermore, the adjuvant medications administered in the studies are those usually given to PD patients in clinical settings, thus supporting external validity of the present results.

Fifth, it was difficult to control for the baseline cognitive functioning of the patients. Only studies including non-demented patients were selected, but it was not always possible to distinguish cognitively compromised from cognitively intact patients at baseline, notably because the clinical criteria for PD-MCI are relatively recent.[1] Furthermore, some patients could have had co-pathology, especially with Alzheimer's disease, which would have an impact on cognition irrespective of the PD. Without the use of biomarkers, it is difficult to exclude this possibility. Thus, comparability between studies could be hindered. Nonetheless, global scores from MMSE or MoCA, when available, indicated that patients from different studies had overall similar scores.

Sixth, there are some concerns regarding the methodological quality assessment. The scale had some limitations during its validation: it was developed for psychiatric studies, and our MQ score is not validated. The principal limitation listed by the CCDAN authors is the portion of subjectivity that affects the score, as suggested by the lower-than-expected inter-rater reliability of 0.5.[32] To minimize the impact of this limit, the two evaluators of the present review held meetings before and after they independently performed the quality assessment. The goals of these meetings were to reach consensus on the operationalization of more subjective criteria before starting the assessment of the studies (CCDAN authors' suggestions were used when available), and to solve all discrepancies following the evaluation of the studies. Whenever a consensus could not be reached, an experienced researcher in the domain (MS) was consulted. Regarding the psychiatric purposes of the CCDAN, our rationale was that the literature on cognitive effects of antiparkinson medications shared some similar limitations to psychiatric studies (e.g., designs of highly variable quality, small sample sizes). To better differentiate our studies and avoid a floor effect, we aimed to find a scale that allowed for a detailed and systematic consideration of the various aspects of trials' quality. As for our MQ score, actions have been taken to ensure transparency of the process (see Method section). We advise caution regarding over-interpretation of the scores.

Finally, the choice of the antiparkinson medications to be reviewed in the present paper, based on the CNSF guidelines[23] (Level A quality of evidence) to initiate dopaminergic treatment in PD, might be perceived by non-Canadian readers as being too specific. However, American Academy of Neurology (AAN)[125] and National Institute for Health and Care Excellence (NICE)[126] guidelines have generally the same Level A treatment recommendations to initiate PD treatment. The only differences between the AAN and NICE versus CNSF recommandations is the addition of cabergoline (AAN) and rotigotine (NICE). There was only one study[41] on the cognitive effects of cabergoline and rotigotine together with L-D, and it has been reviewed in this paper. The authors of this study report that the three drugs did not significantly modify their patients' cognitive performance compared with an Off-treatment condition.

Future Research

Even with low MQ, the data extracted from the reviewed studies provided interesting information regarding the possible

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

association between dopaminergic antiparkinson medications and cognitive changes in patients with mild-to-moderate PD without dementia. However, one should keep in mind that some studies showed no impact at all, whereas some of the observed effects were of questionable validity. Thus, to better understand this association, especially during longer periods of time, studies with significantly higher MQ are required. To achieve this, future studies will have to use the most robust research designs (double-blind randomized placebo-controlled trials), use sample sizes of 100 or more, as indicated by the FDA for phase 2 clinical trials,[127] and have a study duration of ≥6 months to clearly assess the evolution of effects, because the results of studies with longer follow-ups have shown that some effects are only detectable after 2 to 6 months of treatment, whereas other effects seem to vary over time. This is also supported by a trial[128] that compared the cognitive effects of L-D in demented and non-demented PD patients, whose results suggest that some effects differed more than others between those two milestones of PD progression.

Regarding the cognitive evaluations, the current problems reside in the tendency of using only a few tests to measure several cognitive functions and of using several tests to assess only one or two cognitive domains. For instance, domains such as attention and executive functions were almost always assessed, whereas domains such as language, perception or visuo-constructive praxis were neglected. Although attention and executive functions are the cognitive functions most likely to be affected by DA antiparkinsonians, other domains might also be affected (after all, visuo-construction praxis and syntax involves some executive functioning, and can be both impaired by alterations of basal ganglia[129]). Furthermore, some antiparkinson medications, notably monoamine oxidase inhibitors, do not exclusively affect dopaminergic neurotransmission, which could in turn affect cognitive domains underlied by different brain regions. Therefore, all cognitive domains should be assessed minimally with one test, whereas the cognitive domains hypothesized to be more responsive to treatment should be thoroughly evaluated (> one test).

As the cognitive effects of the antiparkinson medications seem to vary over time and with disease progression, it is important to better differentiate the patient's cognitive status before treatment. Interestingly, the only study that recruited exclusively cognitively compromised patients[36] reported some of the most beneficial results. Thus, differentiating PD-MCI[1] from non-MCI patients would provide better insight regarding the most favorable contexts to prescribe the different drugs for healthcare providers.

Acknowledgments

The authors thank Mr Gaétan Daigle, Statistician, Mathematics and Statistics Department at Laval University, for his help with the statistical analyses performed in this study. The authors are also thankful to Ariane Giguère-Rancourt, Éva Racine, Laila El-Amrani, Marika Plourde, Marianne Couture, Élodie Thériault and Valérie Coulombe for their assistance in the preparation of the manuscript. Part of the results were presented at the 13th International Conference on Alzheimer's and Parkinson's Diseases and Related Neurological Disorders (AD/PD™ 2017).

Funding

None.

Disclosures

M-AR has nothing to disclose. MD reports grants from Parkinson Society Canada and grants from Canadian Institutes of Health Research—Institute of Neurosciences, Mental Health and Addiction (201210), outside the submitted work. JT-C reports grants from Canadian Institutes of Health Research, outside the submitted work. ND reports grants from Quebec Parkinson Network and grants from Actelion Pharmaceuticals, outside the submitted work. MS reports grants from Canadian Institutes of Health Research-Big Data on Dementia and grants from Alzheimer Society Research Program/Pacific Alzheimer Research Foundation, outside the submitted work.

Statement of Authorship

M-AR carried out conception, organization, and execution of the research project. He was also responsible for data extraction, including article selection and analysis, and tables; statistical analysis including design and execution; and writing of the drafts. MD contributed to data extraction, including article selection and analysis. JT-C contributed to data extraction for the preparation of tables. ND contributed to the conception of the research project. MS was responsible for the supervision of the research project; resolving discrepancies in article selection; and for the study design. MD, JT-C, ND, and MS carried out critical review of the manuscript.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/cjn.2018.21

References

1. Litvan I, Goldman JG, Tröster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. Mov Disord. 2012;27:349-356.
2. Lawrence BJ, Gasson N, Loftus AM. Prevalence and subtypes of mild cognitive impairment in Parkinson's disease. Sci Rep, 2016; 6.
3. Weintraub D, Simuni T, Caspell-Garcia C, et al. Cognitive performance and neuropsychiatric symptoms in early, untreated Parkinson's disease. Mo Disord. 2015;30(7):919-27.
4. Santangelo G, Vitale C, Picillo M, et al. Mild Cognitive Impairment in newly diagnosed Parkinson's disease: a longitudinal prospective study. Parkinsonism Relat Disord. 2015;21(10): 1219-1226.
5. Wang YQ, Tang BS, Yan XX, et al. A neurophysiological profile in Parkinson's disease with mild cognitive impairment and dementia in China. J Clin Neurosci. 2015;22(6):981-5.
6. Hobson P, Meara J. Mild cognitive impairment in Parkinson's disease and its progression onto dementia: a 16-year outcome evaluation of the Denbighshire cohort. Int J Geriatr Psychiatry. 2015;30:1048-55.
7. Kalbe E, Rehberg SP, Heber I, et al. Subtypes of mild cognitive impairment in patients with Parkinson's disease: evidence from the LANDSCAPE study. J Neurol Neurosurg Psychiatry. 2016;87(10): 1099-105.
8. Pfeiffer HCV, Løkkegaard A, Zoetmulder M, Friberg L, Werdelin L. Cognitive impairment in early-stage non-demented Parkinson's disease patients. Acta Neurol Scand. 2014;129:307-18.
9. Yarnall AJ, Breen DP, Duncan GW, et al. Characterizing mild cognitive impairment in incident Parkinson disease: the ICICLE-PD Study. Neurology. 2014;82:308-16.
10. Pigott K, Rick J, Xie SX, et al. Longitudinal study of normal cognition in Parkinson disease. Neurology. 2015;85(15):1276-82.
11. Bohnen NI, Kaufer DI, Hendrickson R, et al. Cognitive correlates of cortical cholinergic denervation in Parkinson's disease and Parkinsonian dementia. J Neurol. 2006;253:242-7.

12. Calabresi P, Picconi B, Parnetti L, Di Filippo M. A convergent model for cognitive dysfunctions in Parkinson's disease: the critical dopamine–acetylcholine synaptic balance. Lancet Neurol. 2006; 5:974-983.
13. Kish SJ, Tong J, Hornykiewicz O, et al. Preferential loss of serotonin markers in caudate versus putamen in Parkinson's disease. Brain. 2008;131:120-31.
14. Vazey E, Aston-Jones G. The emerging role of norepinephrine in cognitive dysfunctions of Parkinson's disease. Front Behav Neurosci. 2012;6:48.
15. Ye Z, Altena E, Nombela C, et al. Selective serotonin reuptake inhibition modulates response inhibition in Parkinson's disease. Brain. 2014;137(4):1145-55.
16. Cools R. Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. Neurosci Biobehav Rev. 2006;30:1-23.
17. Poletti M, Bonuccelli U. Orbital and ventromedial prefrontal cortex functioning in Parkinson's disease: neuropsychological evidence. Brain Cogn. 2012;79:23-33.
18. Agid Y, Javoy-Agid F, Ruberg M. Biochemistry of neurotransmitters in Parkinson's disease. Mov Disord. 1987;2:166-230.
19. Gotham AM, Brown RG, Marsden CD. 'Frontal' cognitive function in patients with Parkinson's disease 'on' and 'off' levodopa. Brain. 1988;111:299-321.
20. Jellinger KA. Mild cognitive impairment in Parkinson disease: heterogenous mechanisms. J Neural Transm. 2013;120:157-67.
21. Matsumoto M. Dopamine signals and physiological origin of cognitive dysfunction in Parkinson's disease. Mov Disord. 2015;30:472-83.
22. Poletti M, Bonuccelli U. Acute and chronic cognitive effects of levodopa and dopamine agonists on patients with Parkinson's disease: a review. Ther Adv Psychopharmacol. 2013;3:101-13.
23. Grimes D, Gordon J, Snelgrove B, et al. [Canadian guidelines on Parkinson's Disease]. Can J Neurol Sci. 2012;39:S1-30; French.
24. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry. 1988;51:745-52.
25. Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol. 1999;56:33-9.
26. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. Neurology. 1967;17:427-42.
27. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations the Movement Disorder Society Task Force on rating scales for Parkinson's disease. Mov Disord. 2004;19:1020-8.
28. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed. Revised. Washington, DC: American Psychiatric Association; 1987, 608pp.
29. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. text revision. Washington, DC: American Psychiatric Association; 2000, 943pp.
30. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12:189-98.
31. Dubois B, Burn D, Goetz C, et al. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. Mov Disord. 2007;22:2314-24.
32. Moncrieff J, Churchill R, Drummond DC, McGuire H. Development of a quality assessment instrument for trials of treatments for depression and neurosis. Int J Meth Psychiatr Res. 2001;10:126-33.
33. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Edu Behav Stat. 1981;6:107-28.
34. Cohen J. Statistical analysis for the behavioral sciences, 2nd ed, Hillsdale: Lawrance Erlbaum Associates; 1988, 567pp.
35. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;4:863.
36. Hanagasi HA, Gurvit H, Unsalan P, et al. The effects of rasagiline on cognitive deficits in Parkinson's disease patients without dementia: a randomized, double-blind, placebo-controlled, multicenter study. Mov Disord. 2011;26:1851-8.
37. Growdon JH, Kieburtz K, McDermott MP, Panisset M, Friedman JH. Levodopa improves motor function without impairing cognition in mild non-demented Parkinson's disease patients. Neurology. 1998;50:1327-31.
38. Parkinson Study Group. DATATOP: a multicenter controlled clinical trial in early Parkinson's disease. Arch Neurol. 1989;46: 1052-60.
39. Kulisevsky J, Garcia-Sanchez C, Berthier ML, et al. Chronic effects of dopaminergic replacement on cognitive function in Parkinson's disease: a two-year follow-up study of previously untreated patients. Mov Disord. 2000;15:613-26.
40. Brusa L, Bassi A, Stefani A, et al. Pramipexole in comparison to l-dopa: a neuropsychological study. J Neural Trans. 2003;110: 373-380.
41. Brusa L, Pavino V, Massimetti MC, Bove R, Iani C, Stanzione P. The effect of dopamine agonists on cognitive functions in non-demented early-mild Parkinson's disease patients. Funct Neurol. 2013;28:13-7.
42. Fera F, Nicoletti G, Cerasa A, et al. Dopaminergic modulation of cognitive interference after pharmacological washout in Parkinson's disease. Brain Res Bull. 2007;74:75-83.
43. Brusa L, Tiraboschi P, Koch G, et al. Pergolide effect on cognitive functions in early-mild Parkinson's disease. J Neural Trans. 2005;112:231-7.
44. Kwak Y, Müller ML, Bohnen NL, Dayalu P, Seidler RD. l-DOPA changes ventral striatum recruitment during motor sequence learning in Parkinson's disease. Behav Brain Res. 2012; 230:116-24.
45. Costa A, Peppe A, Dell'Agnello G, Caltagirone C, Carlesimo GA. Dopamine and cognitive functioning in de novo subjects with Parkinson's disease: effects of pramipexole and pergolide on working memory. Neuropsychologia. 2009;47:1374-81.
46. Relja M, Klepac N. A dopamine agonist, pramipexole, and cognitive functions in Parkinson's disease. J Neurol Sci. 2006;248:251-4.
47. Antonelli F, Ko JH, Miyasaki J, et al. Dopamine-agonists and impulsivity in Parkinson's disease: impulsive choices vs. impulsive actions. Human Brain Mapping. 2014;35:2499-506.
48. Kieburtz K, McDermott M, Como P, et al. The effect of deprenyl and tocopherol on cognitive performance in early untreated Parkinson's disease. Neurology. 1994;44:1756.
49. Dalrymple-Alford JC, Jamieson CF, Donaldson IM. Effects of selegiline (deprenyl) on cognition in early Parkinson's disease. Clin Neuropharmacol. 1995;18:348-59.
50. Dixit SN, Behari M, Ahuja GK. Effect of selegiline on cognitive functions in Parkinson's disease. J Assoc Physicians India. 1999;47:784-6.
51. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatrics Soc. 2005;53:695-9.
52. Wechsler D. Wechsler memory scale (WMS-III). San Antonio, TX: Psychological Corporation; 1997.
53. Milner B. Interhemispheric differences in the localization of psychological processes in man. Br Med Bull. 1971;27:272-7.
54. Wechsler D. A standardized memory scale for clinical use. J Psychol. 1945;19:87-95.
55. Wechsler D. Measurement of adult intelligence. Baltimore, MD: Williams & Wilkins; 1958.
56. Wechsler D. WAIS-R manual: Wechsler adult intelligence scale-revised. New York, NY: Psychological Corporation; 1981, 156pp.
57. Werheid K, Hoppe C, Thöne A, Müller U, Müngersdorf M, Von Cramon DY. The Adaptive Digit Ordering Test: clinical application, reliability, and validity of a verbal working memory test. Arch Clin Neuropsychol. 2002;17:547-65.
58. Canivez GL. [Test review of the Weschler Adult Intelligence Scale-Fourth Edition]. In: Spies RA, Carlson JF, Geisinger KF, editors. The eighteenth mental measurements yearbook. Lincoln, NE: Buros Center for Testing; 2010.
59. Smith A. Symbol Digit Modalities Test. Los Angeles, CA: Western Psychological Services; 1973.
60. Iverson GL, Lovell MR, Collins MW. Validity of ImPACT for measuring processing speed following sports-related concussion. J Clin Exp Neuropsychol. 2005;27(6):683-9.
61. Koh CL, Lu WS, Chen HC, Hsueh IP, Hsieh JJ, Hsieh CL. Test–retest reliability and practice effect of the oral-format Symbol

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press

Digit Modalities Test in patients with stroke. Arch Clin Neuropsychol. 2011;26(4):356-63.
62. Pereira DR, Costa P, Cerqueira JJ. Repeated assessment and practice effects of the written Symbol Digit Modalities Test using a short inter-test interval. Arch Clin Neuropsychol. 2015;30(5):424-34.
63. Spinnler H, Tognoni G. Italien Groupe on the Neuropsychological Study of Ageing: Italian standardization and classification of neuropsychological tests. Ital J Neurol Sci. 1987;6:1-120.
64. Albert M, Moss M. The assessment of memory disorders in patients with Alzheimer's disease. In: Squire L, Butters N, editors. Neuropsychology of memory. New York, NY: Guilford Press; 1984, pp. 236-46.
65. Uc EY, McDermott MP, Marder KS, et al. Incidence of and risk factors for cognitive impairment in an early Parkinson disease clinical trial cohort. Neurology. 2009;73:1469-77.
66. van Zomeren AH, Brouwer WH. Assessment of attention. In: Crawford JR, Parker, DM, McKinlay WW, editors. A handbook of neuropsychological assessment. Hove, UK: Lawrence Erlbaum; 1992, pp. 241-66.
67. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol. 1935;18:643-62.
68. Homack S, Riccio CA. A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. Arch Clin Neuropsychol. 2004;19(6):725-43.
69. Army Individual Test Battery. Manual of directions and scoring. Washington, DC: War Department, Adjutant General's Office; 1944.
70. Giovagnoli AR, Del Pesce M, Mascheroni S, Simoncelli M, Laiacona M, Capitani E. Trail making test: normative values from 287 normal adult controls. Italian J Neurol Sci. 1996;17(4): 305-309.
71. Benton AL, Hamsher K, Sivan AB. Multilingual aphasia examination, 3rd ed. Iowa City, IA: AJA Associates; 1983, 30pp.
72. Levin BE, Llabre MM, Reisman S, et al. Visuospatial impairment in Parkinson's disease. Neurology. 1991;41(3):365.
73. McCaffrey RJ, Duff K, Westervelt HJ. editors Practitioner's guide to evaluating change with neuropsychological assessment instruments. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2000.
74. Riccio CA, Hynd GW. Validity of Benton's judgement of Line Orientation Test. J Psychoeduc Assess. 1992;10(3):210-8.
75. Buschke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology. 1974;24:1019.
76. Benton AL. Visual retention test, 4th ed. New York, NY: Psychological Corporation; 1974.
77. Delis DC, Kramer JH, Kaplan E, Ober BA., Fridlund A. California Verbal Learning Test, Research Edition (CVLT) Manual. San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich, Inc; 1983.
78. Delis DC, Kramer JH, Kaplan E, Thompkins BAO. CVLT: California Verbal Learning Test-adult version: manual. San Antonio, TX: Psychological Corporation; 1987.
79. Lindskog CO. [Test review of the California Verbal Learning Test, Second Edition, Adult Version]. In: Spies RA, Plake BS editors The sixteenth mental measurements yearbook. Lincoln, NE: Buros Center for Testing; 2005.
80. Rey A. [The clinical psychological examination]. Paris: Presses Universitaires de France; 1964, French.
81. Delaney RC, Prevey ML, Cramer J, Mattson RH. Test-retest comparability and control subject data for the rey-auditory verbal learning test and rey-osterrieth/taylor complex figures. Arch Clin Neuropsychol. 1992;7(6):523-8.
82. Wilson B, Cockburn J, Baddeley A. The Rivermead Behavioral Memory Test. Bury St. Edmunds, UK: Thames Valley Test; 1985.
83. Wilson B, Cockburn J, Baddeley A, Hiorns R. The development and validation of a test battery for detecting and monitoring everyday memory problems. J Clin Exp Neuropsychol. 1989; 11(6):855-870.
84. Tupler LA, Welsh KA, Asare-Aboagye Y, Dawson DV. Reliability of the Rey-Osterrieth Complex Figure in use with memory-impaired patients. J Clin Exp Neuropsychol. 1995;17(4): 566-579.
85. Wechsler D. Weschler memory scale-revised manual. San Antonio, TX: The Psychological Corporation; 1987, 150pp.
86. Moore PM, Baker GA. Psychometric properties and factor structure of the Wechsler Memory Scale-Revised in a sample of persons with intractable epilepsy. J Clin Exp Neuropsychol. 1997; 19(6):897-905.
87. Kaplan EF, Goodglass H, Weintraub S. The Boston naming test, 2nd ed. Philadelphia: Lea & Febiger; 1983.
88. Harry A, Crowe SF. Is the Boston Naming Test still fit for purpose? Clin Neuropsychologist. 2014;28(3):486-504.
89. Freedman M, Leach L, Kaplan E, Winocur G, Shulman KI, Delis DC. Clock drawing: a neuropsychological analysis. Oxford, UK: Oxford University Press; 1994.
90. Matthews CG, Klove K. Instruction manual for the adult neuropsychology test battery. Madison, WI: University of Wisconsin Medical School; 1964.
91. Rey A. [The psychological examination in cases of traumatic encepholopathy. Problems.]. Arch Psychol. 1941;28:215-85; French.
92. Osterrieth PA. [The test of copying a complex figure]. Arch Psychol. 1944;30:206-356; French.
93. Hovestadt A, De Jong GJ, Meerwaldt JD. Spatial disorientation in Parkinson's disease: no effect of levodopa substitution therapy. Neurology. 1998;38:1802.
94. Reitan RM. Manual for administration of neuropsychological test batteries for adults and children. Indianapolis, IN: Indiana University medical Center; 1969.
95. Raven JC, Court JH, Raven J, Kratzmeier H. Advanced progressive matrices. London, UK: HK Lewis; 1988.
96. Raven JC. Guide to the standard progressive matrices: sets A, B, C, D and E. London, UK: HK Lewis; 1960.
97. Flowers KA, Robertson C. The effect of Parkinson's disease on the ability to maintain a mental set. J Neurol Neurosurg Psychiatry. 1985;48:517-29.
98. Heaton RK. Wisconsin card sorting test manual. Odessa: Psychological Assessment Ressources; 1981.
99. Nelson HE. A modified card sorting test sensitive to frontal lobe defects. Cortex. 1976;12:313-24.
100. Bird CM, Papadopoulou K, Ricciardelli P, Rossor MN, Cipolotti L. Monitoring cognitive changes: psychometric properties of six cognitive tests. Br J Clin Psychol. 2004;43(2):197-210.
101. Ferland MB, Ramsay J, Engeland C, O'Hara P. Comparison of the performance of normal individuals and survivors of traumatic brain injury on repeat administrations of the Wisconsin Card Sorting Test. J Clin Exp Neuropsychol. 1998;20(4):473-482.
102. Kirby KN, Petry NM, Bickel WK. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. J Exp Psychol General. 1999;128:78.
103. Smith CL, Hantula DA. Methodological considerations in the study of delay discounting in intertemporal choice: a comparison of tasks and modes. Behav Res Methods. 2008;40(4):940-53.
104. Ballanger B, van Eimeren T, Moro E, et al. Stimulation of the subthalamic nucleus and impulsivity: release your horses. Ann Neurol. 2009;66:817-24.
105. Luria AR. The frontal lobes and the regulation of behavior. In: Primbram, KH, Luria AR, editors. Psychophysiology of the frontal lobes. New York, NY: Academic Press; 1973, pp 3-26.
106. Christensen AL. Luria's neuropsychological investigation. Madrid: Pablo del Rio; 1974.
107. Penner IK, Kobel M, Stöcklin M, Weber P, Opwis K, Calabrese P. The Stroop task: comparison between the original paradigm and computerized versions in children and adults. Clin Neuropsychologist. 2012;26(7):1142-53.
108. Krikorian R, Bartok J, Gay N. Tower of London procedure: a standard method and developmental data. J Clin Exp Neuropsychol. 1994;16:840-50.
109. Michalec J, Bezdicek O, Nikolai T, et al. A comparative study of tower of London scoring systems and normative data. Arch Clin Neuropsychol. 2017;32(3):328-38.
110. Lezak MD. Neuropsychological assessment, 3rd ed. Oxford, UK: Oxford University Press; 1995, 1046pp.
111. Spreen O, Strauss E. A compendium of neuropsychological tests: administration, norms, and commentary. Oxford, UK: Oxford University Press; 1991, 442pp.
112. Benton AL. Multi-lingual aphasia examination. Neuropsychologia. 1967;5:135-40.

113. Ruff RM, Light RH, Parker SB, Levin HS. Benton controlled oral word association test: reliability and updated norms. Arch Clin Neuropsychol. 1996;11(4):329-38.
114. Lezak MD, Howieson DB, Bigler ED, Tranel D. Neuropsychological assessment, 5th ed. New York, NY: Oxford University Press; 2012.
115. Strauss E, Sherman EM, Spreen O. A compendium of neuropsychological tests: administration, norms, and commentary. Oxford, NY: American Chemical Society; 2006.
116. Baddeley AD, Hitch G. Working memory. In: Bower GA, editor. Recent advances in learning and motivation, Vol. VIII, New York: Academic Press; 1974, pp. 47-89.
117. Baddeley AD. The episodic buffer: a new component of working memory? Trends Cogn Sci. 2000;4:417-23.
118. Gotham A, Brown R, Marsden C. Levodopa treatment may benefit or impair frontal function in Parkinson's disease. Lancet. 1986;2:970-1.
119. Müller T. Drug therapy in patients with Parkinson's disease. Transl Neurodegener. 2012;1:10.
120. Poletti M, Logi C, Lucetti C, et al. A single-center, cross-sectional prevalence study of impulse control disorders in Parkinson disease: association with dopaminergic drugs. J Clin Psychopharmacol. 2013;33:691-4.
121. Riedel O, Klotsche J, Spottke A, et al. Cognitive impairment in 873 patients with idiopathic Parkinson's disease. J Neurol. 2008;255: 255-64.
122. Hoops S, Nazem S, Siderowf AD, et al. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. Neurology. 2009;73:1738-45.
123. Weintraub D, Hauser RA, Elm JJ, Pagan F, Davis MD, Choudhry A. Rasagiline for mild cognitive impairment in Parkinson's disease: a placebo-controlled trial. Mov Disord. 2016;31(5):709-14.
124. Frakey LL, Friedman JH. Cognitive effects of rasagiline in mild-to-moderate stage Parkinson's disease without dementia. J Neuropsychiat Clin Neurosci. 2016;29(1):22-5.
125. Miyasaki JM, Martin W, Suchowersky O, Weiner WJ, Lang AE. Practice parameter: initiation of treatment for Parkinson's disease: an evidence-based review Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology. 2002;58(1):11-7.
126. National Institute for Health and Care Excellence. Parkinson's diease in over 20s: diagnosis and management. London: National Institute for Health and Care Excellence; 2006. Available at https://www.nice.org.uk/guidance/cg35/ifp/chapter/Medical-treatments#early-parkinsons-disease. Accessed May 20, 2017.
127. U.S. Food and Drug Administration. U.S. Food and Drug Administration Home Page. Silver Spring, MD: U.S. Food and Drug Administration [updated May 24, 2017]. Available at https://www.fda.gov/#start. Accessed May 26, 2017.
128. Molloy SA, Rowan EN, O'Brien JT, McKeith IG, Wesnes K, Burn DJ. Effect of levodopa on cognitive function in Parkinson's disease with and without dementia and dementia with Lewy bodies. J Neurol Neurosurg Psychiatry. 2006;77(12):1323-28.
129. Bocanegra Y, García AM, Pineda D, et al. Syntax, action verbs, action semantics, and object semantics in Parkinson's disease: dissociability, progression, and executive influences. Cortex. 2015;69:237-54.

https://doi.org/10.1017/cjn.2018.21 Published online by Cambridge University Press