Physical exercise improves quality of life, depressive symptoms, and cognition across chronic brain disorders: a transdiagnostic systematic review and meta-analysis of randomized controlled trials

Meenakshi Dauwan1,2,5 · Marieke J. H. Begemann1 · Margot I. E. Slot1 · Edwin H. M. Lee3 · Philip Scheltens4 · Iris E. C. Sommer1,5,6

Received: 19 March 2019 / Revised: 29 July 2019 / Accepted: 30 July 2019 © The Author(s) 2019

Abstract
We performed a meta-analysis to synthesize evidence on the efficacy and safety of physical exercise as an add-on therapeutic intervention for quality of life (QoL), depressive symptoms and cognition across six chronic brain disorders: Alzheimer’s disease, Huntington’s disease, multiple sclerosis, Parkinson’s disease, schizophrenia and unipolar depression. 122 studies (k) (n = 7231) were included. Exercise was superior to treatment as usual in improving QoL (k = 64, n = 4334, ES = 0.40, p < 0.0001), depressive symptoms (k = 60, n = 2909, ES = 0.78, p < 0.0001), the cognitive domains attention and working memory (k = 21, n = 1313, ES = 0.24, p < 0.009), executive functioning (k = 14, n = 977, ES = 0.15, p = 0.013), memory (k = 12, n = 994, ES = 0.12, p = 0.038) and psychomotor speed (k = 16, n = 896, ES = 0.23, p = 0.003). Meta-regression showed a dose–response effect for exercise time (min/week) on depressive symptoms (β = 0.007, p = 0.012). 69% of the studies that reported on safety, found no complications. Exercise is an efficacious and safe add-on therapeutic intervention showing a medium-sized effect on QoL and a large effect on mood in patients with chronic brain disorders, with a positive dose–response correlation. Exercise also improved several cognitive domains with small but significant effects.

Keywords Alzheimer’s disease · Multiple sclerosis · Parkinson’s disease · Depression · Schizophrenia · Physical exercise

Introduction
Chronic brain disorders are associated with reduced quality of life (QoL) [1–4], high prevalence of low mood and depression, stress sensitivity and cognitive dysfunction [5, 6]. These sequelae are interdependent, as depressive mood and cognitive impairment are two main factors influencing QoL [1, 2, 4–8], while cognition is negatively influenced by depression [9]. Moreover, these general sequelae are associated with various adverse consequences such as poor treatment compliance, loss of independence and even mortality [10]. In treatment of brain disorders, current clinical practice tends to focus on improving disease-specific symptoms (e.g., tremor and rigidity in Parkinson’s disease, psychosis in schizophrenia). Notably, however, patients with brain disorders regard QoL and depressive mood as more important for their health status than disease-specific physical and mental symptoms [11]. Therefore, improvement of these common features should become an important target in treatment of chronic brain disorders.

Exercise therapy may positively affect QoL, depression and cognition across disorders. A leading example is stroke, in which physical exercise has shown favorable effects in improving a wide range of symptoms, such that it has now been incorporated and recommended in guidelines as part of the standard treatment [12–16]. In contrast, research on the efficacy of physical exercise in treatment of other brain disorders is still in its infancy and therefore not part of the standard care. Although several studies have investigated the effect of physical exercise in different chronic brain disorders such as Alzheimer’s disease (AD) [17, 18], multiple sclerosis (MS) [19–21], Parkinson’s disease (PD) [22, 23],

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00415-019-09493-9) contains supplementary material, which is available to authorized users.

Marieke J. H. Begemann and Margot I. E. Slot contributed equally.
Schizophrenia (Sz) [24, 25] and unipolar depression (UD) [26–28], results and mainly recommendations for clinical practice have been highly diverse [29]. As a consequence, current evidence for efficacy of exercise therapy is still disputed and exercise is not part of the regular care offer for patients with aforementioned disorders in most countries.

Of note, the above-mentioned chronic brain disorders share underlying pathophysiological mechanisms. As such, neuroinflammation [30–33], imbalance in same neurotransmitter (e.g., dopamine in Sz and PD [34, 35], serotonin in Sz and UD [36]) and growth factors (e.g., brain-derived neurotrophic factor; BDNF) [37, 38], and disturbed connectivity (e.g., in default-mode network) [39–42] have been implicated in the pathophysiology of many of these brain disorders. Furthermore, a recent genome-wide association study (GWAS) showed high degree of genetic overlap among many psychiatric disorders stating that the different psychiatric disorders do not reflect independent diseases but rather represent different overlapping phenotypes of the same clinical spectra [43].

The aforementioned shows how disease-specific research has de-emphasized and limited our understanding of substantial commonalities that exist across disorders. Considering the overlap in pathophysiology and clinical picture across chronic brain disorders, commonalities across disorders outweigh the differences indicating that transdiagnostic and disease-specific treatments might be at least equally effective. Therefore, by targeting the common functional relationships across disorders with transdiagnostic treatments, both disease-specific and common shared factors can be targeted during treatment. Physical exercise can be such a transdiagnostic treatment for chronic brain disorders.

The objective of this study is to quantitatively review the effect of any type of physical exercise as an add-on intervention on QoL, depressive symptoms and/or cognition across many psychiatric disorders stating that the different psychiatric disorders do not reflect independent diseases but rather represent different overlapping phenotypes of the same clinical spectra [43].

Method

Literature search

This meta-analysis was performed according to the Preferred Reporting for Systematic Reviews and Meta-analysis (PRISMA) Statement [44]. A systematic search was performed in Pubmed (Medline), Embase, PsychInfo and Cochrane Database of Systematic Reviews (independently by MD, MS, and EL), using combinations of the following search terms: ‘Alzheimer’, ‘AD’, ‘Huntington’, ‘HD’, ‘multiple sclerosis’, ‘MS’, ‘Parkinson’, ‘PD’, ‘PDD’, ‘schizophrenia’, ‘psychosis’, ‘psychotic’, ‘depression’, ‘depressive’, ‘mood’, ‘affective’, ‘exercise’, ‘physical’, ‘training’, ‘endurance’, ‘aerobic’, ‘anaerobic’, ‘resistance’, ‘sport’ and ‘yoga’ (Online Resource 1), with no year or language limits. Additionally, the Web of Sciences databases and review articles were examined for cross-references. The search cutoff date was 15th of September 2018. When necessary, corresponding authors were contacted to provide full text details of the study outcome measures.

Inclusion criteria

By consensus (between MD, MS, EL, and IS), the following studies were included:

1. Randomized controlled trials (RCTs) investigating the effect of any type of physical exercise as an add-on intervention on QoL, depressive symptoms and/or cognition
2. Studies investigating whole-body, or upper- or lower-body exercise (i.e., organ-specific exercise such as respiration muscle or pelvic muscle training were excluded)
3. Studies including patients with a diagnosis of AD, HD, MS (idiopathic) PD, Sz [24] and UD (according to a diagnostic interview) in both the intervention and control group (i.e., mixed study populations were excluded)
4. RCTs with a cross-over design providing data for the first study period
5. Studies investigating combined interventions when the control group received the same non-exercise component of the intervention (e.g., exercise + medication versus medication only)
6. Studies investigating rehabilitation programs, provided that physical exercise constituted a main part of the program
7. Studies reported sufficient information to compute common effect size (ES) statistics [i.e., mean and standard deviations (SDs), exact F, p, t, or z values] or corresponding authors could provide these data upon request
8. If multiple publications were retrieved that described the same cohort, only the sample with largest overall sample size and/or original data was included

Exclusion criteria

1. Studies investigating same type of physical exercise in both the intervention and control group
2. Abstracts of studies (without full-text available) with insufficient information about the physical exercise intervention and/or outcome measures to calculate ES and untraceable corresponding information of the authors
Outcome measures

The outcome measures included pre- and post-intervention assessments (i.e., measured directly after finishing the intervention and thus does not include follow-up measurements) of QoL, depressive symptom severity and/or cognition. For measurements of depressive symptoms, observer-rated scales were preferred over self-rated questionnaires because of its higher validity [43]. The scales used to measure depression comprised Hamilton Depression Rating Scale (HDRS) [46], Beck Depression Inventory (BDI) [47], Montgomery Asberg Depression Rating Scale (MADRS) [48], Geriatric Depression Scale (GDS) [49], Patient Health Questionnaire-9 (PHQ-9) [50], and Profile of Mood States (POMS) [51].

Based on the cognitive domains and/or cognitive tests investigated across studies and disorders, the following six cognitive domains were classified: attention and working memory (A&WM), executive functioning (EF), memory (M), psychomotor speed (PS), verbal fluency (VF) and global cognition (GC) (Online Resource 2). To combine studies across disorders, the most stringent control group per disorder [i.e., treatment as usual (TAU) allowing treatments such as disease-specific medication, reading newspapers, educational sessions but no active treatments such as occupational therapy] was used as a reference group.

Assessment of risk of bias

According to the Cochrane Handbook of Systematic Reviews of Interventions [52], risk of bias was assessed for all eligible studies regarding selection bias, detection bias, attrition bias and reporting bias. Attrition bias was divided into assessment of incomplete outcome data (i.e., drop-out and exclusions) and intention-to-treat (ITT) analysis as ITT is considered the least biased method to measure intervention effects in RCTs [52]. Performance bias was not assessed, as it is usually not possible to blind study participants to whether or not exercise intervention is performed.

Data analysis

All analyses were performed using Comprehensive Meta-Analysis Version 2.0. Per outcome measure, the effect of additional exercise (versus control group) was quantified for each study using Hedges’g based on change scores (end of treatment minus baseline). When these were not reported, pre- and post-treatment mean values and SDs, or exact F, p, t, or z values were used. For studies that did not report exact SDs, these were calculated using the 95% confidence intervals (SD = sqrt(N) × [upper limit-lower limit]/[2 × 1.96]) or standard error (SE) (SD = SE × sqrt(N)).

To achieve a single pair-wise comparison between exercise and TAU, if a study investigated two or more types of exercise intervention, groups were combined for the main analysis [53] but studied separately in the moderator analysis (see further). The ES of the individual intervention groups were combined to calculate a composite ES by incorporating the ES and variance of each individual intervention while taking into account the correlation among the different interventions [54]. Likewise, when a study used more than one questionnaire to measure QoL or depressive symptoms, or multiple neuropsychological tests to measure a cognitive domain, a composite ES was calculated. As the correlation among interventions or test measures was mostly not reported, a correlation of 0.5 was taken for all the computations to avoid under- and overestimation of the overall ES [54].

Studies were combined in meta-analysis to calculate a mean weighted ES for each outcome measure (see Online Resource 3 for formulas). A random-effects model was considered appropriate given the heterogeneity across studies and diagnoses. Moreover, a random-effects model allows generalization of the results on population level [55]. ES were interpreted according to Cohen [56], with an ES of 0.2 indicating a small effect, 0.5 a medium and ≥ 0.8 a large effect. First, analyses were performed including all suitable studies per outcome measure. Subsequently, analyses were repeated by excluding outlier studies, defined as studies with standardized residual z scores of ES exceeding ± 1.96 (p < 0.05, two-tailed; shown in Figs. 2, 3, 4), studies with small total sample sizes (n < 20) because of high risk of sampling error in effect estimates [57] and studies with high risk of bias (i.e., considering the aim of the meta-analysis to study RCTs, studies classified as having high risk of bias on randomization and allocation concealment were excluded). ES with p < 0.05 were considered significant. Heterogeneity of results across studies was assessed by calculating the Q-statistic and I²-statistic. Q-Statistic tests the existence of heterogeneity and displays a Chi-square distribution with k−1 degrees of freedom (k = number of studies). Q values higher than the degrees of freedom indicate significant between-studies variability. I² describes the percentage of total variation across studies due to heterogeneity rather than chance. I² values of 25%, 50%, and 75% are considered as low, moderate, and high heterogeneity, respectively [58].

Potential publication bias was investigated by visual inspection of the funnel plots, with asymmetrical funnel plots indicating publication bias. When appropriate, the funnel plot asymmetry was tested with Egger’s test (p < 0.05, two-tailed) [59]. Additionally, Rosenthal’s fail-safe number (Nsf) was calculated for significant ES, estimating the
number of unpublished studies with non-significant results needed to bring the observed result to non-significance [60].

Moderator analyses

Subgroup analyses were performed for ‘type of exercise’ classified as aerobic, resistance, or neuromotor exercise (e.g., yoga) according to the American College of Sports Medicine (ACSM) Guideline [61].

Since an insufficient number of studies examined the effect of flexibility exercise only, analysis was not feasible for this type of exercise.

Random effects meta-regression analyses were conducted to evaluate the effect of the following continuous moderator variables using the unrestricted maximum likelihood model:

- Exercise time (min/week)
- Total length of the intervention period (weeks)
- Age (overall mean age across study groups per study)

If a study reported a range for any of these variables, the mean value of the variable was calculated from the upper and lower bounds. To include each pair-wise comparison separately in these sensitivity analyses, for studies with multiple intervention groups but one shared control group, the total number of participants in the control group were evenly divided up among the comparisons [53].

Since a large number of the included studies did not provide sufficient information about the intensity and safety of the exercise intervention and most of the included studies (80%) investigated supervised exercise intervention, a subgroup meta-regression analysis was not possible to investigate the effect of these parameters. The intensity and safety of the exercise interventions were assessed qualitatively.

Results

A total of 400 articles investigating the effect of any type of exercise intervention for patients with chronic brain disorders were retrieved from the literature search (AD: \(k = 40\), HD: \(k = 6\), MS: \(k = 137\), PD: \(k = 124\), Sz: \(k = 29\), UD: \(k = 64\), see Fig. 1.

A descriptive overview of these studies is provided in Online Resource 4. Of these, 163 studies fulfilled the inclusion criteria and were eligible for meta-analysis [62–224]. Forty-one studies provided insufficient information to compute common effect size. Therefore, a final total of 122 studies could be combined in meta-analysis. Risk of bias of all the eligible studies is shown in Online Resource 5 with a corresponding elaborative assessment of the studies included in the meta-analysis.

Quality of life

Sixty-four studies (\(n = 4334\)) examined the effect of exercise on QoL. Exercise showed a significant medium-size effect (ES = 0.40, 95% CI 0.27–0.52, \(p < 0.0001\); Fig. 2, Table 1). Heterogeneity was high \((Q(63) = 250.18, p < 0.0001; \hat{I}^2 = 75\%)\), indicating that 75% of the dispersion seen in Fig. 2 reflects difference in the true effect sizes while the remaining 25% can be attributed to random sampling error. Five studies [68, 142, 186, 200, 217] were identified as outliers, six studies [68, 119, 173, 200, 208, 216] had small sample sizes (\(n < 20\)) and another four studies [135, 140, 165, 193] were classified as having high risk of bias. After exclusion, ES decreased, but remained significant (\(k = 51, n = 3895\), ES = 0.31, 95% CI 0.19–0.43, \(p < 0.0001\)). Heterogeneity decreased, but remained moderate to high \((Q(50) = 159.13, p < 0.0001; \hat{I}^2 = 69\%)\). Funnel plot and Egger’s test indicated potential publication bias before \(t(62) = 5.00, p < 0.0001, N_R = 1898\), and after exclusion of the studies \(t(49) = 3.39, p < 0.010, N_R = 847\) but with very high fail-safe numbers (Table 1).

Within-disorder analysis showed a positive effect of exercise on QoL in patients with MS, PD and Sz (Table 2).

Depressive symptoms

Sixty studies (\(n = 2909\)) showed a significant large-size effect of exercise on depressive symptoms (ES = 0.78, 95% CI 0.58–0.98, \(p < 0.0001\); Fig. 3), with high heterogeneity \((Q(59) = 367.90, p < 0.0001; \hat{I}^2 = 84\%\); Table 1). Excluding eight outliers [75, 101, 104, 108, 112, 159, 220, 221] and seven small studies (\(n < 20\)) [68, 82, 87, 95, 190, 207, 225] and two studies [99, 193] with high risk of bias decreased the overall ES to a medium effect (\(k = 43 n = 2430, ES = 0.47, 95\% CI 0.32–0.62, p < 0.0001\)). Heterogeneity reduced to moderate to high \((Q(42) = 130.55, p < 0.0001; \hat{I}^2 = 68\%)\). Funnel plot and Egger’s test indicated potential publication bias \(t(58) = 6.10, p < 0.0001, N_R = 3937\), which remained after exclusion of the outliers \(t(41) = 3.97, p < 0.001, N_R = 1088\) (Table 1).

Within-disorder analysis showed a positive effect of exercise on depressive symptoms in AD, MS, Sz and UD (Table 2).

Cognition

Of the 120 studies, 36 studies (AD: \(k = 12\), HD: \(k = 3\), MS: \(k = 7\), PD: \(k = 7\), Sz: \(k = 3\), UD: \(k = 4\), examining 2125 patients, evaluated cognitive function and were included.
Attention and working memory

Exercise showed a significant small effect on attention and working memory ($k = 21, n = 1313, ES = 0.24, 95\% CI 0.06–0.41, p = 0.009; Fig. 4) with moderate heterogeneity ($Q(20) = 40.83, p = 0.004; I^2 = 51\%$). Eight (40\%) out of 20 studies comprised AD, HD or PD. The funnel plot and Egger’s test indicated potential publication bias [$t(19) = 2.14, p = 0.046, N_{ES} = 55$] (Table 1). The ES remained significant after excluding one outlier study [219], four small studies ($n < 20$) [163, 181, 190, 225] and one study [193] with high risk of bias ($k = 14, n = 923, ES = 0.25, 95\% CI 0.08–0.42, p = 0.004$).
<table>
<thead>
<tr>
<th></th>
<th>Studies (N)</th>
<th>Patients (IG/CG)</th>
<th>Mean age (years) (range)</th>
<th>Exercise time (min/week) (range)</th>
<th>Intervention duration (weeks) (range)</th>
<th>Hedges’ g</th>
<th>95% CI</th>
<th>P value</th>
<th>Q statistic (df)</th>
<th>I² (%)</th>
<th>Egger’s test</th>
<th>N_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>QoL</td>
<td>64</td>
<td>2349/1985</td>
<td>53.3 (15.4–78.0)</td>
<td>116.50 (40.0–142.5)</td>
<td>12.20 (4.0–52.0)</td>
<td>0.40</td>
<td>0.27 to 0.52</td>
<td>&lt; 0.0001</td>
<td>Q(63) = 250.18, 75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Without outliers</td>
<td>51</td>
<td>2091/1804</td>
<td>54.6 (15.4–78.0)</td>
<td>116.49 (40.0–360.0)</td>
<td>0.31</td>
<td>0.19 to 0.43</td>
<td>&lt; 0.0001</td>
<td>Q(50) = 159.13, 69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic exercise</td>
<td>9</td>
<td>257/250</td>
<td>54.7 (15.4–83.0)</td>
<td>128.75 (40.0–300.00)</td>
<td>13.31 (1.4–52.0)</td>
<td>0.45</td>
<td>0.16 to 0.75</td>
<td>0.003</td>
<td>Q(8) = 27.36, 71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuromotor exercise</td>
<td>10</td>
<td>254/215</td>
<td>54.7 (15.4–83.0)</td>
<td>128.75 (40.0–300.00)</td>
<td>13.31 (1.4–52.0)</td>
<td>0.35</td>
<td>0.07 to 0.64</td>
<td>0.013</td>
<td>Q(9) = 22.63, 60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resistance exercise</td>
<td>6</td>
<td>118/109</td>
<td>54.7 (15.4–83.0)</td>
<td>128.75 (40.0–300.00)</td>
<td>13.31 (1.4–52.0)</td>
<td>0.57</td>
<td>0.20 to 0.94</td>
<td>0.003</td>
<td>Q(5) = 4.19, 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All types of exercise</td>
<td>8</td>
<td>288/275</td>
<td>54.7 (15.4–83.0)</td>
<td>128.75 (40.0–300.00)</td>
<td>13.31 (1.4–52.0)</td>
<td>0.37</td>
<td>0.08 to 0.67</td>
<td>0.014</td>
<td>Q(7) = 26.93, 74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td>60</td>
<td>1635/1274</td>
<td>54.7 (15.4–83.0)</td>
<td>128.75 (40.0–300.00)</td>
<td>13.31 (1.4–52.0)</td>
<td>0.78</td>
<td>0.58 to 0.98</td>
<td>&lt; 0.0001</td>
<td>Q(59) = 367.90, 84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Without outliers</td>
<td>43</td>
<td>1364/1066</td>
<td>54.3 (15.4–83.0)</td>
<td>118.14 (40.0–210.0)</td>
<td>0.47</td>
<td>0.32 to 0.62</td>
<td>&lt; 0.0001</td>
<td>Q(42) = 130.55, 68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic exercise</td>
<td>17</td>
<td>493/415</td>
<td>55.8 (24.6–82.0)</td>
<td>118.57 (60.0–360.0)</td>
<td>15.36 (3.0–104.0)</td>
<td>0.40</td>
<td>0.16 to 0.65</td>
<td>0.001</td>
<td>Q(16) = 49.41, 68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuromotor exercise</td>
<td>8</td>
<td>176/143</td>
<td>55.8 (24.6–82.0)</td>
<td>118.57 (60.0–360.0)</td>
<td>15.36 (3.0–104.0)</td>
<td>0.55</td>
<td>0.18 to 0.91</td>
<td>0.001</td>
<td>Q(7) = 7.90, 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resistance exercise</td>
<td>4</td>
<td>75/69</td>
<td>55.8 (24.6–82.0)</td>
<td>118.57 (60.0–360.0)</td>
<td>15.36 (3.0–104.0)</td>
<td>0.96</td>
<td>0.44 to 1.48</td>
<td>&lt; 0.001</td>
<td>Q(3) = 6.22, 52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All types of exercise</td>
<td>2</td>
<td>135/139</td>
<td>55.8 (24.6–82.0)</td>
<td>118.57 (60.0–360.0)</td>
<td>15.36 (3.0–104.0)</td>
<td>0.06</td>
<td>– 0.53 to 0.64</td>
<td>0.854</td>
<td>Q(1) = 2.35, 57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention and working memory</td>
<td>21</td>
<td>794/519</td>
<td>55.8 (24.6–82.0)</td>
<td>118.57 (60.0–360.0)</td>
<td>15.36 (3.0–104.0)</td>
<td>0.24</td>
<td>0.06 to 0.41</td>
<td>0.009</td>
<td>Q(20) = 40.83, 51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Without outliers</td>
<td>14</td>
<td>547/376</td>
<td>57.8 (24.6–82.0)</td>
<td>100.18 (60.0–180.0)</td>
<td>0.25</td>
<td>0.08 to 0.42</td>
<td>0.004</td>
<td>Q(13) = 20.83, 38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic exercise</td>
<td>8</td>
<td>287/184</td>
<td>57.8 (24.6–82.0)</td>
<td>100.18 (60.0–180.0)</td>
<td>12.82 (6.0–24.0)</td>
<td>0.06</td>
<td>– 0.16 to 0.29</td>
<td>0.575</td>
<td>Q(7) = 13.27, 47</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 1 (continued)

<table>
<thead>
<tr>
<th></th>
<th>Studies (N)</th>
<th>Patients (IG/CG)</th>
<th>Mean age (years) (range)</th>
<th>Exercise time (min/week) (range)</th>
<th>Intervention duration (weeks) (range)</th>
<th>Hedges’ g</th>
<th>95% CI</th>
<th>P value</th>
<th>Q statistic (df)</th>
<th>ˆI² (%)</th>
<th>Egger’s test</th>
<th>N_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuromotor exercise</td>
<td>8</td>
<td>241/171</td>
<td>56.3 (24.6–78.8)</td>
<td>165.0 (60.0–480.0)</td>
<td>17.71 (3.0–52.0)</td>
<td>0.39</td>
<td>0.17 to 0.60</td>
<td><strong>0.001</strong></td>
<td>Q(7) = 6.84, p = 0.446</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive functioning</td>
<td>14</td>
<td>596/381</td>
<td>0.15</td>
<td>0.03 to 0.27</td>
<td><strong>0.013</strong></td>
<td>0.08</td>
<td>0.005 to 0.27</td>
<td>0.127</td>
<td>Q(13) = 12.30, p = 0.503</td>
<td>0</td>
<td>r(12) = 0.48, p = 0.641</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>10</td>
<td>565/351</td>
<td>0.17</td>
<td>0.04 to 0.29</td>
<td><strong>0.009</strong></td>
<td>0.08</td>
<td>0.04 to 0.29</td>
<td>0.127</td>
<td>Q(9) = 4.58, p = 0.869</td>
<td>0</td>
<td>r(8) = 1.54, p = 0.163</td>
<td></td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic exercise</td>
<td>7</td>
<td>316/241</td>
<td>0.20</td>
<td>0.06 to 0.35</td>
<td><strong>0.007</strong></td>
<td>0.08</td>
<td>– 0.13 to 0.29</td>
<td>0.465</td>
<td>Q(6) = 1.92, p = 0.927</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuromotor exercise</td>
<td>3</td>
<td>164/118</td>
<td>0.12</td>
<td>0.07 to 0.24</td>
<td><strong>0.038</strong></td>
<td>0.08</td>
<td>– 0.03 to 0.21</td>
<td>0.127</td>
<td>Q(11) = 10.74, p = 0.465</td>
<td>0</td>
<td>r(10) = 0.59, p = 0.568</td>
<td></td>
</tr>
<tr>
<td>Memory</td>
<td>12</td>
<td>609/385</td>
<td>0.12</td>
<td>0.07 to 0.24</td>
<td><strong>0.038</strong></td>
<td>0.08</td>
<td>– 0.03 to 0.21</td>
<td>0.127</td>
<td>Q(8) = 4.81, p = 0.777</td>
<td>0</td>
<td>r(7) = 0.90, p = 0.399</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>9</td>
<td>582/357</td>
<td>0.09</td>
<td>– 0.03 to 0.21</td>
<td>0.127</td>
<td>0.08</td>
<td>– 0.03 to 0.21</td>
<td>0.127</td>
<td>Q(8) = 4.81, p = 0.777</td>
<td>0</td>
<td>r(7) = 0.90, p = 0.399</td>
<td></td>
</tr>
<tr>
<td>Psychomotor speed</td>
<td>16</td>
<td>509/387</td>
<td>0.23</td>
<td>0.08 to 0.38</td>
<td><strong>0.003</strong></td>
<td>0.14</td>
<td>0.005 to 0.27</td>
<td>0.042</td>
<td>Q(15) = 19.02, p = 0.213</td>
<td>21</td>
<td>r(14) = 2.36, p = 0.035</td>
<td>42</td>
</tr>
<tr>
<td>Without outlier</td>
<td>10</td>
<td>454/332</td>
<td>0.14</td>
<td>0.005 to 0.27</td>
<td><strong>0.042</strong></td>
<td>0.14</td>
<td>0.005 to 0.27</td>
<td>0.042</td>
<td>Q(9) = 8.56, p = 0.479</td>
<td>0</td>
<td>r(8) = 1.02, p = 0.338</td>
<td></td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic exercise</td>
<td>8</td>
<td>338/247</td>
<td>0.09</td>
<td>– 0.07 to 0.24</td>
<td>0.276</td>
<td>0.14</td>
<td>– 0.07 to 0.24</td>
<td>0.276</td>
<td>Q(7) = 7.04, p = 0.425</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Neuromotor exercise</td>
<td>2</td>
<td>60/26</td>
<td>0.32</td>
<td>– 0.08 to 0.71</td>
<td>0.116</td>
<td>0.32</td>
<td>– 0.08 to 0.71</td>
<td>0.116</td>
<td>Q(1) = 0.66, p = 0.416</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>6</td>
<td>303/237</td>
<td>0.24</td>
<td>– 0.07 to 0.134</td>
<td>0.134</td>
<td>0.24</td>
<td>– 0.07 to 0.134</td>
<td>0.134</td>
<td>Q(5) = 14.36, p = 0.014</td>
<td>65</td>
<td>r(4) = 3.09, p = 0.037</td>
<td>3</td>
</tr>
<tr>
<td>Without outlier</td>
<td>5</td>
<td>288/222</td>
<td>0.06</td>
<td>– 0.15 to 0.27</td>
<td>0.569</td>
<td>0.06</td>
<td>– 0.15 to 0.27</td>
<td>0.569</td>
<td>Q(4) = 5.55, p = 0.236</td>
<td>28</td>
<td>r(3) = 2.48, p = 0.089</td>
<td></td>
</tr>
</tbody>
</table>
Heterogeneity turned low to moderate \( [Q(13) = 20.83, p = 0.076; I^2 = 38\%] \). Egger’s test was non-significant (Table 1).

### Executive functioning

Fourteen studies \((n = 977)\) showed a significant small effect of exercise on executive functioning (ES = 0.15, 95% CI 0.03–0.27, \(p = 0.013\); Fig. 4). Five (35.7%) out of 14 studies investigated physical exercise in AD, HD or PD. Studies were homogenous \([Q(13) = 12.30, p = 0.503; I^2 = 0\%]\). Egger’s test was non-significant (Table 1). After excluding one outlier \([63]\) and three small studies \([68, 163, 190]\), ES remained significant \((k = 10, n = 916, ES = 0.17, 95% CI 0.04–0.29, p = 0.009)\). There were no studies with high risk of bias.

### Memory

Twelve studies \((n = 994)\) examined the effect of physical exercise on memory and showed a beneficial small effect of exercise (involving mainly aerobic exercise) (ES = 0.12, 95% CI 0.07–0.24, \(p = 0.038\); Fig. 4). Four (33.3%) out of 2 studies comprised AD, HD or PD. Studies were homogenous \([Q(11) = 10.74, p = 0.465; I^2 = 0\%]\). Egger’s test was non-significant (Table 1). After excluding one outlier study \([128]\) and one small study \([225]\), ES was non-significant \((k = 9, n = 939, ES = 0.09, 95% CI −0.03 to 0.21, p = 0.127)\), while studies remained homogenous (Table 1).

### Psychomotor speed

Exercise showed a significant small effect on psychomotor speed \((k = 16, n = 896, ES = 0.23, 95% CI 0.08 to 0.38, p = 0.003; Fig. 4)\). Five (31.3%) out of 16 studies were based on AD, HD or PD. Heterogeneity among studies was low \([Q(15) = 19.02, p = 0.213; I^2 = 21\%]\). Funnel plot and Egger’s test indicated potential publication bias \([t(14) = 2.36, p = 0.035, N_R = 42]\). After excluding one outlier \([65]\) and four small studies \([162, 163, 190, 225]\), ES remained significant \((k = 10, n = 786, ES = 0.14, 95% CI 0.005–0.27, p = 0.042)\). Studies showed complete homogeneity and Egger’s test was non-significant (Table 1).

### Verbal fluency

Exercise showed no significant effect on verbal fluency \((k = 6, n = 540, ES = 0.24, 95% CI −0.07 to 0.55, p = 0.134; Fig. 4)\) and remained non-significant after excluding one outlier study \([65]\) \((k = 5, n = 510, ES = 0.06, 95% CI −0.15 to 0.27, p = 0.569)\). Four (66.7%) out of six studies comprised AD, HD or PD. Heterogeneity among studies was moderate.
### Table 2 Results per disorder for all outcome measures

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Studies (N)</th>
<th>Patients (IGiCG)</th>
<th>Hedges’ $g$</th>
<th>95% CI</th>
<th>$P$ value</th>
<th>$Q$ statistic (df)</th>
<th>$I^2$ (%)</th>
<th>Egger’s test$^b$</th>
<th>$N_{bg}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>QoL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>5</td>
<td>234/224</td>
<td>0.40</td>
<td>−0.10 to 0.91</td>
<td>0.119</td>
<td>$Q(4) = 23.51, p &lt; 0.0001$</td>
<td>83</td>
<td>$t(3) = 1.30, p = 0.283$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>4</td>
<td>227/217</td>
<td>0.22</td>
<td>−0.24 to 0.68</td>
<td>0.345</td>
<td>$Q(3) = 15.90, p = 0.001$</td>
<td>81</td>
<td>$t(2) = 0.47, p = 0.688$</td>
<td></td>
</tr>
<tr>
<td>Huntington’s disease</td>
<td>3</td>
<td>35/32</td>
<td>0.31</td>
<td>−0.25 to 0.88</td>
<td>0.280</td>
<td>$Q(2) = 3.39, p = 0.184$</td>
<td>41</td>
<td>$t(1) = 5.05, p = 0.124$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>2</td>
<td>26/23</td>
<td>0.05</td>
<td>−0.46 to 0.56</td>
<td>0.850</td>
<td>$Q(1) = 0.24, p = 0.626$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>25</td>
<td>909/641</td>
<td>0.41</td>
<td>0.24 to 0.58</td>
<td>&lt; 0.0001</td>
<td>$Q(24) = 72.61, p &lt; 0.0001$</td>
<td>67</td>
<td>$t(23) = 2.20, p = 0.038$</td>
<td>380</td>
</tr>
<tr>
<td>Without outlier</td>
<td>21</td>
<td>749/551</td>
<td>0.39</td>
<td>0.25 to 0.54</td>
<td>&lt; 0.0001</td>
<td>$Q(20) = 34.99, p = 0.020$</td>
<td>43</td>
<td>$t(19) = 1.15, p = 0.263$</td>
<td></td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>19</td>
<td>887/852</td>
<td>0.31</td>
<td>0.08 to 0.54</td>
<td>0.009</td>
<td>$Q(18) = 81.45, p &lt; 0.0001$</td>
<td>78</td>
<td>$t(17) = 2.94, p = 0.009$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>14</td>
<td>825/793</td>
<td>0.18</td>
<td>−0.04 to 0.41</td>
<td>0.112</td>
<td>$Q(13) = 52.43, p &lt; 0.0001$</td>
<td>75</td>
<td>$t(12) = 2.05, p = 0.063$</td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>5</td>
<td>130/88</td>
<td>0.89</td>
<td>0.22 to 1.55</td>
<td>0.009</td>
<td>$Q(4) = 21.02, p &lt; 0.0001$</td>
<td>81</td>
<td>$t(3) = 1.67, p = 0.194$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>3</td>
<td>110/72</td>
<td>0.43</td>
<td>−0.13 to 0.99</td>
<td>0.130</td>
<td>$Q(2) = 6.35, p = 0.042$</td>
<td>68</td>
<td>$t(1) = 0.11, p = 0.931$</td>
<td></td>
</tr>
<tr>
<td>Unipolar depression</td>
<td>7</td>
<td>154/148</td>
<td>0.34</td>
<td>−0.04 to 0.72</td>
<td>0.082</td>
<td>$Q(6) = 10.08, p = 0.004$</td>
<td>69</td>
<td>$t(5) = 0.64, p = 0.552$</td>
<td></td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>5</td>
<td>264/254</td>
<td>0.80</td>
<td>0.12 to 1.49</td>
<td>0.022</td>
<td>$Q(4) = 48.15, p &lt; 0.0001$</td>
<td>92</td>
<td>$t(3) = 3.46, p = 0.041$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>3</td>
<td>237/227</td>
<td>0.05</td>
<td>−0.16 to 0.24</td>
<td>0.653</td>
<td>$Q(2) = 2.38, p = 0.305$</td>
<td>16</td>
<td>$t(1) = 0.005, p = 0.997$</td>
<td></td>
</tr>
<tr>
<td>Huntington’s disease</td>
<td>2</td>
<td>24/24</td>
<td>0.40</td>
<td>−0.76 to 1.56</td>
<td>0.496</td>
<td>$Q(1) = 4.03, p = 0.045$</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>14</td>
<td>327/249</td>
<td>0.45</td>
<td>0.12 to 0.79</td>
<td>0.007</td>
<td>$Q(13) = 47.78, p &lt; 0.0001$</td>
<td>73</td>
<td>$t(12) = 2.30, p = 0.040$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>13</td>
<td>291/231</td>
<td>0.23</td>
<td>0.06 to 0.40</td>
<td>0.010</td>
<td>$Q(12) = 9.61, p = 0.650$</td>
<td>3</td>
<td>$t(11) = 3.59, p = 0.004$</td>
<td></td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>5</td>
<td>116/100</td>
<td>0.05</td>
<td>−0.36 to 0.45</td>
<td>0.822</td>
<td>$Q(4) = 7.91, p = 0.095$</td>
<td>49</td>
<td>$t(3) = 0.83, p = 0.469$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>3</td>
<td>89/77</td>
<td>-0.04</td>
<td>−0.63 to 0.55</td>
<td>0.895</td>
<td>$Q(2) = 6.22, p = 0.045$</td>
<td>68</td>
<td>$t(1) = 0.20, p = 0.874$</td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>2</td>
<td>46/21</td>
<td>0.73</td>
<td>0.20 to 1.26</td>
<td>0.007</td>
<td>$Q(1) = 0.89, p = 0.347$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>1</td>
<td>42/15</td>
<td>0.62</td>
<td>0.04 to 1.19</td>
<td>0.037</td>
<td>$Q(0) = 0.00, p = 1.000$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unipolar depression</td>
<td>32</td>
<td>858/626</td>
<td>1.08</td>
<td>0.78 to 1.38</td>
<td>&lt; 0.0001</td>
<td>$Q(31) = 210.96, p &lt; 0.0001$</td>
<td>85</td>
<td>$t(30) = 4.83, p &lt; 0.0001$</td>
<td></td>
</tr>
<tr>
<td>Without outliers</td>
<td>23</td>
<td>736/523</td>
<td>0.88</td>
<td>0.62 to 1.14</td>
<td>&lt; 0.0001</td>
<td>$Q(22) = 101.96, p &lt; 0.0001$</td>
<td>78</td>
<td>$t(21) = 4.18, p &lt; 0.001$</td>
<td></td>
</tr>
<tr>
<td>Cognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention and working memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>3</td>
<td>44/43</td>
<td>0.28</td>
<td>−0.13 to 0.69</td>
<td>0.185</td>
<td>$Q(2) = 2.30, p = 0.317$</td>
<td>13</td>
<td>$t(1) = 5.29, p = 0.119$</td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>5</td>
<td>117/95</td>
<td>0.23</td>
<td>−0.04 to 0.49</td>
<td>0.089</td>
<td>$Q(4) = 4.16, p = 0.384$</td>
<td>4</td>
<td>$t(3) = 0.67, p = 0.550$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>4</td>
<td>112/90</td>
<td>0.24</td>
<td>−0.07 to 0.56</td>
<td>0.134</td>
<td>$Q(3) = 4.16, p = 0.245$</td>
<td>28</td>
<td>$t(2) = 1.16, p = 0.365$</td>
<td></td>
</tr>
<tr>
<td>Outcome measure</td>
<td>Studies (N)</td>
<td>Patients (IG/CG)</td>
<td>Hedges' $g$</td>
<td>95% CI</td>
<td>$P$ value</td>
<td>$Q$ statistic (df)</td>
<td>$I^2$ (%)</td>
<td>Egger's test</td>
<td>$N_R$</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>-----------</td>
<td>--------------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Parkinson’s disease</strong></td>
<td>5</td>
<td>89/82</td>
<td><strong>0.50</strong></td>
<td><strong>0.20 to 0.80</strong></td>
<td><strong>0.001</strong></td>
<td>$Q(4) = 2.63$, $p = 0.622$</td>
<td>0</td>
<td>$t(3) = 0.05$, $p = 0.962$</td>
<td></td>
</tr>
<tr>
<td>Without outliers</td>
<td>2</td>
<td>57/54</td>
<td>0.41</td>
<td>− 0.12 to 0.94</td>
<td>0.129</td>
<td>$Q(1) = 1.61$, $p = 0.205$</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>4</td>
<td>373/184</td>
<td>0.07</td>
<td>− 0.41 to 0.55</td>
<td>0.776</td>
<td>$Q(3) = 14.57$, $p = 0.002$</td>
<td>79</td>
<td>$t(2) = 0.54$, $p = 0.642$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>3</td>
<td>365/174</td>
<td>0.04</td>
<td>− 0.51 to 0.60</td>
<td>0.879</td>
<td>$Q(2) = 14.32$, $p = 0.001$</td>
<td>86</td>
<td>$t(1) = 1.44$, $p = 0.386$</td>
<td></td>
</tr>
<tr>
<td>Unipolar depression</td>
<td>4</td>
<td>171/115</td>
<td>0.22</td>
<td>− 0.24 to 0.68</td>
<td>0.351</td>
<td>$Q(3) = 8.72$, $p = 0.033$</td>
<td>66</td>
<td>$t(2) = 1.14$, $p = 0.373$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>3</td>
<td>163/107</td>
<td>0.17</td>
<td>− 0.36 to 0.70</td>
<td>0.540</td>
<td>$Q(2) = 7.79$, $p = 0.020$</td>
<td>74</td>
<td>$t(1) = 0.78$, $p = 0.578$</td>
<td></td>
</tr>
<tr>
<td><strong>Executive functioning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>3</td>
<td>78/82</td>
<td>0.03</td>
<td>− 0.58 to 0.64</td>
<td>0.921</td>
<td>$Q(2) = 5.21$, $p = 0.074$</td>
<td>62</td>
<td>$t(1) = 0.0005$, $p = 1.000$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>2</td>
<td>71/75</td>
<td>-0.17</td>
<td>− 0.86 to 0.52</td>
<td>0.628</td>
<td>$Q(1) = 3.16$, $p = 0.076$</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>4</td>
<td>76/56</td>
<td>0.15</td>
<td>− 0.18 to 0.47</td>
<td>0.370</td>
<td>$Q(3) = 2.00$, $p = 0.572$</td>
<td>0</td>
<td>$t(2) = 0.49$, $p = 0.673$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>3</td>
<td>71/51</td>
<td>0.21</td>
<td>− 0.13 to 0.56</td>
<td>0.223</td>
<td>$Q(2) = 0.74$, $p = 0.692$</td>
<td>0</td>
<td>$t(1) = 0.82$, $p = 0.564$</td>
<td></td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>2</td>
<td>24/16</td>
<td>0.28</td>
<td>− 0.25 to 0.80</td>
<td>0.306</td>
<td>$Q(1) = 0.70$, $p = 0.402$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>1</td>
<td>15/8</td>
<td>0.08</td>
<td>− 0.62 to 0.78</td>
<td>0.827</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>2</td>
<td>263/125</td>
<td>0.17</td>
<td>− 0.21 to 0.55</td>
<td>0.386</td>
<td>$Q(1) = 2.88$, $p = 0.090$</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unipolar depression</td>
<td>2</td>
<td>146/91</td>
<td>0.20</td>
<td>− 0.01 to 0.42</td>
<td>0.065</td>
<td>$Q(1) = 0.04$, $p = 0.835$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>3</td>
<td>127/110</td>
<td>0.05</td>
<td>− 0.18 to 0.28</td>
<td>0.666</td>
<td>$Q(2) = 0.75$, $p = 0.688$</td>
<td>0</td>
<td>$t(1) = 0.31$, $p = 0.811$</td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>2</td>
<td>48/30</td>
<td>0.48</td>
<td>− 0.53 to 1.48</td>
<td>0.352</td>
<td>$Q(1) = 4.64$, $p = 0.031$</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>3</td>
<td>271/135</td>
<td>0.13</td>
<td>− 0.07 to 0.33</td>
<td>0.201</td>
<td>$Q(2) = 0.89$, $p = 0.641$</td>
<td>0</td>
<td>$t(1) = 1.01$, $p = 0.496$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>2</td>
<td>263/125</td>
<td>0.12</td>
<td>− 0.09 to 0.33</td>
<td>0.250</td>
<td>$Q(1) = 0.79$, $p = 0.376$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unipolar depression</td>
<td>3</td>
<td>154/99</td>
<td>0.17</td>
<td>− 0.04 to 0.38</td>
<td>0.104</td>
<td>$Q(2) = 0.77$, $p = 0.680$</td>
<td>0</td>
<td>$t(1) = 0.69$, $p = 0.615$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>2</td>
<td>146/91</td>
<td>0.16</td>
<td>− 0.05 to 0.38</td>
<td>0.136</td>
<td>$Q(1) = 0.67$, $p = 0.413$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychomotor speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>3</td>
<td>127/113</td>
<td>0.49</td>
<td>− 0.32 to 1.29</td>
<td>0.237</td>
<td>$Q(2) = 10.38$, $p = 0.006$</td>
<td>81</td>
<td>$t(1) = 1.62$, $p = 0.352$</td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>6</td>
<td>133/113</td>
<td>0.24</td>
<td>− 0.008 to 0.48</td>
<td>0.058</td>
<td>$Q(5) = 3.22$, $p = 0.667$</td>
<td>0</td>
<td>$t(4) = 0.68$, $p = 0.533$</td>
<td></td>
</tr>
<tr>
<td>Without outliers</td>
<td>4</td>
<td>118/99</td>
<td>0.22</td>
<td>− 0.04 to 0.48</td>
<td>0.099</td>
<td>$Q(3) = 2.63$, $p = 0.452$</td>
<td>0</td>
<td>$t(2) = 0.20$, $p = 0.858$</td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>2</td>
<td>77/43</td>
<td><strong>0.45</strong></td>
<td><strong>0.07 to 0.83</strong></td>
<td><strong>0.020</strong></td>
<td>$Q(1) = 0.02$, $p = 0.886$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>1</td>
<td>69/33</td>
<td><strong>0.44</strong></td>
<td><strong>0.02 to 0.85</strong></td>
<td><strong>0.040</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unipolar depression</td>
<td>3</td>
<td>154/99</td>
<td>0.18</td>
<td>− 0.05 to 0.41</td>
<td>0.133</td>
<td>$Q(2) = 1.84$, $p = 0.398$</td>
<td>0</td>
<td>$t(1) = 1.74$, $p = 0.332$</td>
<td></td>
</tr>
<tr>
<td>Without outlier</td>
<td>2</td>
<td>146/91</td>
<td>0.14</td>
<td>− 0.10 to 0.38</td>
<td>0.238</td>
<td>$Q(1) = 0.48$, $p = 0.487$</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
to high \( Q(5) = 14.36, p = 0.014; I^2 = 65\% \); Table 1] but decreased after excluding the outlier (Table 1).

**Global cognition**

Fifteen studies \((n = 725)\), all comprising AD, HD or PD, showed a trend of exercise in improving global cognition (ES = 0.30, 95% CI − 0.03 to 0.63, \( p = 0.076 \); Fig. 4). ES increased and showed significance \((k = 10, n = 620, ES = 0.39, 95\% CI 0.09–0.68, p = 0.010)\) after excluding two outliers [63, 74], three small studies [68, 119, 190] and one study [193] with high risk of bias. Heterogeneity was high \[ Q(14) = 60.79, p < 0.0001; I^2 = 77\% \] but decreased after exclusion of the studies \[ Q(9) = 50.92, p < 0.0001 \] (Table 1).

Egger’s test was non-significant (Table 1).

Separate analyses per disorder showed beneficial effects of exercise on A and WM in PD, PS in Sz and on GC in AD and PD (Table 2).

The study by Oertel Knöchel et al. [105] and Maci et al. [68] investigated physical exercise in combination with a cognitive intervention. Exclusion of these studies did not change results for any of the outcome measures.

**Studies with ITT-analysis**

Additional analyses with studies with only low or unclear risk of bias on ITT analyses showed even larger effect of exercise on both QoL (ES = 0.56) and depressive symptoms (ES = 0.90), while effect on the cognitive domain psychomotor speed remained small (ES = 0.24) but significant. Effect of physical exercise on all the other cognitive domains was no longer significant. See Online Resource 6 for a detailed overview of these results.

**Moderator analysis**

Subgroup analysis showed a significant medium effect of aerobic and neuromotor exercise and a medium-to-large effect of resistance exercise on QoL and depressive symptoms. Furthermore, a comprehensive program including all types of exercises according to ACSM was also effective in improving QoL. For cognition, aerobic and neuromotor exercises showed significant effects (Table 1).

Meta-regression analysis showed a small but positive dose–response effect for the amount of weekly exercise in min/week in reducing depressive symptoms \((\beta = 0.007, 95\% CI 0.002–0.013, p = 0.012; Online Resource 7–8)\), indicating that every 1-min increase in exercise intervention per week corresponds to an 0.007 unit increase is ES. No significant effect was found for the moderator total length of intervention (range 1.4–104 weeks). Additional meta-regression results are shown in Online Resource 7.

**Intensity**

With regard to intensity of the exercise intervention as possible moderator, 50 of the analyzed studies (41.0%) did not report any information. Of the remaining 59.0%, 18 studies (25.0%) investigated neuromotor exercises and therefore possibly could not report any intensity level. 36 studies (50.0%) applied low-to-moderate intensity of exercise, while 16 studies (22.2%) investigated moderate-to-high intensity.
exercise. Two studies (2.8%) investigated low-to-high intensity exercise (Online Resource 9).

Safety

Sixty-five studies (53.3%) reported on safety aspects of the exercise intervention (Online Resource 10). Forty-five of these studies (69.2%) found no physical injuries related to exercise. Eighteen studies (27.7%) found physical injuries that were related to the exercise intervention. These consisted mainly of muscle/joint pain (17.5%), fall incidents (11.4%, all with complete recovery) and ankle sprain (1.9%). In 83.3% of these studies (k = 15), physical injuries were short-lasting and/or had no consequences for participation in and completion of the exercise intervention.

Fig. 3 Meta-analysis of the effect of physical exercise on depressive symptoms. Effect sizes (ES) per study and the overall ES are in Hedges’ g with corresponding p values and sample size of the intervention and control group. Standardized residual z scores of ES were used to detect outlier studies
**Fig. 4** Meta-analysis of the effect of physical exercise on the cognitive domains (from top to down) attention and working memory, executive functioning, memory, psychomotor speed, verbal fluency and global cognition. Effect sizes (ES) per study and the overall ES are in Hedges’ *g* with corresponding *p* values and sample size of the intervention and control group. Standardized residual *z* scores of ES were used to detect outlier studies.
Discussion

One hundred and twenty-two studies, including 7231 patients, showed a significant medium-size effect (ES = 0.40) of exercise as an add-on therapeutic intervention on QoL (k = 64, n = 4334), a large effect (ES = 0.78) on depressive symptoms (k = 60, n = 2909) and a small but significant effect (ES = 0.12–0.24) on improving function in several cognitive domains. The effects for QoL and depression were well powered. The included number of patients was lower for cognition (k = 36, n = 2125), which makes these results more sensitive for new findings. From the studies that reported on safety (k = 18), low incidences of complications related to the exercise interventions were found, which had no lasting consequences for participation in and completion of the exercise interventions.

Current clinical practice

In present clinical practice, the role of physical exercise as an add-on therapy in the management of QoL, depressive symptoms and cognitive impairment in chronic brain disorders remains elusive [226–228]. Management guidelines sometimes suggest physical exercise in treatment of, e.g., physical health, motor symptoms, falls and fatigue in chronic brain disorders but lack in clarity over the effectiveness of physical exercise on the studied symptoms [229–235].

Chronic brain disorders commonly affect well being and QoL. Therefore, improvement of QoL is a main care objective in these disorders. Depressed mood and cognitive inabilities are important contributors to reduce QoL. Currently, evidence for treatment designed specifically to target QoL is lacking. Most treatments for chronic brain disorders alleviate disease-specific symptoms, progression or relapse. In contrast, exercise therapy targets overall well-being, mood and cognition, independent of type of disease.

At present, physical exercise is not generally viewed as an effective intervention. For example, in a recent review, Kok et al. evaluated treatment of depression in older adults and stated that depressive symptoms can be effectively treated with antidepressants whereas physical exercise may not be a mainstream treatment modality, yet might be considered as a complementary therapy [236]. In contrast, Turner et al., showed that the efficacy of antidepressants is subject to selective publication of positive studies with a precipitous drop in ES to an overall ES of 0.32 when non-published FDA approved drug trials of antidepressants were combined with published drug trials [237].

For dementia, there are still no disease-modifying agents available and treatment is limited to amelioration of symptoms [238]. The effects for cognition found in our meta-analysis are statistically small but significant and similar or larger than effects of cognitive therapy [239–244] or drug treatment [245–248], which makes these effects relevant for cognitive outcomes.

Heterogeneity and moderators

To our knowledge, this is the first meta-analysis to assess the effect of physical exercise interventions across chronic brain disorders. Since heterogeneity between studies is a valid reason of concern in meta-analyses, our study shows that when we consider brain disorders to share underlying mechanisms, it is feasible to combine disorders and studies across disorders in a joint analysis. We found lower heterogeneities in the joint analysis compared to within-disorder analysis. High heterogeneity across studies and disorders was accounted for using the random-effects model and excluding outlier studies, small studies and studies with high risk of bias. As a consequence, for QoL and depressive symptoms, both heterogeneity and ES decreased, but exercise still showed a significant medium effect. Moderator analyses, performed to assess potential sources of heterogeneity, showed moderate variability between studies that investigated aerobic exercises whereas studies that evaluated the efficacy of resistance or neuromotor exercises on QoL and depressive symptoms showed higher ES and no heterogeneity. Largest effects were found for resistance exercise. Better performance of resistance exercise on these outcomes might be mediated by an increase in peripheral blood levels of Insulin-growth-factor-1 (IGF-1), which can cross the blood–brain barrier and has been shown to regulate the effects of exercise on depression, learning, angiogenesis and hippocampal neurogenesis [249, 250]. As one study evaluated the role of resistance exercise only on cognition, this result should be interpreted with caution. Heterogeneity across studies assessing cognition was low or completely lacking for all but two cognitive domains (i.e., attention and working memory and global cognition) that showed significant results. For cognition, neuromotor exercise resulted in higher effects than aerobic exercise. Neuromotor exercises involve multifaceted exercises that target different brain systems involved in the regulation of attention, balance, coordination, mood, motor functioning and cognition, amongst others. Hence, neuromotor exercises are suggested to improve synchronization between different brain areas, which might explain their efficacy on a wide variety of clinical symptoms [251].

We found a positive dose–response effect for the weekly time spent on exercise in min/week in reducing depressive symptoms, indicating that the more time spent on exercise per week, the larger the reduction in depressive symptoms. However, no significant dose–response effect was found for the total length of the exercise intervention (i.e., the number
of weeks spend on exercise), suggesting that both short- and long-term exercise interventions might be beneficial in improving QoL, depressive symptoms, and cognition. Patient groups ranged in mean age from 15.4 to 84.0 years, but no significant effect of this moderator was found on the outcome measures indicating that the effect of exercise on the examined outcome measure is not age-dependent.

Regarding exercise intensity, most of the studies that provided information on the intensity of the studied exercise intervention, applied moderate exercise intensity. Additionally, we found that risk of possible complications due to exercise is low, which should not be considered a limiting factor for exercise intervention.

While all aforementioned moderators were expected to be an explanatory factor for the high heterogeneity in QoL, depressive symptoms and the cognitive domain global cognition, the role of exercise intensity and safety could not be assessed quantitatively. One other explanation for the high heterogeneity could be the different questionnaires used in the separate studies. For both QoL and depressive symptoms, 13 different rating scales were used. For global cognition, six different tests were used.

**Implications for clinical practice**

Currently, physical exercise is not a standard part of the treatment of the six chronic brain disorders included in this study. Based on our work, it is likely that patients with any of the investigated brain disorders could benefit from additional physical exercise therapy. As safety issues and age constraints do not seem to be a limiting factor, healthcare professionals could use the present findings to provide patients with a tailored intervention in terms of type of exercise, exercise time and duration of intervention period. We showed a positive dose–effect interaction for exercise time, indicating that longer exercise programs are better for mood improvement. Most studies included in our meta-analysis assessed supervised exercise. Therefore, our results cannot be generalized to unsupervised exercise.

**Implications for further research**

Given the purpose and transdiagnostic character of the present study, we chose to compare exercise intervention only to TAU control condition. Evaluation of any differential effects of other components of the interventions such as adherence, setting (e.g., home-based vs. gym-based), monitoring of exercise sessions with instruments (e.g., heart rate meters), cost-effectiveness and comparison with other control groups (e.g., active control conditions) is required to provide detailed recommendations on physical exercise interventions for the clinical practice.

**Strengths and limitations**

The greatest strength of the present study is that it provides an up-to-date and extensive quantitative overview of the literature regarding the efficacy of different exercise interventions in patients with chronic brain disorders. Second, our findings are largely in accordance with previous (quantitative) reviews that synthesized evidence on the efficacy of physical exercise in the studied brain disorders [20, 22, 24, 25, 28, 252]. However, in contrast to previous work, we performed both transdiagnostic and within-disorder analyses and evaluated the effect of several moderators providing evidence that physical exercise can be considered as an effective add-on and transdiagnostic treatment.

This study has some limitations. First, several studies could not be included in the cognitive meta-analyses, so that the overall effect of exercise on cognition was based on fewer studies than the other meta-analyses, making these findings more susceptible to change over time (when more studies become available). Notably, a recent RCT of 4-month aerobic and resistance exercise of moderate to high intensity added to usual care found that physical exercise did not slow cognitive decline in patients with mild-to-moderate dementia [18]. The authors measured global cognition with Alzheimer’s disease assessment scale-cognitive subscale (ADAS-cog) and found a small average difference with uncertain clinical relevance. This study did not fulfill the inclusion criteria of our study to be included in the quantitative review. However, considering the fact that we included four RCTs [65, 68, 74, 224] with negative outcomes of exercise on global cognition in AD (see Fig. 4) and did not find a significant overall effect of exercise on global cognition, we do not expect that adding this study would have changed our findings. Second, the analysis regarding the effect of physical exercise on depressive symptoms included studies with different disorders, and the included studies also differed in the severity of depression, ranging from mild depression to the presence of major depressive disorder. This might have biased the findings and resulted in a high effect size. However, both low and high effect sizes were found in mild and major depression, which suggests that physical exercise is effective for depressive symptoms in general, irrespective of the underlying severity. Third, publication bias is an important possible drawback in meta-analytical studies. Egger’s test showed potential publication bias for QoL and depressive symptoms. However, the fail-safe numbers of these tests were extremely large, increasing the validity of the results. Fourth, heterogeneity among studies was high, possibly due to combining studies with largely different interventions offered to different groups. However, heterogeneity values of the joint analysis were lower than the within-disorder heterogeneities (Tables 1, 2), indicating consistency in studies across disorders so that joint analysis
of disorders deemed sensible. Moreover, one of the main inter-study differing variables, age, did not affect the efficacy of exercise on the outcome measures. Besides, $Q$- and $I^2$-statistic cannot be used to estimate the magnitude of true dispersion [253]. Fifth, for all outcome measures, the risk of bias assessment indicated highest risk in terms of attrition. Incomplete outcome data and lack of ITT-analysis in studies could have biased the observed results. However, to account for possible attrition bias, we performed separate analyses on studies that performed ITT-analysis and thus had low risk of bias and studies with unclear risk of bias on ITT analysis (i.e., insufficient information to judge). These results showed even higher effects of exercise on QoL and depressive symptoms, while effects on cognition remained similar for the cognitive domain PS, but turned to non-significance for the cognitive domains A and WM, EF and M. The latter is likely due to the moderate to high heterogeneity among studies after inclusion of the study by [219]. Finally, we randomly selected six brain disorders of various etiology (e.g., neurodegenerative, neurodevelopmental, inflammatory) to demonstrate the generalizability of efficacy of exercise. Since we did not find any RCTs evaluating the effect of physical exercise in bipolar disorder, we decided to only include unipolar depression in the present study. Other brain disorders, such as epilepsy, traumatic brain injury and migraine have been investigated as well, but given restriction in time and capacity (as well as wordcount), this paper was confined to the chronic brain disorders summed above.

Conclusion

Additional therapy with physical exercise in patients with chronic brain disorders seems safe and has a medium-sized effect on QoL and a large beneficial effect on depressive symptoms, with a positive dose–response correlation. The evidence for the efficacy on cognition is small, but clinically relevant. Therefore, to improve the health status of patients with chronic brain disorders, add-on exercise therapy should be considered as an essential part of the treatment.

Acknowledgements The authors gratefully acknowledge the corresponding authors of the studies that provided additional information upon request. See Online Resource 3 for a detailed list of author names.

Author contributions MD, corresponding author, was involved in designing the study, literature search and data collection, performed data analysis, led manuscript preparation and discussion of the results, wrote the manuscript and prepared all the figures and tables and contributed to the manuscript revision. MJHB was involved in data analysis and contributed to the manuscript revision. MIES was involved in literature search and contributed to the manuscript revision. EHML was involved in literature search and contributed to the manuscript revision. PS contributed to the manuscript revision. IES was involved in designing the study, data analysis, discussion of the results, contributed to the manuscript revision and supervised the study. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding This study was partly supported by ZONMW TOP Grant 40-00812-98-13009.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

88. Danielsson L, Papoulia I, Peterson EL, Carlsson J, Waern M (2014) Exercise or basic body awareness therapy as add-on


Springer


244. Nair R, Kj M, Nk L (2016) Memory rehabilitation for people with multiple sclerosis (review) 3


Affiliations

Meenakshi Dauwan\textsuperscript{1,2,5} · Marieke J. H. Begemann\textsuperscript{1} · Margot I. E. Slot\textsuperscript{1} · Edwin H. M. Lee\textsuperscript{3} · Philip Scheltens\textsuperscript{4} · Iris E. C. Sommer\textsuperscript{1,5,6}

Marieke J. H. Begemann
m.j.h.begemann@umcutrecht.nl

Margot I. E. Slot
i.e.slot-3@umcutrecht.nl

Edwin H. M. Lee
edwinlhm@hku.hk

Philip Scheltens
p.scheltens@vumc.nl

Iris E. C. Sommer
i.e.c.sommer@umcg.nl

1 Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Postbus 85500, 3508 GA Utrecht, The Netherlands
2 Department of Clinical Neurophysiology and MEG Center, Amsterdam UMC, VU University Medical Center Amsterdam, Neuroscience Campus, Postbus 7057, 1007 MB Amsterdam, The Netherlands
3 Department of Psychiatry, 2/F, New Clinical Building, University of Hong Kong, Queen Mary Hospital, Hong Kong, China
4 Alzheimer Center and Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Center, Postbus 7057, 1007 MB Amsterdam, The Netherlands
5 Department of Biomedical Sciences of Cells and Systems Section, University of Groningen, University Medical Center Groningen, Neuroimaging Center 3111, Deusinglaan 2, 9713 AW Groningen, The Netherlands
6 Department of Biological and Medical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway