Review

The Effectiveness of Exercise Interventions for People with Parkinson’s Disease: A Systematic Review and Meta-Analysis

Victoria A. Goodwin, MSc,1* Suzanne H. Richards, PhD,1 Rod S. Taylor, PhD,1 Adrian H. Taylor, PhD,2 and John L. Campbell, MD1

1Primary Care Research Group, Peninsula Medical School, Exeter, United Kingdom
2School of Sport and Health Sciences, University of Exeter, United Kingdom

Abstract: Parkinson’s disease (PD) is a neurodegenerative disorder affecting the physical, psychological, social, and functional status of individuals. Exercise programs may be an effective strategy to delay or reverse functional decline for people with PD and a large body of empirical evidence has emerged in recent years. The objective is to systematically review randomized controlled trials (RCTs) reporting on the effectiveness of exercise interventions on outcomes (physical, psychological or social functioning, or quality of life) for people with PD. RCTs meeting the inclusion criteria were identified by systematic searching of electronic databases. Key data were extracted by two independent researchers. A mixed methods approach was undertaken using narrative, vote counting, and random effects meta-analysis methods. Fourteen RCTs were included and the methodological quality of most studies was moderate. Evidence supported exercise as being beneficial with regards to physical functioning, health-related quality of life, strength, balance and gait speed for people with PD. There was insufficient evidence support or refute the value of exercise in reducing falls or depression. This review found evidence of the potential benefits of exercise for people with PD, although further good quality research is needed. Questions remain around the optimal content of exercise interventions (dosing, component exercises) at different stages of the disease. © 2008 Movement Disorder Society

Key words: Parkinson’s disease; exercise; systematic review; meta-analysis.

Exercise is a planned, structured physical activity which aims to improve one or more aspects of physical fitness.1 Current models of rehabilitation often use compensatory strategies as the basis of therapeutic management. However, there is a growing body of evidence regarding the benefits of exercise in terms of neuroplasticity and the ability of the brain to self repair.2 Animal models have found that exercise has protective benefits against the onset of symptoms in Parkinson’s disease (PD).3 This appears to be due to the release of neurotrophic factors, and greater cerebral oxygenation, which together promote new cell growth and cell survival.4,5 In PD, it has been found that exercise stimulates dopamine synthesis in remaining dopaminergic cells and thus reducing symptoms.6 Fox et al.3 suggest there are five key principles of exercise that enhance neuroplasticity in relation to PD, these being: (a) intensive activity maximizes synaptic plasticity; (b) complex activities promote greater structural adaptation; (c) activities that are rewarding increase dopamine levels and therefore promote learning/relearning; (d) dopaminergic neurones are highly responsive to exercise and inactivity (“use it or lose it”); (e) where exercise is introduced at an early stage of the disease, progression can be slowed.

It has been well documented that physical activity levels decline with advancing age and these reductions contribute to functional decline.7 People with PD have been shown to reduce levels of physical activity more
quickly than their healthy peers and have lower levels of strength and functional ability. \(^9,10\) However, the observation of muscle weakness is not simply a secondary consequence of ageing and inactivity, but also a primary symptom of PD.\(^11\) This is due to impaired basal ganglia having an inadequate effect on the cortical motor centers which in turn lead to less activation of motor neurones and therefore muscle weakness.\(^8,10,12\) This mechanism also contributes to impaired balance, falls, and disability.\(^13\) People with PD are three times more likely to sustain a hip fracture as a result of a fall when compared to those without the condition.\(^14,15\)

A number of systematic reviews and a meta-analysis have been undertaken to investigate the efficacy of physiotherapy among people with PD. The earlier reviews, with literature searches up until 1999\(^19\) and 2000,\(^16,17\) evaluated a range of physiotherapeutic techniques including nonexercise interventions, such as sensory cueing and behavioral therapy, in addition to exercise strategies. The Cochrane reviews\(^16,17\) were limited to randomized controlled trials and reported that they were unable to combine the studies for meta-analysis given the clinical and methodological heterogeneity. The studies were not conclusive in respect to the physiotherapy management of people with PD. De Goede et al\(^19\) included studies that adopted a quasi-experimental design. They reported significant benefits in respect of activities of daily living, walking speed, and stride length. However, only one study in each of these domains was a randomized controlled trial evaluating an exercise intervention, with the others using less rigorous study designs or other physiotherapy techniques.

A recent review of physiotherapy for people with PD (literature searches to June 2006) concluded that there were positive benefits associated with gait, transfers, balance and functional ability.\(^18\) However, by including a range of management strategies, such as exercise and cueing, under the umbrella of “physiotherapy” it is difficult to extract information regarding the contribution of individual therapeutic components. Lim et al.\(^20\) for example, reviewed the literature on the effects of cueing on gait in people with PD and reported that while auditory cueing may be beneficial to gait speed, no conclusions could be made as to the effects of visual cueing due to a lack of evidence. Nieuwboer et al.\(^21\) recently reported significant improvements in gait and balance using external cueing devices and recommended cueing training as an adjunct to gait management.

This systematic review examines the potential benefit of exercise interventions for people with PD, focusing specifically on evaluations adopting experimental, randomized designs. We did not include studies explicitly evaluating cueing strategies as cueing itself is not an exercise but an external temporal or spatial stimulation to facilitate gait.\(^21\)

**PATIENTS AND METHODS**

**Search Strategy**

One researcher undertook the initial literature search, scanning abstracts to identify eligible studies. If it was unclear as to whether the study met the selection criteria, advice was sought from a second researcher and a consensus opinion made. The following electronic databases were searched: Cinahl (1982 to Dec 2006); Embase (1974 to Dec 2006); Allied and Complementary Medicine AMED (1985 to Dec 2006); PubMed (1980 to Dec 2006); SPORTDiscus (1980 to Dec 2006); and Cochrane Library (1980 to Dec 2006). Literature was also identified by citation tracking using reference lists from papers and Internet searching. The following keywords were used in combinations: Parkinson’s disease, Parkinsonism, exercise, physical activity, physical therapy.

**Inclusion Criteria**

A study was included if it met the following criteria:

i. A randomized controlled trial methodology was used;
ii. The target population was people with PD;
iii. The effects of an exercise/physical activity intervention were compared with any comparator, including other forms of exercise/physical activity;
iv. The outcomes included at least one of the following: physical performance or functioning, falls, or health-related quality of life;
v. The paper was available in English.

**Exclusion Criteria**

A study was excluded if:

i. The effects of a nonexercise intervention were evaluated (examples include behavioral interventions, cueing strategies, music therapy)
ii. The paper did not report outcomes for the first assessment period (cross-over studies only) so as to prevent any bias of carry over or order effects.\(^22\)

**Data Extraction and Quality Assessment**

Key data were independently extracted from the identified papers by two researchers using a structured form.\(^23\) Data extraction forms included the key components of general study information (title, author, and country of study), study characteristics (population data, intervention, comparator and outcomes) and findings,
including length of follow up. The quality of each study was assessed in terms of (i) internal validity: randomization had been appropriately carried out and concealed; presence of blinding; reporting losses to follow up, and the use of intention to treat analysis; (ii) external validity: reporting inclusion and exclusion criteria; and (iii) study power: reporting of sample size calculation. A numerical quality score was calculated using a modified Jadad scale. This tool awards one point for each of the following: (a) being described as randomized; (b) random allocation concealed from clinician/care provider; (c) appropriate blinding of outcome assessment; and (d) describing withdrawals. A score of one was considered to be low quality; two or three was considered moderate; and a score of four was considered to reflect a high quality trial. Any discrepancies in data extraction or quality assessment were resolved by reference to the original paper and discussion between the researchers.

Data Analysis and Synthesis

A mixed methods approach was undertaken. The principle approach to data synthesis was a narrative review of the results supplemented by vote counting. Vote counting is a method of reporting study outcomes used to synthesize the results whereby all outcomes relating to each study were listed and the direction of effect for each one was identified. Where a statistically significant difference was reported in favor of the intervention it was noted as positive, and if in favor of the control it was recorded as negative. Where no significant difference was found between groups, an equivocal rating was noted.

Meta-analysis was undertaken using STATA Version 8 (Stata Corp, College Station, TX). For two of the most commonly reported outcome domains (physical functioning and health-related quality of life) a standardized effect size was calculated for each study and expressed in standard deviation units. Given the methodological diversity between studies, $\chi^2$ test for heterogeneity was applied and data pooled using random effects meta-analysis using the DerSimonian and Laird method.

RESULTS

The search results can be seen in Figure 1. One study was excluded as it did not report the first assessment period and randomized the order of four different interventions over four consecutive days, with assessments at the end of each day. However, the authors reported the outcomes for each intervention over the whole study rather than at the end of each day and did not take into account the potential accumulative effect by allowing a “washout” period between treatments. The quality of studies is summarized in Table 1 and the individual study characteristics in Table 2.

Methodological Quality

Methodological details reported in the papers were varied and often poor (Table 1). Only two studies reported power and sample size calculations. While two studies did not clearly report participant selection criteria, the criteria reported in the other studies were varied, although all included a diagnosis of PD and being medically stable. Similarly, participant exclusion criteria also varied between the studies. One study was unclear in respect of reporting the randomization procedure, describing controls as “matched” (abstract), but then stating that controls were randomly allocated in the methods text. Only three papers explicitly reported concealment of randomization and we cannot discount the possibility of selection bias in the other papers. With regards to assessor blinding, all but four studies reported that assessors were blinded to participant allocation. Ashburn et al., however, reported that by the six month follow up assessment the assessor was aware of allocation for 39% of those in the intervention group and 17% of the controls. Withdrawals were not described in two studies although both papers reported that analysis was undertaken on an intention to treat basis. Two studies were found to be of high quality, 10 of moderate quality and two of low quality.

Participants

The characteristics of included studies are described in Table 2. A total of 495 participants contributed to the
studies reported in this review. The minimum number of participants in a study was 11 and the maximum was 142. Of the studies reporting the sex of participants 284/423 (67%) were male. PD status was described in all but one study using Hoehn and Yahr’s measure of disease severity. All studies included participants at stages 2 and 3 with the exception of Hirsch who included those at stages 1 and 2 only. Three studies included participants at stage 4 of the disease.

### Intervention

Four studies failed to report a rationale behind the development of the intervention, although those that did varied in the level of detail. This was noted in particular in the two studies evaluating Qigong, a form of Chinese physiotherapy and exercise. Five studies compared the exercise intervention with no intervention or usual care, two of the studies evaluated an exercise intervention against a nonexercise control, while four compared two different exercise interventions and one study compared three different exercise interventions. Two studies did not report details of the comparison group.

Across all of the studies, the interventions were clinically heterogeneous with regards to the type of exercise and to the frequency and duration of exercise being undertaken (between 6 and 36 hours spread over 4–12 weeks). All except four studies reported exercise interventions being delivered by physiotherapists. Palmer et al. compared flexibility exercises delivered by a “corrective therapist” with seated karate delivered by a student nurse with a black belt in karate, Hirsch et al. used a trained exercise leader, and Schmitz-Hubsch et al. used a Qigong teacher to deliver the intervention. All interventions took place within an outpatient setting except Ashburn et al. who implemented a home based intervention and Hirsch who utilized a leisure setting. One study did not report details of who delivered the intervention or the setting. Six studies used a group intervention while the other eight used an individual approach.

### Outcomes

The results of the vote counting method of data synthesis have been summarized and presented in narrative form. Most studies evaluated the short-term effect of interventions with only four studies monitoring patients over the longer term of 6 months or more. Nine studies reassessed outcomes immediately post intervention and again at a later date to observe any detraining effect.
<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Mean age (SD)</th>
<th>Gender</th>
<th>Disease stage</th>
<th>Intervention (hours/per week/weeks)</th>
<th>Control (hours/per week/weeks)</th>
<th>Follow up (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmer et al.</td>
<td>14</td>
<td>C = 6.59 (7.2)</td>
<td>12M</td>
<td>2–4</td>
<td>Stretching (1/3/12)</td>
<td>Seated karate (1/3/12)</td>
<td>12</td>
</tr>
<tr>
<td>Comella et al.</td>
<td>18</td>
<td>66 (8)</td>
<td>Incomplete data</td>
<td>2–3</td>
<td>Progressive exercise training (1/3/4)</td>
<td>Wait list control for 6 months then intervention</td>
<td>26</td>
</tr>
<tr>
<td>Bridgewater and Sharpe</td>
<td>26</td>
<td>C = 67.3 (3.9)</td>
<td>12M</td>
<td>1–3</td>
<td>Aerobic exercise (0.75/2/12)</td>
<td>Interest talks (0.75/1/4)</td>
<td>16</td>
</tr>
<tr>
<td>Bridgewater and Sharpe</td>
<td>26</td>
<td>C = 65.9 (10.2)</td>
<td>12M</td>
<td>1–3</td>
<td>Trunk strength and aerobic exercise (0.75/2/12)</td>
<td>Interest talks (0.75/1/4)</td>
<td>16</td>
</tr>
<tr>
<td>Schenkman et al.</td>
<td>51</td>
<td>Not specified</td>
<td>Incomplete data</td>
<td>2–3</td>
<td>Relaxation and muscle activation (1/2/10)</td>
<td>No intervention</td>
<td>10</td>
</tr>
<tr>
<td>Toole et al.</td>
<td>11</td>
<td>C = 71</td>
<td>12M</td>
<td>1–3</td>
<td>Strength and balance training (1/5/10)</td>
<td>Not reported</td>
<td>10</td>
</tr>
<tr>
<td>Miyai et al.</td>
<td>24</td>
<td>C = 69.8 (1.5)</td>
<td>12M</td>
<td>2.5–3</td>
<td>BWSTT (0.75/3/12)</td>
<td>Physiotherapy (0.75/3/12)</td>
<td>26</td>
</tr>
<tr>
<td>Hirsch et al.</td>
<td>15</td>
<td>C = 75.7 (1.8)</td>
<td>12M</td>
<td>1–2</td>
<td>Balance and strength training (0.75/5/1.2)</td>
<td>Balance training (0.5/3/12)</td>
<td>16</td>
</tr>
<tr>
<td>Toole et al.</td>
<td>23</td>
<td>74.5 (9.7)</td>
<td>19M</td>
<td>1–4</td>
<td>Unweighted or weighted treadmill walking (0.33/3/6)</td>
<td>Treadmill walking 0.33/3/6</td>
<td>10</td>
</tr>
<tr>
<td>Protas et al.</td>
<td>18</td>
<td>C = 73.7 (8.5)</td>
<td>18M</td>
<td>2–3</td>
<td>Gait training and BWSTT (varied/3/8)</td>
<td>Not described</td>
<td>12</td>
</tr>
<tr>
<td>Ellis et al.</td>
<td>68</td>
<td>Group A = 64 (8.4)</td>
<td>51M</td>
<td>2–3</td>
<td>Group A Physiotherapy (1.5/2/6) and medication + medication only (6 weeks)</td>
<td>Group B Medication only (6 weeks) + Physiotherapy (1.5/2/6) and medication</td>
<td>24</td>
</tr>
<tr>
<td>Schmitz-Hubsch et al.</td>
<td>56</td>
<td>C = 63 (8)</td>
<td>43M</td>
<td>Not reported</td>
<td>Qigong (1/1/8) then repeated after 8 week rest</td>
<td>No intervention</td>
<td>52</td>
</tr>
<tr>
<td>Burini et al.</td>
<td>26</td>
<td>I = 65.7 (7)</td>
<td>9M</td>
<td>2–3</td>
<td>Aerobic training 0.75/3/7 then 2 month rest then Qigong 0.75/3/7</td>
<td>Reverse order of intervention group</td>
<td>22</td>
</tr>
<tr>
<td>Ashburn et al.</td>
<td>142</td>
<td>C = 71.6 (8.8)</td>
<td>86M</td>
<td>2–4</td>
<td>Home based physiotherapy (1/1/6)</td>
<td>Usual care</td>
<td>26</td>
</tr>
</tbody>
</table>

C, control group; I, Intervention group; M, Male; BWSTT, Body weight supported treadmill training.
Physical Functioning

Nine studies reported findings across three outcomes measuring physical functioning including the Unified Parkinson’s Disease Rating Scale (UPDRS), the North Western University Disability Scale (NUDS), and the Self Assessment Parkinson’s disease Disability Scale (SAS). The latter is also known as Brown’s Disability Scale (BDS). Each of the tools is disease specific to PD. Of these, four out of nine studies reported a statistically significant benefit in favor of the exercise intervention.

Seven of the nine studies (n = 360 participants) reported sufficient data to enable extraction relating to physical functioning (see Fig. 2). Burini et al. used two physical function outcomes (UPDRS and BDS), but only one (UPDRS) was included in the meta-analysis. It was notable that the meta-analysis found a different result to the ones reported in the papers in three studies. Burini et al. and Miyai et al. reported an equivocal outcome but the meta-analysis found a significant benefit in the direction of the exercise intervention. However, Schmitz-Hubsch et al. (2005) reported a significant benefit in terms of disability, but the meta-analysis did not support this difference. One study reported a variance that was not commensurate with standard deviation (SD). We therefore used two methods of imputation to derive a SD for this study (from the P value reported in the paper and using formulae calculating SD from a standard error). The overall pooled result was insensitive to the method of imputation. Pooled data identified evidence of improvement in physical functioning with exercise (mean SMD 0.47, 95% CI 0.12 to 0.82) (see Fig. 2). The studies were found to be statistically heterogeneous ($\chi^2 = 12.89$ (d.f. = 6), $P = 0.045$).

Health-Related Quality of Life

Four studies (n = 292 participants) reported findings across three quality of life outcomes including the Sickness Impact Profile (SIP-68), the Parkinson’s disease questionnaire (PDQ-39), and the EuroQOL (EQ-5D). Of these, only one reported a statistically significant benefit in favor of the exercise intervention group.

We were able to extract relevant data from all four papers relating to health-related quality of life (HRQOL) and data suggest that exercise interventions are likely to result in improvements in HRQOL (mean SMD 0.27, 95% CI 0.04 to 0.51) (see Fig. 3). Testing for heterogeneity was statistically insignificant ($\chi^2 = 0.43$ (d.f. = 3), $P = 0.93$).
Strength

Muscle strength was reported in four studies.29,33,35,36 A significant improvement in leg muscle strength was reported by Hirsch et al.35 and Toole et al.33 Palmer et al.29 did not report between group differences although did report some improvements in grip strength in both the karate and stretching groups.

Balance

Five studies32,33,35,36,41 reported balance as an outcome using three tools (Berg Balance Scale, Functional Reach and Sensory Orientation Test-SOT). A significant improvement in balance was reported in four of the five of the studies; however, Toole et al.36 reported a favorable improvement on the SOT but not on the Berg Balance Scale, and Ashburn et al.41 found an improvement in functional reach but not in the Berg Balance Scale.

Gait

Four studies34,36-38 reported outcomes relating to gait with three studies reporting a significant improvement in walking speed following an exercise intervention.

Depression

Four papers reported depression as an outcome measure30,31,39,40 using four different tools (Geriatric Depression Scale, Levine Pilowsky Depression Questionnaire, Montgomery-Asperg Depression Rating Scale, and the Beck Depression Inventory). No study reported a significant improvement in depression as a result of exercise.

FIG. 3. Meta-analysis for exercise and health-related quality of life (random effects model).
Adverse Events

Adverse events or side effects of the interventions were not generally reported. One study reported that no falls occurred during the implementation of the intervention and another reported data on injuries occurring during strength testing.

DISCUSSION

The aim of this systematic review was to evaluate the effectiveness of exercise interventions in randomized controlled trials undertaken with people with PD. Our study supports and updates the findings of previous reviews, and, through refining our scope to one aspect of physiotherapy (i.e., exercise-based interventions) we have identified that exercise is of benefit to people with PD in respect of physical functioning, HRQOL, strength, balance and gait speed. Our findings add to the growing body of evidence regarding the effectiveness of physiotherapy for people with PD. There is currently insufficient evidence to support or refute the value of exercise in reducing falls or depression, or its safety with people with PD. We however acknowledge that there is some potential for publication bias as we limited our selection criteria to those studies available in English. We also recognize that there may be confounding from other physiotherapy techniques, such as cueing, as it is difficult to control for this when delivering complex, multi-faceted interventions.

Most studies provided an inadequate description of their methods to allow a full assessment of their methodological quality. Where details reported were available, most studies were found to be of moderate quality using the modified Jadad score. Deane et al. had reported that many studies were of poor methodological quality and had small participant numbers when reviewing the effectiveness of physiotherapy techniques (which may include exercise) in people with PD, suggesting that methodological quality can be an issue in studies of this type. Given the generally relatively small sample size of most trials, a lack of statistically significant difference between groups may simply reflect a lack of statistical power rather than the absence of a real lack of difference. However, our pooling of the results for some outcomes across studies allowed us to at least partly overcome the criticism of inadequate power.

The participants were mainly men although the prevalence of PD is said to be similar for men and women suggesting that males are somewhat over-represented in these studies. Similarly no studies reported the ethnicity of participants. These raise the questions as to whether these interventions are acceptable to women with PD and the generalizability of the results.

The failure to report a clear rationale behind the development of the intervention may contribute to some of the equivocal findings. It is essential that a “complex intervention” such as exercise training have a theoretical basis in order to inform the design of a study. The interventions described in this review were often short in duration with six studies providing an intervention of 8 weeks or less in duration. This dose of exercise may be insufficient to significantly affect the outcomes. Some studies described their intervention as “physiotherapy” which may be considered by some to not be a form of exercise. However in these studies the authors describe in more detail the content of the interventions which utilize exercise as the main component supporting their inclusion in the review.

The studies reviewed in this paper were comparable, in that they targeted the same population (people with PD) using exercise as an intervention and reported outcomes that displayed some similarities, although the length of follow-up varied widely. Most of the studies assessed outcomes at three time points in order to establish any detraining effects after the intervention period had ceased. This is an important factor in clinical practice. Even though we attempted to tighten the focus of this review, the degree to which the studies are clinically and methodologically homogeneous remains debatable. Although the test for statistical heterogeneity was not significant for the papers reporting HRQOL it cannot be assumed that they are homogenous. Vote counting was used to supplement the narrative and meta-analyses and to synthesize the results of the included studies given their substantial heterogeneity in outcome reporting. The method provides an overall summary of direction of effect although it does not consider the magnitude of the effect size and the precision of the estimated effects. However, in this review it did provide an approach for summarizing the effect of exercise reported across all studies whereas meta-analysis (which formally takes into account both directionality and precision of studies) could only be performed on a proportion of the studies and outcomes.

In three of the seven studies reporting physical functioning as an outcome, and in three of the four studies reporting HRQOL, we identified a discrepancy between the author-reported results and the results we generated in the meta-analysis. In some studies, we reported a significant effect size derived from the random effects model when the individual study had reported equivocal findings. This may be due to the model awarding relatively more weight to smaller studies thus effectively increasing the power to detect significant changes in key outcomes in individual studies. Conversely, we found...
an equivocal outcome when the study had reported a significant improvement\(^9\); a lack of assessor blinding may have contributed to detection bias and an exaggeration in effect size. However, as vote counting takes into account the only the direction of effect and not the size of effect, the results of the meta-analyses are considered to be superior.\(^{20}\) The meta-analyses provide support for exercise as an effective intervention for improving physical functioning and HRQOL for people with PD, but the generalizability of these positive meta-analysis results should be interpreted with some caution.

**Implications**

We have found exercise to be effective at improving physical functioning and HRQOL, leg strength, balance, and walking but there is currently insufficient evidence with regards effectiveness in the areas of falls prevention and the management of depression. Future research needs to establish what elements constitute an optimal exercise intervention for people with PD such as the dosage, component parts of intervention, and the targeted stage of the disease. This is of particular importance given the deteriorating nature of this condition. In addition, researchers need to provide a theory-driven rationale for the development of their intervention, ensure studies are adequately powered with a sample size sufficient to be able to detect a statistically significant difference, and report their findings, in accordance with currently internationally agreed standards such as CONSORT.\(^{47}\) It is also important that study populations reflect the general PD population in terms of gender and ethnicity in order to support the generalizability of findings.

**Acknowledgments:** This work was supported initially by the Clinical Academic Fellowship Scheme funded by Devon and Cornwall Workforce Development Confederation and Plymouth Teaching Primary Care Trust. Subsequently, VG was funded by a Personal Award Scheme Researcher Development Award from the National Institute of Health Research (Department of Health).

**REFERENCES**


