Cognitive training in Parkinson disease
A systematic review and meta-analysis

ABSTRACT

Objective: To quantify the effects of cognitive training (CT) on cognitive and behavioral outcome measures in patients with Parkinson disease (PD).

Methods: We systematically searched 5 databases for randomized controlled trials (RCTs) of CT in patients with PD reporting cognitive or behavioral outcomes. Efficacy was measured as standardized mean difference (Hedges’ g) of post-training change.

Results: Seven studies encompassing 272 patients with Hoehn & Yahr Stages 1–3 were included. The overall effect of CT over and above control conditions was small but statistically significant (7 studies: g = 0.23, 95% confidence interval [CI] 0.014–0.44, p = 0.037). True heterogeneity across studies was low (I² = 0%) and there was no evidence of publication bias. Larger effect sizes were noted on working memory (4 studies: g = 0.74, CI 0.32–1.17, p = 0.001), processing speed (4 studies: g = 0.31, CI 0.01–0.61, p = 0.04), and executive function (5 studies: g = 0.30, CI 0.01–0.58, p = 0.042), while effects on measures of global cognition (4 studies), memory (5 studies), visuospatial skills (4 studies), and depression (5 studies), as well as attention, quality of life, and instrumental activities of daily living (3 studies each), were not statistically significant. No adverse events were reported.

Conclusions: Though still small, the current body of RCT evidence indicates that CT is safe and modestly effective on cognition in patients with mild to moderate PD. Larger RCTs are necessary to examine the utility of CT for secondary prevention of cognitive decline in this population.

Neurology® 2015;85:1843-1851

GLOSSARY

CI = confidence interval; CT = cognitive training; MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; PD = Parkinson disease; PEDro = Physiotherapy Evidence Database; RCT = randomized controlled trial; SMD = standardized mean difference.

Cognitive impairment is increasingly recognized as an important nonmotor symptom of Parkinson disease (PD).1 Neuropsychological impairments in PD are common, with the majority of patients showing at least some evidence of cognitive decline, while many progress to mild cognitive impairment (MCI)2 or dementia.3 These changes have a significant impact on quality of life in patients, as well as increasing caregiver burden and health care costs. Therefore, investigating potential methods of cognitive restoration is vital.4 Medications have been shown to have only limited benefit in the treatment of cognitive impairment in PD5 and nonpharmacologic interventions are of interest because the majority of patients with advanced PD are already burdened by complex polypharmacy. Cognitive training (CT) is one such option, which involves structured and theoretically driven teaching of strategies or guided practice on tasks that target particular cognitive domains.6 A recent meta-analysis has shown that computerized CT is efficacious on cognition in healthy older adults when supervised,7 and a systematic review found evidence for efficacy in MCI.8

*These authors contributed equally to this work.
From the Regenerative Neuroscience Group (I.H.K.L., H.H., M.V., A.L.) and Parkinson’s Disease Research Clinic (C.C.W., S.J.G.L.), Brain and Mind Centre, University of Sydney, Australia.
Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.
The Article Processing Charge was paid by the University of Sydney.
This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially.

© 2015 American Academy of Neurology 1843
© 2015 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.
Reviews of CT in PD have been reported previously. However, the literature has grown considerably since their publication, and meta-analytic techniques had not been previously employed. Furthermore, these reviews were not restricted to randomized controlled trials (RCTs) and combined results from CT studies with other cognitive interventions. Therefore, this study aims to quantitatively and systematically examine whether RCTs of strictly defined CT can improve cognitive and psychosocial outcomes in patients with PD.

METHODS This review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, was prospectively registered with PROSPERO, CRD42014012936, and follows methods established in our previous review of computerized CT in healthy elderly.

Eligibility criteria. We included published reports of RCTs examining behavioral effects (cognition, instrumental activities of daily living, quality of life, and depression) of CT in patients with PD. CT was defined as repeated practice on cognitively challenging tasks, including strategy training or drill-and-practice exercises using computers or pencil-and-paper approaches, for at least 4 hours. For studies that used CT in combination with other interventions (e.g., occupational therapy), we included only those that had CT as the differentiating condition between the study groups and where CT comprised at least 50% of the intervention.

Information sources and study selection. We systematically searched Medline (Ovid), Embase, PsycINFO, CINAHL, and CENTRAL for the term Parkinson’s in combination with widely used terms describing cognition-based intervention (see full search strategy in appendix e-1 on the Neurology Web site at Neurology.org) from inception to November 6, 2014. There was no limit on publication language. Reference and citation lists of relevant studies were manually scanned for potential eligible articles. One reviewer (I.H.K.L.) performed initial eligibility screening by assessing titles and abstracts of all results. Following initial screening, 2 independent reviewers (I.H.K.L. and A.L.) assessed full-text versions of potentially eligible articles.

Data collection and coding. Two reviewers (I.H.K.L. and C.C.W.) coded each outcome measure into one cognitive or behavioral domain based on the categorization provided in Straus et al., A Compendium of Neuropsychological Tests, whenever possible. Other outcomes were coded by consensus and approved by A.L. (see table e-1 for categorization of outcomes by domains). Outcomes were recorded as mean and SD for each group at baseline and follow-up or as means and SDs of pre-post change. All data were entered into Comprehensive Meta-Analysis (CMA) version 2 (Biostat, Englewood, NJ), and utilized an a priori pre-post correlation of 0.6.

Risk of bias in individual studies and study appraisal. Risk of bias in individual studies was conducted in accordance with the Cochrane Collaboration’s risk of bias tool. This tool assesses high, low, or unclear risk of bias in 6 categories: sequence generation; allocation concealment; blinding of participants, therapists, and assessors; incomplete outcome data; selective outcome reporting; and other sources of bias. We did not assess blinding of participants and therapists as such blinding is impractical in CT trials. We considered studies that lacked assessor blinding or did not adhere to intention-to-treat analysis (i.e., those with high risk for incomplete outcome data) as having high risk of bias. In addition, we used an adapted version of the Physiotherapy Evidence Database (PEDro-P) Rating Scale to assess the methodologic quality of the individual studies. The original PEDro scale consists of 11 items; however, 2 of the scale’s items that assessed blinding of therapists and patients were not considered due to impracticability in CT trials, and so the maximum possible PEDro score was set at 9. The assessment of each article was conducted by multiple independent reviewers (I.H.K.L., C.C.W., and H.H.). Disagreements were solved by a senior reviewer (A.L.). Table e-2 provides the results of risk of bias and PEDro assessments for each trial.

Statistical analysis. The unit of analysis was standardized mean difference (SMD) between CT and control groups of change from baseline to immediately post-training. We calculated SMD as Hedges’g with a 95% confidence interval (CI) for each outcome measure. We analyzed overall effects by calculating the mean SMD of all outcomes in each study and correcting for intercorrelation among outcomes by adjusting the mean variance by a factor of 0.7. SMD and variance from each study were then pooled using random-effects model. We analyzed domain-specific effects using a similar method, but intercorrelations among tests were assumed at 0.8. We used CMA for all analyses.

Positive values imply training-induced improvement in the CT group over and above control. An effect size of ≥ 0.50 was considered small, ≥ 0.60 was considered moderate, and ≥ 0.80 was considered large. We used the F statistic with 95% CI to quantify the proportion of true variance (i.e., variance from the true effect size rather than due to sampling error) from total observed variance. F values of 25%, 50%, and 75% imply low, moderate, and large proportions of variance from the true effect size (true heterogeneity), respectively.

Finally, we generated funnel plots for each analysis by charting SMDs against their standard errors in order to inspect for asymmetry that might suggest small study effect (publication bias). Planned analysis of funnel plot asymmetry using Egger test of the intercepts was not conducted as there were fewer than 10 studies in the review, which does not provide sufficient power for such an analysis. As a pragmatic alternative, we performed sensitivity analyses for domains with potential asymmetry by repeating the random-effects analysis after removal of outliers. Similarly, a planned series of subgroup analyses based on our previously published methods was not conducted due to an insufficient number of studies.

RESULTS Study selection. After removing duplicate entries, we screened 1,109 articles for initial eligibility, and excluded 1,014 articles based on their abstract and title. We then assessed the full-text versions of 95 full-text articles and found 8 studies eligible for inclusion. We requested summary data or clarifications from authors of 3 studies; 1 responded to and provided information, 1 responded but did not provide the requested information, and 1 did not respond. Finally, 1 study was excluded from the review as the original article did not report group summary data and these could not be obtained from the authors (figure 1). A study by Petrelli et al. presented data from 2 intervention groups, namely structured and unstructured training, and a passive
control group. Analysis of this study compared the structured to the passive control group due to lack of appropriate control for the second intervention.

**Characteristics of included studies.** The 7 studies included in this review included an overall number of 272 participants (CT, n = 133, mean group size = 19; controls, n = 139, mean group size = 20; table 1). Eighty-one outcomes were used to generate effect sizes. Mean age across samples ranged from 59.8 to 69.1 years. Approximately 57% of patients were male. Participants’ disease severity ranged between Hoehn & Yahr Stages 1 and 3. Five studies were conducted in Europe,20,22–24,27 1 in the United States,21 and 1 in Brazil.25 Five of the studies compared CT to an active control intervention (for description of individual studies, see table 1). The average PEDro score was 6.57/9 (SD 1.18). Six of the 7 studies were found to have a high risk of bias due to lack of adherence to intention-to-treat analysis, and lack of assessor blinding was noted in 2 studies.21,27 (table e-2).

Intervention design varied across studies (table 1). Four studies used computerized CT, 2 used paper-based CT, and 1 used a combination of paper-based and computerized exercises. Five studies trained participants in a center (group) settings, 1 provided training at home, and 1 included both. Session length ranged from 30 to 90 minutes, and total training time ranged from 7 to 36 hours. All studies provided 2–3 training sessions per week.

**Overall efficacy on cognitive outcomes.** The overall effect of CT on cognitive outcomes was small and statistically significant ($g = 0.23, 95\% CI 0.04–0.44, p = 0.037$; figure 2). True heterogeneity across studies was low ($F = 0\%, 95\% CI 0\%–68.58\%$). The funnel plot did not show substantial asymmetry (figure e-1).

**Domain-specific efficacy.** *Executive functions.* Five studies reported outcomes with measures of executive functions. The combined effect size was moderate and statistically significant ($g = 0.30, 95\% CI 0.01–0.58, p = 0.042$; figure 3). True heterogeneity across studies was low ($F = 0\%, 95\% CI 0\%–59.9\%$). The funnel plot did not show substantial asymmetry (figure e-1).

**Processing speed.** Four studies reported processing speed outcomes. The combined effect size was moderate and statistically significant ($g = 0.31, 95\% CI 0.01–0.61, p = 0.04$; figure 3). True heterogeneity across studies was low ($F = 0\%, 95\% CI 0\%–73.18\%$). The funnel plot did not show significant asymmetry (figure e-1), but one outlier with large effect size and low precision was detected.20 A sensitivity analysis excluding the outlier revealed large and statistically significant effect ($g = 0.62, 95\% CI 0.25–0.99, p = 0.001; F = 0\%, 95\% CI 0\%–95.16\%$).

**Global cognition.** Four studies reported global cognition outcomes. The combined effect size was moderate and statistically nonsignificant ($g = 0.32, 95\% CI −0.03 to 0.67, p = 0.065$; figure 3). True heterogeneity across studies was low ($F = 0\%, 95\% CI 0\%–61.58\%$), and the funnel plot did not show evidence of asymmetry (figure e-1).

**Memory.** Four studies reported memory outcomes. The combined effect size was small and statistically nonsignificant ($g = 0.13, 95\% CI −0.29 to 0.55, p = 0.55$; figure 4). True heterogeneity across studies was moderate ($F = 55.13\%, 95\% CI 0\%–83.34\%$), and the funnel plot did not show evidence of asymmetry (figure e-1).

**Visuospatial skills.** Four studies reported visuospatial outcomes. The combined effect size was negligible and statistically nonsignificant ($g = 0.01, 95\% CI −0.58 to 0.61, p = 0.96$; figure 4). True heterogeneity across studies was moderate ($F = 65.19\%, 95\% CI 0\%–74.59\%$). The funnel plot showed that the 2 least precise studies yielded the biggest effect sizes (figure e-1).
CI 0%–88.18%), and the funnel plot did not show evidence of asymmetry (figure e-1).

**Depression.** Five studies reported depression outcomes. The combined effect size was moderate and statistically nonsignificant ($g = 0.50$, 95% CI $-0.28$ to $1.28$, $p = 0.21$; figure 4). True heterogeneity across studies was large ($F = 85.49$, 95% CI 67.98%–93.43%). The funnel plot revealed one conspicuous outlier$^{22}$ (figure e-1). Removal of this study yielded a negligible and statistically nonsignificant combined effect size ($g = 0.11$, 95% CI $-0.28$ to $0.50$, $p = 0.58$; $F = 31.42$, 95% CI 0% to 75.37).

### Table 1 Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample characteristics</th>
<th>H&amp;Y range</th>
<th>Mean (SD) years since diagnosis</th>
<th>Study design: program description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. 20</td>
<td>15 (60) 59.70 (10.9)$^a$ 29.05 (1.1)$^a$ 1-3</td>
<td>3.35 (0.9)$^a$</td>
<td>CT: computerized CT program (RehaCom), 2 × 60 minutes per week for 6 weeks (group-based); control: a simple computerized visuomotor tapping task</td>
<td></td>
</tr>
<tr>
<td>Ref. 21</td>
<td>73 (69) 68.78 (8.1)$^a$ 28.07 (1.5)$^a$ 1-3</td>
<td>6.94 (5.5)$^a$</td>
<td>CT: computerized CT (InSight), 1-3 × 60 minutes per week for 13 weeks (home-based); control: no contact</td>
<td></td>
</tr>
<tr>
<td>Ref. 22</td>
<td>42 (68) 67.84 (8.4)$^a$ 27.05 (2.7)$^a$ 1-3</td>
<td>6.50 (5.2)$^a$</td>
<td>CT: structured paper-pencil tasks that target multiple domains (REHACOP), 3 × 60 minutes session per week for 12 weeks (group-based); control: basic occupational activities</td>
<td></td>
</tr>
<tr>
<td>Ref. 23</td>
<td>28 (50) 65.04 (9.2) 27.89 (1.4) 1-3</td>
<td>7.5 (6.8)</td>
<td>CT: multidomain training combining paper-pencil with computerized exercises (SmartBrain Tool), 3 × 45-minute per week for 4 weeks (group-based, in addition to home exercises); control: speech therapy</td>
<td></td>
</tr>
<tr>
<td>Ref. 24</td>
<td>43 (69) 69.15 (8.7) 27.9 (2.0) 1-3</td>
<td>5.47 (3.2)</td>
<td>CT: group-based multidomain training (NEUROvitalis), 2 × 90 minutes session per week for 6 weeks; control: no contact</td>
<td></td>
</tr>
<tr>
<td>Ref. 25</td>
<td>32 (67) 67.40 (8.1) 26.80 (2.4)$^b$ 1-2</td>
<td>5 (4.5)</td>
<td>CT: multidomain training with an integrative computerized CT program combining motor training with attention and working memory, 2 × 30 minutes per week for 7 weeks (group-based); control: balance exercises</td>
<td></td>
</tr>
<tr>
<td>Ref. 27</td>
<td>39 (68) 68.05 (8.3) 29$^c$ 2$^c$</td>
<td>5.15$^d$</td>
<td>CT: multidomain computerized training (CogniPlus), 3 × 40 minutes session per week for 4 weeks (group-based); control: Exergames (Nintendo Wi)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CT = cognitive training; H&Y = Hoehn & Yahr; MMSE = Mini-Mental State Examination.

$a$ Means for the complete sample (comprising subjects who were not included in the final analysis).

$b$ Measured with the Montreal Cognitive Assessment (1–30 range).

$c$ Average of median scores.

$d$ Based on subtracting mean age at diagnosis from mean age at baseline.

---

**Figure 2 Overall efficacy of cognitive training on all cognitive outcomes**

Tests for heterogeneity: $X^2 = 5.48, df = 6, p = 0.484, I^2 = 0$

Test for overall random effect: $Z = 2.09, p = 0.037$

Effect estimates are based on a random-effects model. CI = confidence interval; CT = cognitive training.
Other outcomes. Analyses of domains that were reported in only 3 studies each did not reveal statistically significant results (attention: $g = -0.13$, $p = 0.72$; instrumental activities of daily living: $g = 0.01$, $p = 0.93$; quality of life: $g = -0.10$, $p = 0.64$).

Adverse events. No adverse events related to CT were reported.

DISCUSSION Following previous findings from systematic reviews establishing the efficacy of CT on cognition in healthy older adults and MCI, we report that this intervention could potentially help to attenuate cognitive deficits in patients with PD. The current body of RCT evidence is small but of reasonable quality, and synthesis of outcomes found clinically meaningful improvements.
in overall cognition, as well as moderate to large effect sizes on measures of working memory, processing speed, and executive functions. Overall, our review provides the first high-level evidence that CT is efficacious on cognition in patients with PD.

The effect on Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment as measures of global cognition did not reach the threshold of statistical significance, though this may reflect relatively high baseline scores (average MMSE score range 26.8–29) as well as the insensitivity of these measures.
global tools as outcome measures. The MMSE in particular is known to be an unreliable tool in patients with PD, and this may have impacted the finding as 3 of the 4 studies assessing global cognitive outcomes reported MMSE scores.

In patients with PD without dementia, cognitive deficits are typically frontostriatal by nature. Thus, executive skills such as planning, cognitive flexibility, verbal fluency, and inhibitory control in addition to working memory type tasks have all been shown to be impaired in this patient group. That these functions are improved in response to CT is encouraging, providing strong support for continuation of CT trials in this population. Interestingly, improvements in executive functions were not shown in a recent meta-analysis in healthy older adults, possibly because deterioration in these domains is less pronounced in normal cognitive aging than in PD. Processing speed is another cognitive domain that is vitally important for everyday functioning, and typically shows declines in PD. Effects in this domain are consistent with those found in healthy elderly.

Memory did not demonstrate any statistically significant effect. Similarly, lack of effect on visuospatial skills, a key domain of PD-related cognitive deficits that responds well to CT in healthy older adults, warrants development of new CT exercises that target these domains. An analysis of depression yielded a negligible effect size after one outlier study was removed. However, depression scores in the samples were low; for example, the average Geriatric Depression Scale–15 score of the CT in the 2 studies that reported this outcome was 2.33 (SD 1.51), while the average Beck Depression Inventory–II score was 8.90 (SD 3.1) in 2 other studies; neither indicates depressive symptoms in the cohorts and thus a ceiling effect is likely.

The findings of the current meta-analysis are of particular interest given the lack of efficacy illustrated in pharmacologic treatments for cognitive decline in PD. An extensive evidence-based medical review of the area in 2011 showed that with the possible exception of rivastigmine there is insufficient evidence for pharmacologic therapy for dementia in PD. Thus, given its efficacy, safety, and relatively low cost, implementation of CT should be pushed forward as a pragmatic approach for maintaining cognition in PD.

The current body of evidence is thus compelling and warrants further studies aiming at establishing standards for CT in PD populations and clinical implementation. To achieve this goal, future studies will need to ensure adherence to the highest RCT standards, as several recent studies were excluded for non-RCT criteria (e.g., references 34–36), and while RCTs in this review reported relatively low attrition rates (all ≤ 15%), intention-to-treat analyses were not performed in 6 studies, thereby potentially inflating the results to some extent. Not least important is to ensure assessor blinding, mask interventions as much as possible by using active control groups or head-to-head comparisons of different CT interventions, and include large enough sample sizes to sufficiently power studies to detect effects on key clinical outcomes. Indeed, the typical trial in this review was modest in size (median n = 39), while the sample size needed to provide 80% power at the 0.05 level for an anticipated effect size of $g = 0.23$ is approximately 129, allowing for 15% attrition rate. Given objective difficulties recruiting and working with a PD population on a consistent basis as required for CT, the field might benefit from the inception of large, multicenter trials.

Clearly, the relatively small number of RCTs and their typically small sample sizes limited the precision of our findings and our ability to perform planned analyses in several domains. Similarly, we could not perform subgroup analyses that could indicate the relative efficacy of intervention design elements, due to the small number of studies as well as the lack of true heterogeneity (i.e., $F = 0\%$) in overall cognitive results. Further, neuropsychological classification into independent domains cannot accurately reflect the complex nature of these tasks, which often tap into multiple areas of cognition. Replication of this meta-analysis in the future will be crucial once further studies have been completed.

This meta-analysis suggests that CT leads to measurable improvements in cognitive performance in individuals with PD, particularly in working memory, executive functioning, and processing speed, which are typically impaired in the disease. Future RCTs employing large samples are required. The efficacy of CT in more cognitively impaired PD cohorts remains to be investigated.
support in the form of no-cost software from BrainTrain Inc. (USA) and HAPPYneuron Inc. (USA/France) for projects unrelated to this work. A. Lampit receives in-kind research support in the form of no-cost software from BrainTrain Inc. (USA) and HAPPYneuron Inc. (USA/France) for projects unrelated to this work. Go to Neurology.org for full disclosures.

Received February 6, 2015. Accepted in final form June 25, 2015.

REFERENCES


Introducing EBM Online—FREE to AAN Members!

The American Academy of Neurology again demonstrates its commitment to high-quality neurology education by converting and refining its popular classroom evidence-based medicine training into a convenient on-demand, self-paced online program. Only EBM Online from the AAN provides the trusted expertise of the source of the world’s most respected neurology guidelines.

EBM Online:
• Features five hours of convenient, interactive courses in 10 modules
• Uses common, real-life neurologic clinical examples
• Measures outcomes with pre-test, post-test, and module evaluations
• Provides helpful feedback on individual knowledge and more

This program is available now to residents and fellows; practicing neurologists are encouraged to register when CME is available in 2016.

Learn more at AAN.com/view/EBMOline.

Visit the Neurology® Web Site at Neurology.org
• Enhanced navigation format
• Increased search capability
• Highlighted articles
• Detailed podcast descriptions
• RSS Feeds of current issue and podcasts
• Personal folders for articles and searches
• Mobile device download link
• AAN Web page links
• Links to Neurology Now®, Neurology Today®, and Continuum®
• Resident & Fellow subsite

Find Neurology® on Facebook: http://tinyurl.com/neurologyfan
Follow Neurology® on Twitter: https://twitter.com/GreenJournal
Cognitive training in Parkinson disease: A systematic review and meta-analysis
Neurology 2015;85;1843-1851 Published Online before print October 30, 2015
DOI 10.1212/WNL.0000000000002145

This information is current as of October 30, 2015