Title: Walking Exercise for Chronic Musculoskeletal Pain: Systematic Review and Meta-Analysis

Seán R O'Connor PhD, Mark A Tully PhD, Brigid Ryan BSc (Hons), Chris M Bleakley PhD, George D Baxter MPhil, Judy M Bradley PhD, Suzanne M McDonough PhD

1 Centre for Public Health, Queens’ University Belfast, United Kingdom
2 UKCRC Centre of Excellence for Public Health (Northern Ireland), United Kingdom
3 Centre for Physiotherapy Research, University of Otago, New Zealand
4 Sport and Exercise Sciences Research Institute, University of Ulster, United Kingdom
5 Institute of Nursing and Health Research, University of Ulster, United Kingdom

Acknowledgment of financial support
This study was funded by a PhD award from the Department of Employment and Learning, Northern Ireland, UK

Conflicts of interest
None declared

* Corresponding author:
Professor Suzanne M McDonough, Room 01F118, School of Health Sciences, University of Ulster, Shore Road, Newtownabbey, Co. Antrim, BT37 0QB, United Kingdom
e-mail: s.mcdonough@ulster.ac.uk; Telephone: 00 44 2890 366459
Walking for Musculoskeletal Pain: Review

ABSTRACT

Objective: To systematically review the evidence examining effects of walking interventions on pain and self-reported function in individuals with chronic musculoskeletal pain.

Data Sources: Six electronic databases (Medline, CINAHL, PsychINFO, PEDro, SportDiscus and the Cochrane Central Register of Controlled Trials) were searched from January 1980 up to March 2014.

Study Selection: Randomized and quasi-randomized controlled trials in adults with chronic low back pain, osteoarthritis or fibromyalgia comparing walking interventions to a non-exercise or non-walking exercise control group.

Data Extraction: Data were independently extracted using a standardized form. Methodological quality was assessed using the United States Preventative Services Task Force (USPSTF) system.

Data Synthesis: Twenty-six studies (2384 participants) were included and suitable data from 17 were pooled for meta-analysis with a random effects model used to calculate between group mean differences and 95% confidence intervals. Data were analyzed according to length of follow-up (short-term: ≤8 weeks post randomization; medium-term: >2 months - 12 months; long-term: > 12 months). Interventions were associated with small to moderate improvements in pain at short (mean difference (MD) -5.31, 95% confidence interval (95% CI) -8.06 to -2.56) and medium-term follow-up (MD -7.92, 95% CI -12.37 to -3.48). Improvements in function were observed at short (MD -6.47, 95% CI -12.00 to -0.95), medium (MD -9.31, 95% CI -14.00 to -4.61) and long-term follow-up (MD -5.22, 95% CI 7.21 to -3.23).
Walking for Musculoskeletal Pain: Review

Conclusions: Evidence of fair methodological quality suggests that walking is associated with significant improvements in outcome compared to control interventions but longer-term effectiveness is uncertain. Using the USPSTF system, walking can be recommended as an effective form of exercise or activity for individuals with chronic musculoskeletal pain but should be supplemented with strategies aimed at maintaining participation. Further work is also required examining effects on important health related outcomes in this population in robustly designed studies.

Key words: Meta-analysis, walking, exercise, chronic musculoskeletal pain.
Walking for Musculoskeletal Pain: Review

Chronic musculoskeletal pain (CMP) is a major cause of morbidity. Given the changing age profile of the population it is probable that its prevalence and associated costs will continue to rise. Chronic low back pain (CLBP), osteoarthritis (OA) and fibromyalgia syndrome (FMS) are reported as being among the most common types of musculoskeletal disorder. These conditions may be associated with significant functional limitations. There is also evidence that they can exert a substantial influence on long-term health status and overall quality of life.

Current treatment recommendations support various non-pharmacological interventions, including aerobic exercise, in order to reduce pain and maintain or increase functional status. However, randomized controlled trials have tended to report only short-term improvements in outcome with relatively small effect sizes. This may be due to a number of factors, including heterogeneity of interventions.

Walking may represent an ideal form of aerobic activity, due to its ease of accessibility and relatively low impact. It has a low risk of musculoskeletal injury and is considered safe to recommend for previously sedentary individuals. Low to moderate intensity walking (described as exercising at a MET value of between 3-4 or a pace that results in an increased respiratory and heart rate, but where the individual can still carry out a conversation) has been shown to lead to improvements in aerobic capacity, body mass index, systolic/diastolic blood pressures, triglyceride, and high density lipoprotein cholesterol levels in healthy sedentary individuals, as well as in those with established cardiovascular disease and type 2 diabetes.
Walking for Musculoskeletal Pain: Review

Although it is widely recommended, there is currently limited evidence relating to the effectiveness of walking exercise for management of musculoskeletal disorders. (17)

The aim of this systematic review was to examine the effects of walking interventions on pain and self-reported function in adults with CMP.

METHODS

Data sources, searches and extraction
Comprehensive search strategies were carried out by at least two independent reviewers according to the preferred reporting items for systematic reviews and meta-analyses (PRISMA) recommendations and those of the Cochrane Musculoskeletal Review Group. (18,19) A review protocol was developed 'a priori' using the PICOS framework to define the research question and inclusion criteria. Six electronic databases (Medline, CINAHL, PsychINFO, PEDro, Sport Discus and the Cochrane Central Register of Controlled Trials) were searched for relevant papers published between January 1980 and March 2014 using combinations of key terms which included “walking”, “aerobic exercise”, “musculoskeletal pain”, “low back pain”, “arthritis” and “fibromyalgia” (A full list of the MeSH terms used is included in Supplementary data: Appendix A). Reference lists of included articles and key systematic reviews were also checked by hand.

All randomized or quasi-randomized studies published in full were considered for inclusion. No language restrictions were applied. Studies were required to include
Walking for Musculoskeletal Pain: Review

adults aged 18 years or over, with a diagnosis of CLBP, OA or FMS made according to clinical judgement or accepted diagnostic criteria.(6,20,21)

All land or treadmill based walking interventions were considered for inclusion. Studies were required to include a comparative non-exercise or non-walking exercise control group. Those including any form of assisted walking were excluded. Studies were also excluded if they involved peri-operative or post-operative interventions. Primary outcomes of interest were pain and self-reported function.

At least two reviewers independently examined titles and abstracts of identified studies. Full text copies of potentially eligible studies were assessed to determine whether walking formed at least half of the overall intervention. Final inclusion was determined by consensus between review authors. Data were extracted independently using a standardised form. Disagreements were resolved by consensus and involved a third author if required. Intervention and control group sample size, plus mean and standard deviation (SD) values for pain and function were extracted. Where the SD was not provided it was calculated from the standard error (SE) or 95% confidence intervals (95% CI). Where tabulated results were not presented, an attempt was made to extract data from graphs. All data were cross checked by a second author. For the purposes of comparability, outcomes were converted to a 0-100 scale (with higher scores indicating greater pain or functional limitation).

Assessment of methodological quality and adequacy of exercise interventions

The United States Preventative Services Task Force (USPSTF) system was used to assess methodological quality and form treatment recommendations based on an
Walking for Musculoskeletal Pain: Review

estimate of net benefit and the overall strength of evidence.(22) Internal validity and external validity were rated as “good”, “fair”, or “poor” according to pre-defined criteria specific to the study design.(23) (See supplementary data: Appendix B). Studies rated as “good” met all relevant criteria. Fair studies did not meet all criteria while “poor” studies were judged to contain a serious methodological flaw. Individual studies were given an overall rating, with internal and external validity considered to have equal weighting. Included studies were also screened for statements indicating sources of funding or support. Reviewers were not blinded with regards to study authors, institution, or journal of publication. All final decisions regarding quality assessment and overall recommendations were reached by consensus. Studies were also scrutinized independently to determine if the interventions met American College of Sports Medicine (ACSM) guidelines for the quantity and quality of aerobic exercise in inactive individuals based on frequency, intensity, timing, mode and duration of interventions.(24)

Data synthesis and analysis
The meta-analysis compared mean values for pain and function between walking intervention and control groups. To avoid double counting, where multiple treatment groups were included walking was compared only to minimal intervention controls. Suitable studies were considered to be clinically homogeneous on the basis of similarities in participant demographics and intervention methods. These data were pooled and analyzed using RevMan (v.5.2.8).(25) Statistical heterogeneity was assessed using the $\chi^2$ and $I^2$ test statistics. Where the P value was less than 0.05 or the $I^2$ value greater than 50%, indicating large heterogeneity,(26) a random effects model for inverse variance was used to calculate the mean difference and 95%CI.
Formal statistical tests were not used to assess publication bias, which was evaluated using visual assessment of funnel plots. Data were analyzed by length of follow-up which was categorized as short (≤8 weeks post randomization), medium (2-12 months) or long-term (>12 months). Sensitivity analyses were carried out excluding studies where walking was combined with a co-intervention.

Nine articles were not included in the meta-analysis for the following reasons: no validated self-reported measure of pain or function (27,28) (one study used a functional scale that contained additional questions related to global health status and these data were therefore not included); unadjusted baseline differences between groups; (29,30) presented median data only; (31,32) change over time only (33) or did not include a measure of variability. (34) One study reported pain as an outcome but did not include these data in the paper. (35)

RESULTS

Description of studies

The electronic database searches revealed a total of 2760 articles after exclusion of duplicates. Thirty seven of these met the inclusion criteria (see Appendix C for a list of excluded studies). Eleven were reports of follow-up data or sub-sample analyses. There were therefore 26 original studies in the review including a total of 2384 participants (Mean: 93) with an average age of 57 years (SD: 15), of whom 77% were female. The complete selection process, including reasons for exclusion is shown in Figure 1.
Walking for Musculoskeletal Pain: Review

Twenty four of the studies were randomized controlled trials. Twelve provided data for OA (27-29,31,36-43), eight for FMS (30,33-35,44-47), five for CLBP (32,48-51) and one included participants with chronic hip, lower back or knee pain.(52) Demographic details and study characteristics are summarized in Table 1 and Table 2.

In the majority of interventions (19/26, 73%) walking was supervised in a hospital clinic, gymnasium or other setting (Table 2). Some studies combined supervised walking with instructions to walk at home;(31,37,40) six were home-based only.(28,30,32,38,43,51) Three used pedometers to assist with step-based walking goals (28,43,51) while three used time-based walking goals.(30,32,38)

Thirteen studies included a walking only intervention group. The remaining combined walking with a co-intervention. The most common of which were educational interventions or alternative forms of exercise (Table 2). A range of controls were used including education; usual care; alternative forms of exercise; a passive intervention (relaxation/massage) and a 6-8 week pre-intervention baseline phase. Mean length of final follow-up was 1.8 months (SD: 0.4) for studies with short term outcomes (≤8 weeks post randomization); 4.9 months (SD: 1.9) for medium-term outcomes (>2-12 months); and 18.4 months (SD: 7.6) for long-term outcomes (>12 months).

Eleven studies included a statement of associated adverse events. These included two falls resulting in distal radial fractures, one fall resulting in a hip fracture, one case of plantar fasciitis and two cases of allergic skin reactions to metal pedometer clips. Two studies including participants with fibromyalgia reported a general increase in reporting of pain and muscle stiffness in the intervention group. One study including participants
Walking for Musculoskeletal Pain: Review

with CLBP reported temporary exacerbations in pain levels in a small number of participants which was attributed to unaccustomed activity levels.

Methodological quality and exercise interventions

Overall, the included evidence was judged to be of at least fair methodological quality (Supplementary data: Appendix B). Six studies met all criteria for internal validity and were rated as good. (32,33,37,47,48,51) A small number of studies (n=5) contained serious potential sources of methodological bias and were therefore rated as poor. (27,28,30,31,43) This was as a result of inadequate allocation concealment during randomization, (28,30) unequal distribution of important confounding variables at baseline not accounted for during analysis, (27,30) no masking of outcome assessment, (31) or due to a substantial (>50%) drop-out rate and subsequent post hoc revision of the intervention groups examined. (29) For external validity most studies were rated as fair, with nine rated as good. (32,33,36,38,42,43,47,50,51) Studies generally included similar populations in terms of demographics and clinical presentation, as well as interventions that would be routinely available or feasible in clinical practice. Visual assessment of funnel plots indicated that there was no substantial evidence of publication bias. Only one study (27) did not include a statement indicating sources of funding or support. Ten studies (35-37,40,41,44-46,49,50) included interventions that met all ACSM criteria. (24) (Supplementary data: Appendix B) While the majority met minimum criteria for frequency of exercise and length of intervention, eleven either did not provide enough detail regarding exercise intensity, or it was not sufficient to effect any change in fitness. Eleven of the 26 studies (32,33,36,37,39,40,41,43,47,51,52) reported a measure of participant adherence (Table 2). These included attendance at exercise classes (n=7), self-
Walking for Musculoskeletal Pain: Review

reported completion of home exercise (n=2) or self-reported adherence to wearing a pedometer (n=2).
Walking for Musculoskeletal Pain: Review

**Meta-analysis**

Data from 17 of the included studies were suitable for meta-analysis. Although applying the alternative fixed effects model did not substantially alter any analyses, data are presented here using the more conservative random effects model. Analysis revealed significant differences in favour of walking interventions in terms of reduced pain at short (mean difference (MD) -5.31, 95% confidence interval (95% CI) -8.06 to -2.56) and medium-term follow-up (MD -7.92, 95% CI -12.37 to -3.48). No effect on pain was observed for long-term data (MD -2.22, 95% CI -6.03 to 1.59) (Figure 2). For self-reported function, improvements were found at short-term (MD -6.47, 95% CI -12.00 to -0.95), medium (MD -9.31, 95% CI -14.00 to -4.61) and long-term follow-up (MD -5.22, 95% CI -7.21 to -3.23) (Figure 3). Sensitivity analyses excluding studies which combined walking with a co-intervention did not alter overall results.

**DISCUSSION**

Overall findings indicated that walking interventions were associated with significant improvements in both pain and self-reported function in individuals with CMP. While effects appeared to be maintained beyond the immediate post-intervention period, only differences in function were observed at long-term follow-up. This was based primarily on data derived from interventions lasting for between six and 12 months. It is therefore unlikely that improvements in outcome would be maintained following the shorter intervention periods included in the majority of other interventions. This is supported by additional sub-sample data from one included study which indicated that
Walking for Musculoskeletal Pain: Review

significant improvements in outcome following an eight week intervention were absent at 12 months. (53)

While it has been suggested that supervised interventions may be required to maintain adherence with exercise, (7) other techniques, including those encouraging self-management, may be of benefit. (54) Walking did appear to have a slightly greater effect on function than pain outcomes. Inclusion of educational and behavioral components alongside walking in many studies may have contributed to this apparent effect; lending support to treatment approaches which place greater emphasis on improving function despite continued pain. (55) These interventions are often based on psychological theories such as operant conditioning which use positive reinforcement to reduce negative pain behaviors; for example through graded activity or pacing. (56) The underlying mechanisms contributing to these effects are uncertain but could be related to reduced fear of movement or increased self-efficacy. (55) Although co-interventions varied, there were commonalities: including that they frequently consisted of hospital or clinic-based group discussions (supplemented with written information), with condition-specific and general information on pain management strategies and advice on maintaining exercise. Some studies included additional strategies including goal setting and self-monitoring. Use of self-monitoring techniques including pedometer feedback represents a potentially useful method to increase walking in individuals with CMP disorders. (43, 51) However, these methods have not been widely tested. This is reflected in the fact that only three of the included studies used pedometers. A recent study examining a remote, web-delivered pedometer intervention (excluded from this review as it compared two forms of walking) found no
Walking for Musculoskeletal Pain: Review

long-term effects on functional outcomes.(57) Further work is required examining pedometer interventions in this population which are delivered within a clinical setting.

To our knowledge, this is one of the first systematic reviews to examine the effects of walking in a range of CMP disorders. A previous review (58) examining walking for LBP (both acute and chronic) found limited evidence to support its use as a primary intervention. Roddy and co-authors (59) found aerobic walking to be equally as effective as strengthening at reducing pain and disability in knee OA. Other reviews examining the effects of general aerobic exercise interventions in CMP (7,8,60,61) have provided conflicting results, with limited evidence to support the use of any one type or intensity of exercise. While aerobic exercise may lead to improved overall well-being and physical function it is often associated with little or no difference in pain.(60,62) In contrast, others have shown slight to moderate intensity aerobic exercise to be effective at reducing pain;(8) however this latter review did not look directly at effects on functional data.

Study strengths and limitations

This review has a number of strengths, including an extensive search of the available evidence, rather than limiting inclusion to studies selected on the basis of experimental design. We also included studies which involved only walking-based interventions, allowing for examination of a more homogenous intervention type than has previously been examined. Studies were considered to be similar on the basis of clinical characteristics and intervention methods. The majority involved supervised treadmill or land-based interventions (commonly within a hospital or clinic gymnasium setting), of between six to eight weeks duration. A number of these studies included more
Walking for Musculoskeletal Pain: Review

independent home-based walking as an additional exercise element and we were
therefore unable to determine the influence of treatment setting on outcomes.

We were unable to use sensitivity analysis to examine studies separately on the basis
of quality as only one study included in the meta-analysis contained a potential serious
methodological flaw which could have compromised its validity. Use of the USPSTF
system allowed a qualitative assessment of the overall evidence to be made, and the
findings and conclusions were broadly similar between this assessment method and
the results of the meta-analysis.

There are some limitations which should be taken into account when considering these
findings. A small number of studies had methodological limitations, including
inadequate allocation concealment in randomized controlled trials or lack of an
appropriate method for dealing with missing data. In six studies there was insufficient
information on masking of outcome assessments and with additional information it is
possible that some studies rated as “fair” may have been rated as “poor” which would
influence the recommendation made on the basis of the evidence included in the
review. Many studies lacked sufficient detail to assess adequacy of the exercise
interventions. The overall effects of the interventions may also have been attenuated
by the small number of non-intervention control groups. Furthermore, few studies
reported whether there were any associated adverse events. Even among the more
supervised interventions, there was limited detail regarding participant adherence.

Further research is required examining interventions which use objective measurement
of overall physical activity as both an important outcome and a method for increasing
motivation and use of self-monitoring. Objective monitors are more accurate than
subjective assessment methods, due to recall and social-desirability biases of subjective reports (63). Objective monitors such as pedometers can give immediate feedback on performance (prompting adherence), however, one limitation is that they require the user to remember to put them on. Other solutions, such as wrist worn, waterproof devices, that don’t need to be removed for sleep or water based activities may offer a solution, but may not provide the same quality of visual feedback that a pedometer does. Such issues should be considered in the design of future research.

Conclusions

Meta-analysis of data from studies of at least fair methodological quality demonstrated that walking may lead to improvements in outcome, comparable to other forms of exercise. Using the USPSTF system to summarize the existing evidence, walking-based exercise can be recommended for individuals with CMP. However, robustly designed research is required examining longer-term maintenance of walking programs and their effects on important health related outcomes in this population.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and design of this review. SOC, BR, CB and SMcD were responsible for conducting the search strategies and extracting study data. SOC, MT, GDB, JB and SMcD were responsible for assessment of study quality and rating the overall strength of evidence. SOC drafted the manuscript and all authors contributed to and approved the finalised version.
Walking for Musculoskeletal Pain: Review

References


6. Delitto A, George SZ, Van Dillen L, et al. Low Back Pain: Clinical Practice Guidelines Linked to the International Classification of Functioning, Disability, and
Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review

540  exercise for chronic musculoskeletal pain in adults. Cochrane Database Syst Rev
541  2010;1:CD005956.

543  review of reviews of intervention components associated with increased effectiveness

545  56. Vincent HK, Seay AN, Montero C, et al. Kinesiophobia and fear-avoidance beliefs
546  in overweight older adults with chronic low-back pain: relationship to walking

549  intervention for adults with chronic low back pain: randomised controlled trial. J Med

551  58. Hendrick P, Te Wake AM, Tikkisetty AS, et al. The effectiveness of walking as an

553  59. Roddy E, Zhang W, Doherty M. Aerobic walking or strengthening exercise for

555  60. Busch AJ, Barber KA, Overend TJ, et al. Exercise for treating fibromyalgia

Walking for Musculoskeletal Pain: Review


Walking for Musculoskeletal Pain: Review

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram showing process of selection for systematic review (16)

Figure 2. Effect of walking on pain (/100) compared to control interventions

Figure 3. Effect of walking on self-reported function (/100) compared to control interventions
Walking for Musculoskeletal Pain: Review

594
Walking for Musculoskeletal Pain: Review
Table 1. Summary of demographic information from individual studies (studies included in the meta-analysis are highlighted in bold)

<table>
<thead>
<tr>
<th>Study</th>
<th>Condition</th>
<th>Diagnostic Criteria</th>
<th>Duration of Symptoms (years) Mean (SD)</th>
<th>Age (years) Mean (SD)</th>
<th>Gender (% Female)</th>
<th>Mass (kg) / BMI Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bautch et al, 2000 (27)</td>
<td>OA knee</td>
<td>Clinical diagnosis according to ACR criteria *</td>
<td>-</td>
<td>69.7 (1.9)</td>
<td>66.7</td>
<td>- / 26.6 (1.0)</td>
</tr>
<tr>
<td>Bautch et al, 1997 (29)</td>
<td>OA knee</td>
<td>Clinical diagnosis according to ACR criteria *</td>
<td>-</td>
<td>69.0 (2.3) ‡</td>
<td>72.7 ‡</td>
<td>- / 28.7 (1.2)</td>
</tr>
<tr>
<td>Bircan et al, 2008 (44)</td>
<td>FM</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>4.2 (4.3)</td>
<td>47.2 (9.5)</td>
<td>100</td>
<td>- / -</td>
</tr>
<tr>
<td>Brosseau et al, 2012 (36)</td>
<td>OA knee</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>10.3 (9.3)</td>
<td>63.4 (8.6)</td>
<td>68.9</td>
<td>82.2 (16.6) / 29.8 (5.4)</td>
</tr>
<tr>
<td>Dias et al, 2003 (31)</td>
<td>OA knee</td>
<td></td>
<td>-</td>
<td>75 § (65-89 #)</td>
<td>86.4</td>
<td>- / -</td>
</tr>
<tr>
<td>Ettinger et al, 1997 (37)</td>
<td>OA knee</td>
<td>Radiographic evidence</td>
<td>-</td>
<td>68.6 (6.1)</td>
<td>70.4</td>
<td>- / -</td>
</tr>
<tr>
<td>Evcik et al, 2002 (38)</td>
<td>OA knee</td>
<td>Clinical and radiographic assessment using Kellgren &amp; Lawrence criteria</td>
<td>8.1 (3.3)</td>
<td>56.3 (6.5)</td>
<td>68.9</td>
<td>- / -</td>
</tr>
<tr>
<td>Ferrell et al, 1997 (52)</td>
<td>cMSK pain **</td>
<td>Clinical diagnosis of ‘stable’ lower extremity / mechanical LBP (&gt;3 months)</td>
<td>-</td>
<td>73.2 (3.7)</td>
<td>21.1</td>
<td>- / -</td>
</tr>
<tr>
<td>Hartvigsen et al, 2010 (48)</td>
<td>cLBP</td>
<td>Clinical diagnosis</td>
<td>-</td>
<td>46.7 (10.9)</td>
<td>71.6</td>
<td>- / -</td>
</tr>
<tr>
<td>Hiyama et al, 2012 (28)</td>
<td>OA knee</td>
<td>Clinical diagnosis</td>
<td>-</td>
<td>72.8 (5.4)</td>
<td>100</td>
<td>59.4 (6.9) / 23.7 (2.1)</td>
</tr>
<tr>
<td>Holtgrafe et al, 2007 (34)</td>
<td>FM</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>4.3 (4.7)</td>
<td>52.3 (18.1)</td>
<td>100</td>
<td>- / 27.9 (5.7)</td>
</tr>
<tr>
<td>Koldas Dogan</td>
<td>cLBP</td>
<td></td>
<td>4.5 (5.5)</td>
<td>42.1 (9.5)</td>
<td>78.2</td>
<td>- / -</td>
</tr>
<tr>
<td>Study</td>
<td>Region</td>
<td>Methodology</td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Concordance</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>et al, 2008 (49)</td>
<td>OA knee</td>
<td>Clinical and Radiographic evidence</td>
<td>-</td>
<td>69.4 (10.2)</td>
<td>83.4 - / -</td>
<td></td>
</tr>
<tr>
<td>Kovar et al, 1992 (39)</td>
<td>FM</td>
<td>Clinical diagnosis</td>
<td>10.1 (15.6)</td>
<td>49.4 (16.3)</td>
<td>84.6 -</td>
<td></td>
</tr>
<tr>
<td>Martin et al, 1996 (35)</td>
<td>FM</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>14.1 (7.2)</td>
<td>44.8 (9.8)</td>
<td>97.4 - / -</td>
<td></td>
</tr>
<tr>
<td>McDonough et al, 2013 (51)</td>
<td>cLBP</td>
<td>Clinical diagnosis</td>
<td>10.7 (7.7)</td>
<td>49.5 (20.1)</td>
<td>55.3 28.5 (6.9)</td>
<td></td>
</tr>
<tr>
<td>Messier et al, 2004 (40)</td>
<td>OA knee</td>
<td>Clinical and radiographic assessment using Kellgren &amp; Lawrence criteria</td>
<td>-</td>
<td>68.6 (0.4)</td>
<td>72.8 95.1 (1.2)</td>
<td></td>
</tr>
<tr>
<td>Meyer et al, 2000 (30)</td>
<td>FM</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>13.1 (15.5)</td>
<td>49.5 (6.3)</td>
<td>100 - / -</td>
<td></td>
</tr>
<tr>
<td>Miller et al, 2006 (41)</td>
<td>OA knee</td>
<td>Self-report + clinical diagnosis</td>
<td>-</td>
<td>69.5 (0.9)</td>
<td>62.1 97.8 (16.6) / 34.6 (4.4)</td>
<td></td>
</tr>
<tr>
<td>Nichols et al, 1994 (45)</td>
<td>FM</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>-</td>
<td>53.1 (11.5)</td>
<td>91.6 - / -</td>
<td></td>
</tr>
<tr>
<td>Rasmussen-Barr et al, 2009</td>
<td>cLBP</td>
<td>Clinical diagnosis</td>
<td>14.5 (1-38) #</td>
<td>57 (11.0)</td>
<td>2.8 76 (15) / 24.8 ¶</td>
<td></td>
</tr>
<tr>
<td>Rooks et al, 2007 (47)</td>
<td>FM</td>
<td>Clinical diagnosis according to ACR criteria †</td>
<td>5.7 (4.7)</td>
<td>49.7 (11.3)</td>
<td>100 76 (16.5) / 29.3 (6.2)</td>
<td></td>
</tr>
<tr>
<td>Schlenk et al, 2011 (42)</td>
<td>OA knee</td>
<td>Physician-confirmed diagnosis</td>
<td>11.3 (12.0)</td>
<td>63.2 (9.8)</td>
<td>96.0 - / 33.3 (6.0)</td>
<td></td>
</tr>
<tr>
<td>Shnayderman et al, 2012 (50)</td>
<td>cLBP</td>
<td>Clinical diagnosis (≥ 12 weeks)</td>
<td>-</td>
<td>45.3 (11.7)</td>
<td>78.8 73.9 (14.5) / 28.3(4.9)</td>
<td></td>
</tr>
<tr>
<td>Talbot et al, 2003 (43)</td>
<td>OA knee</td>
<td>Radiographic assessment using Kellgren &amp; Lawrence criteria</td>
<td>-</td>
<td>70.2 (5.5)</td>
<td>76.5 - / 31.8 (6.4)</td>
<td></td>
</tr>
<tr>
<td>Valim et al, 2003 (46)</td>
<td>FM</td>
<td>Clinical diagnosis</td>
<td>-</td>
<td>45.5 (10.5)</td>
<td>100 - / -</td>
<td></td>
</tr>
</tbody>
</table>

- = Not reported. OA = Osteoarthritis. FM = Fibromyalgia. cLBP = chronic Low Back Pain. * = American College of Rheumatology criteria for diagnosis of osteoarthritis. † = American College of Rheumatology criteria for diagnosis of fibromyalgia. ‡ = presents only demographic data from subjects who completed
study not total sample. § = median value. || = Standard error of mean (SEM). ¶ = Where not stated in paper value calculated based on mean mass and therefore unable to calculate SD. # = only range reported. ** = chronic musculoskeletal pain (hip, lower back and knee pain).
Table 2. Summary of methodological characteristics of individual studies (studies included in the meta-analysis are highlighted in bold)

<table>
<thead>
<tr>
<th>Study design &amp; blinding</th>
<th>Total sample</th>
<th>Walking group</th>
<th>Control group</th>
<th>Duration of intervention (months)</th>
<th>Reported adherence (%) †</th>
<th>Time point of follow-up assessment(s) (post-randomization): Drop out; N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bautch et al, 2000 (27) RCT/B</td>
<td>30</td>
<td>Education + treadmill walking</td>
<td>Education</td>
<td>3 m</td>
<td>-</td>
<td>3 m: 9 (30)</td>
</tr>
<tr>
<td>Bautch et al, 1997 (29) RCT/B</td>
<td>34</td>
<td>Education + treadmill walking</td>
<td>Education</td>
<td>3 m</td>
<td>-</td>
<td>3 m: 4 (11.7)</td>
</tr>
<tr>
<td>Bircan et al, 2008 (44) RCT/-</td>
<td>30</td>
<td>Treadmill walking</td>
<td>General Strengthening exercise</td>
<td>2 m</td>
<td>-</td>
<td>2 m: 4 (13.3)</td>
</tr>
<tr>
<td>Brosseau et al, 2012 (36) RCT/B</td>
<td>222</td>
<td>Supervised walking</td>
<td>Education</td>
<td>12 m</td>
<td>79.0</td>
<td>3 m: 37 (16.6) 6 m: 19 (8.5) 9 m: 17 (7.6) 12 m: 14 (6.3) 15 m: 5 (2.3) 18 m: 8 (3.6)</td>
</tr>
<tr>
<td>Dias et al, 2003 (31) RCT/B</td>
<td>50</td>
<td>Education + supervised exercise + home based walking</td>
<td>Education</td>
<td>1.5 m</td>
<td>-</td>
<td>3 m: - 6 m: 3 (6)</td>
</tr>
<tr>
<td>Ettinger et al, 1997 (37) RCT/B</td>
<td>439</td>
<td>Facility and home based walking</td>
<td>Education</td>
<td>18 m</td>
<td>68.0</td>
<td>3 m: 47 (10.7) 9 m: 82 (18.6) 18 m: 75 (17.1)</td>
</tr>
<tr>
<td>Evcik et al, 2002 (38) NRS/-</td>
<td>90</td>
<td>Home based walking</td>
<td>Instructed to continue with normal daily activities</td>
<td>3 m</td>
<td>-</td>
<td>6 m: 9 (10)</td>
</tr>
<tr>
<td>Study Authors, Year (Reference)</td>
<td>Intervention</td>
<td>Primary Outcome</td>
<td>Duration</td>
<td>Improvement</td>
<td>Comparison</td>
<td>Notes</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>----------</td>
<td>-------------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>Ferrell et al, 1997 (52) RCT/-</td>
<td>Supervised walking on outdoor track or gymnasium</td>
<td>Pain management information</td>
<td>1.5 m</td>
<td>93.0</td>
<td>2 w: -</td>
<td>2 m: 4 (12.2)</td>
</tr>
<tr>
<td>Hartvigsen et al, 2010 (48) RCT/B</td>
<td>Supervised Nordic walking</td>
<td>Education</td>
<td>2 m</td>
<td>-</td>
<td>2 m: 10 (7.4)</td>
<td>6 m: 0</td>
</tr>
<tr>
<td>Hiyama et al, 2012 (28) RCT/B</td>
<td>Home based walking (with pedometer) + ice + general home exercises</td>
<td>Ice + general home exercises</td>
<td>1 m</td>
<td>-</td>
<td>1 m: 0</td>
<td></td>
</tr>
<tr>
<td>Holtgrafe et al, 2007 (34) NRS/B</td>
<td>Hospital based indoor walking</td>
<td>Pre-intervention, baseline phase</td>
<td>2 m</td>
<td>-</td>
<td>2 m: -</td>
<td></td>
</tr>
<tr>
<td>Koldas Dogan et al, 2008 (49) RCT/-</td>
<td>Treadmill based exercise</td>
<td>General home exercises</td>
<td>1.5 m</td>
<td>-</td>
<td>1.5 m: 5 (8.3)</td>
<td>2.5 m: 5 (8.3)</td>
</tr>
<tr>
<td>Kovar et al, 1992 (39) RCT/B</td>
<td>Hospital based supervised walking + education</td>
<td>Contacted by phone to discuss nature of daily activities</td>
<td>2 m</td>
<td>87.5</td>
<td>2 m: 10 (9.8)</td>
<td></td>
</tr>
<tr>
<td>Lemstra et al, 2005 (33) RCT/B</td>
<td>Supervised aerobic exercise + massage + education</td>
<td>Standard care</td>
<td>1.5 m</td>
<td>90.6</td>
<td>1.5 m: 7 (8.8)</td>
<td>15 m: 8 (10.2)</td>
</tr>
<tr>
<td>Martin et al, 1996 (35) RCT/B</td>
<td>Supervised walking + strength and flexibility training</td>
<td>Relaxation sessions</td>
<td>1.5 m</td>
<td>-</td>
<td>1.5 m: 20 (33.3)</td>
<td></td>
</tr>
<tr>
<td>McDonough et al, 2013 (51) fRCT/B</td>
<td>Pedometer based walking + education</td>
<td>Education</td>
<td>2 m</td>
<td>73.0</td>
<td>2 m: 7 (12.5)</td>
<td>6 m: 8 (14.3)</td>
</tr>
<tr>
<td>Messier et al, 2004 (40) RCT/B</td>
<td>Facility and home based aerobic + lower limb resistance training</td>
<td>Usual Care</td>
<td>18 m</td>
<td>60.0</td>
<td>6 m: 41 (12.9)</td>
<td>18 m: 64 (20.2)</td>
</tr>
<tr>
<td>Meyer et al, 2000 (30) RCT/-</td>
<td>Home based walking</td>
<td>Instructed to maintain current activity levels</td>
<td>6 m</td>
<td>-</td>
<td>6 m: -</td>
<td>18 m: - (57.2)</td>
</tr>
<tr>
<td>Study</td>
<td>Sample Size</td>
<td>Intervention Description</td>
<td>Control Description</td>
<td>Duration</td>
<td>Compliance</td>
<td>Adherence</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>----------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Miller et al., 2006</td>
<td>87</td>
<td>Education + facility / home based lower limb strengthening + aerobic training*</td>
<td>Education</td>
<td>6 m</td>
<td>77.5</td>
<td>6 m: 8 (9.1)</td>
</tr>
<tr>
<td>Nichols et al., 1994</td>
<td>24</td>
<td>Supervised indoor walking program</td>
<td>Instructed to continue with usual daily activities</td>
<td>2 m</td>
<td>-</td>
<td>2 m: 5 (20.8)</td>
</tr>
<tr>
<td>Rasmussen-Barr et al., 2009</td>
<td>71</td>
<td>Instructed to walk each day plus given general home exercises</td>
<td>Specific stabilization exercises with bio-pressure unit</td>
<td>2 m</td>
<td>71.0</td>
<td>6 m: 7 (9.8)</td>
</tr>
<tr>
<td>Rooks et al., 2007</td>
<td>207</td>
<td>Treadmill walking + flexibility training</td>
<td>Education</td>
<td>3 m</td>
<td>73.0</td>
<td>3 m: 72 (20.2)</td>
</tr>
<tr>
<td>Schlenk et al., 2011</td>
<td>26</td>
<td>Fitness walking + Education</td>
<td>Usual care + Education</td>
<td>6 m</td>
<td>-</td>
<td>6 m: 5 (19.2)</td>
</tr>
<tr>
<td>Shnayderman et al., 2012</td>
<td>52</td>
<td>Treadmill walking</td>
<td>General strengthening exercise</td>
<td>1.5 m</td>
<td>-</td>
<td>1.5 m: 9 (17.3)</td>
</tr>
<tr>
<td>Talbot et al., 2003</td>
<td>40</td>
<td>Pedometer based walking</td>
<td>Education</td>
<td>3 m</td>
<td>76.0</td>
<td>3 m: -</td>
</tr>
<tr>
<td>Valim et al., 2003</td>
<td>76</td>
<td>Supervised walking</td>
<td>General stretching exercises</td>
<td>5 m</td>
<td>-</td>
<td>2.5 m: -</td>
</tr>
</tbody>
</table>

RCT: Randomized Controlled Trial. fRCT: Feasibility Randomized Controlled Trial. URT: Uncontrolled Randomized Trial. NRS: Non Randomized Study. B: Blinded outcome assessment. U: Unblinded. - = Not reported. * Walking primary mode of aerobic exercise. † Percentage adherence reported as total number of classes attended; self-reported completion of home exercise or self-reported adherence to wearing a pedometer. w: weeks. m: months.
Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram showing process of selection for systematic review (16)

Number of citations identified during electronic searches

- Total = 5449
- Medline = 2689
- CINAHL = 1534
- PsychINFO = 132
- PEDro = 113
- SPORT Discus = 636
- Cochrane Register = 345

Number of citations screened = 2760

Non-relevant citations excluded = 2512

Number of full text papers reviewed = 248

Papers excluded after full text review = 222

- Reasons for exclusion:
  - Other type of aerobic exercise = 134
  - No control group = 15
  - No relevant outcomes = 28
  - Walking not a predominant component of intervention = 34
  - Follow up or sub sample analysis studies = 11

Number of studies included in Systematic review = 26

Number of studies included in meta-analysis = 17

Number of citations identified from other sources (hand searching of reference lists) = 18

Number of citations remaining after removing duplicates = 2760

Number of citations identified during electronic searches

- Total = 5449
- Medline = 2689
- CINAHL = 1534
- PsychINFO = 132
- PEDro = 113
- SPORT Discus = 636
- Cochrane Register = 345

Number of citations screened = 2760

Non-relevant citations excluded = 2512

Number of full text papers reviewed = 248

Papers excluded after full text review = 222

- Reasons for exclusion:
  - Other type of aerobic exercise = 134
  - No control group = 15
  - No relevant outcomes = 28
  - Walking not a predominant component of intervention = 34
  - Follow up or sub sample analysis studies = 11

Number of studies included in Systematic review = 26

Number of studies included in meta-analysis = 17

Number of citations identified from other sources (hand searching of reference lists) = 18

Number of citations remaining after removing duplicates = 2760

Number of citations identified during electronic searches

- Total = 5449
- Medline = 2689
- CINAHL = 1534
- PsychINFO = 132
- PEDro = 113
- SPORT Discus = 636
- Cochrane Register = 345

Number of citations screened = 2760

Non-relevant citations excluded = 2512

Number of full text papers reviewed = 248

Papers excluded after full text review = 222

- Reasons for exclusion:
  - Other type of aerobic exercise = 134
  - No control group = 15
  - No relevant outcomes = 28
  - Walking not a predominant component of intervention = 34
  - Follow up or sub sample analysis studies = 11

Number of studies included in Systematic review = 26

Number of studies included in meta-analysis = 17

Number of citations identified from other sources (hand searching of reference lists) = 18

Number of citations remaining after removing duplicates = 2760
Figure 2. Effect of walking on pain (/100) compared to control interventions

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Walking</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Control</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bircan 2008</td>
<td>21.9</td>
<td>18.8</td>
<td>13</td>
<td>25.5</td>
<td>14.1</td>
<td>13</td>
<td>2.8</td>
<td>1</td>
<td>-4.69</td>
<td>-17.37, 8.17</td>
<td></td>
</tr>
<tr>
<td>Ferrell 1997</td>
<td>36.4</td>
<td>25.6</td>
<td>9</td>
<td>55.9</td>
<td>21</td>
<td>9</td>
<td>1.1</td>
<td>1</td>
<td>-19.50</td>
<td>-41.13, 21.13</td>
<td></td>
</tr>
<tr>
<td>Hartvigsen 2010</td>
<td>38.8</td>
<td>45</td>
<td>45</td>
<td>7.5</td>
<td>54</td>
<td>45</td>
<td>14.2</td>
<td>2</td>
<td>-5.00</td>
<td>-9.20, -0.80</td>
<td></td>
</tr>
<tr>
<td>Koldos Ogden 2008</td>
<td>34.3</td>
<td>30.8</td>
<td>19</td>
<td>40</td>
<td>21.0</td>
<td>18</td>
<td>1.6</td>
<td>1</td>
<td>-5.10</td>
<td>-22.22, 12.02</td>
<td></td>
</tr>
<tr>
<td>Korpi 1992</td>
<td>37.2</td>
<td>17.3</td>
<td>47</td>
<td>7.7</td>
<td>25</td>
<td>45</td>
<td>5.0</td>
<td>1</td>
<td>-10.00</td>
<td>-16.62, -3.18</td>
<td></td>
</tr>
<tr>
<td>McDonald 2002</td>
<td>45.8</td>
<td>28.6</td>
<td>30</td>
<td>26</td>
<td>26.2</td>
<td>17</td>
<td>2.2</td>
<td>1</td>
<td>6.00</td>
<td>-3.84, 20.84</td>
<td></td>
</tr>
<tr>
<td>Nichols 1984</td>
<td>24.5</td>
<td>13.3</td>
<td>12</td>
<td>29.4</td>
<td>13.2</td>
<td>12</td>
<td>2.6</td>
<td>1</td>
<td>-4.93</td>
<td>-18.13, 8.33</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>164</strong></td>
<td></td>
<td></td>
<td><strong>159</strong></td>
<td><strong>29.4</strong></td>
<td><strong>5.31</strong></td>
<td><strong>-4.06, -2.56</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 5.68, df = 6 (P = 0.53); τ² = 0%
Test for overall effect: Z = 3.70 (P = 0.0002)

1.2.2 Mid-term outcomes (>2-12 months)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Walking</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Control</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brosseau 2012</td>
<td>24.6</td>
<td>15.8</td>
<td>44</td>
<td>25</td>
<td>19.4</td>
<td>41</td>
<td>8.3</td>
<td>2</td>
<td>-0.49</td>
<td>7.55, 7.15</td>
<td></td>
</tr>
<tr>
<td>Evkel 2002</td>
<td>34</td>
<td>13</td>
<td>20</td>
<td>60</td>
<td>33</td>
<td>26</td>
<td>2.6</td>
<td>2</td>
<td>-26.00</td>
<td>-39.47, -12.43</td>
<td></td>
</tr>
<tr>
<td>Hartvigsen 2010</td>
<td>38</td>
<td>14</td>
<td>45</td>
<td>42</td>
<td>13.5</td>
<td>45</td>
<td>8.6</td>
<td>2</td>
<td>-4.00</td>
<td>0.88, 8.68</td>
<td></td>
</tr>
<tr>
<td>Koldos Ogden 2008</td>
<td>34.1</td>
<td>27.6</td>
<td>19</td>
<td>33.6</td>
<td>24.3</td>
<td>18</td>
<td>1.7</td>
<td>2</td>
<td>0.50</td>
<td>-16.23, 17.23</td>
<td></td>
</tr>
<tr>
<td>McDonald 2002</td>
<td>38.5</td>
<td>25.3</td>
<td>39</td>
<td>41</td>
<td>26.3</td>
<td>17</td>
<td>2.1</td>
<td>2</td>
<td>-3.00</td>
<td>17.31, 11.31</td>
<td></td>
</tr>
<tr>
<td>Miller 2006</td>
<td>20.5</td>
<td>12.9</td>
<td>44</td>
<td>30.5</td>
<td>16.4</td>
<td>43</td>
<td>0.0</td>
<td>2</td>
<td>-10.00</td>
<td>-16.21, -3.79</td>
<td></td>
</tr>
<tr>
<td>Rooks 2007</td>
<td>48</td>
<td>26</td>
<td>51</td>
<td>56</td>
<td>22</td>
<td>50</td>
<td>4.7</td>
<td>2</td>
<td>-11.00</td>
<td>-20.18, -1.82</td>
<td></td>
</tr>
<tr>
<td>Talent 2003</td>
<td>21.4</td>
<td>16</td>
<td>17</td>
<td>31.4</td>
<td>22.4</td>
<td>17</td>
<td>2.8</td>
<td>2</td>
<td>-10.09</td>
<td>-23.06, 3.09</td>
<td></td>
</tr>
<tr>
<td>Valim 2003</td>
<td>34.2</td>
<td>25</td>
<td>32</td>
<td>46</td>
<td>21.8</td>
<td>28</td>
<td>2.1</td>
<td>2</td>
<td>-11.80</td>
<td>-23.34, 0.04</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>319</strong></td>
<td></td>
<td></td>
<td><strong>285</strong></td>
<td><strong>40.0%</strong></td>
<td><strong>7.92</strong></td>
<td><strong>12.37, -3.48</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 15.64; Chi² = 15.02, df = 8 (P = 0.06); τ² = 41%
Test for overall effect: Z = 3.20 (P = 0.0015)

1.2.3 Long-term outcomes (>12 months)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Walking</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Control</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brosseau 2012</td>
<td>23.5</td>
<td>15.1</td>
<td>43</td>
<td>23.5</td>
<td>17.8</td>
<td>35</td>
<td>6.4</td>
<td>2</td>
<td>0.10</td>
<td>7.33, 7.53</td>
<td></td>
</tr>
<tr>
<td>Ellinger 1987</td>
<td>36.5</td>
<td>10</td>
<td>144</td>
<td>40</td>
<td>10.2</td>
<td>148</td>
<td>16.4</td>
<td>2</td>
<td>-4.40</td>
<td>4.51, 12.31</td>
<td></td>
</tr>
<tr>
<td>Neser 2004</td>
<td>31.2</td>
<td>21.1</td>
<td>84</td>
<td>30.1</td>
<td>19.8</td>
<td>78</td>
<td>7.7</td>
<td>2</td>
<td>1.10</td>
<td>5.26, 7.46</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>267</strong></td>
<td></td>
<td></td>
<td><strong>262</strong></td>
<td><strong>30.6%</strong></td>
<td><strong>-2.27</strong></td>
<td><strong>-6.83, 2.59</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 5.28; Chi² = 3.40, df = 2 (P = 0.10); τ² = 43%
Test for overall effect: Z = 1.14 (P = 0.25)

| Total (95% CI) | 770 | 708 | 100.0% | Mean Difference | 5.29 | 7.58, 3.01 |
Favours [experimental] | Favours [control] |

Heterogeneity: Tau² = 6.69; Chi² = 27.88, df = 18 (P = 0.06); τ² = 36%
Test for subgroup differences: Chi² = 3.75, df = 2 (P = 0.15), τ² = 46.7%
Figure 3. Effect of walking on self-reported function (/100) compared to control interventions

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Walking</th>
<th>Control</th>
<th>Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>3.1.1 Short-term outcomes (≤ 8 weeks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bircan 2008</td>
<td>85</td>
<td>21.2</td>
<td>13</td>
<td>52.3</td>
</tr>
<tr>
<td>Ferret 1997</td>
<td>41.5</td>
<td>27.6</td>
<td>9</td>
<td>57</td>
</tr>
<tr>
<td>Hartingsen 2010</td>
<td>37.1</td>
<td>10.6</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>Koldos Odogan 2008</td>
<td>37.1</td>
<td>28.3</td>
<td>19</td>
<td>56.6</td>
</tr>
<tr>
<td>Koren 1992</td>
<td>33.0</td>
<td>13.1</td>
<td>47</td>
<td>55.1</td>
</tr>
<tr>
<td>McDonough 2013</td>
<td>32.5</td>
<td>17.2</td>
<td>26</td>
<td>28.6</td>
</tr>
<tr>
<td>Shawdellerman 2012</td>
<td>22.8</td>
<td>14.4</td>
<td>26</td>
<td>19.1</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>196</td>
<td>173</td>
<td>34.8%</td>
<td>-6.47 [10.59, -2.85]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 30.06; Chi² = 17.16, df = 6 (P = 0.009); I² = 65%
Test for overall effect: Z = 3.30 (P = 0.00)

3.1.2 Mid-term outcomes (>2.12 months) |         |         |                 |                   |         |       |        |                   |
| Brousseau 2012    | 24.5    | 13.6    | 44              | 25.1              | 13.5    | 41    | 7.2%   | -0.60 [8.41, 5.21] |
| Evik 2002         | 15.3    | 3.6     | 28              | 30.4              | 6.5     | 26    | 9.2%   | -15.43 [16.21, -12.50] |
| Hartingsen 2010   | 34.0    | 13.3    | 45              | 44                 | 12.4    | 45    | 7.7%   | -10.00 [15.31, -4.69] |
| Koldos Odogan 2008| 38.3    | 30.4    | 19              | 55.4              | 30.4    | 18    | 1.9%   | -17.10 [30.70, -2.50] |
| McDonough 2013    | 38.0    | 25.3    | 39              | 41                 | 26.3    | 17    | 2.9%   | -3.00 [17.81, 11.81] |
| Miller 2008       | 23.3    | 14.6    | 44              | 36.1              | 19.3    | 43    | 6.4%   | -12.70 [18.80, -5.50] |
| Brooks 2007       | 41.1    | 20.3    | 51              | 50.7              | 23.3    | 50    | 5.5%   | -9.90 [18.26, -0.54] |
| Valm 2013         | 25.7    | 17.4    | 32              | 31.7              | 21.5    | 26    | 4.8%   | -6.00 [15.96, 3.99] |
| Subtotal (95% CI) | 302     | 268     | 45.8%           | -9.31 [14.00, -4.81] |

Heterogeneity: Tau² = 27.47; Chi² = 24.24, df = 7 (P = 0.001); I² = 71%
Test for overall effect: Z = 3.88 (P = 0.001)

3.1.3 Long-term outcomes (>12 months) |         |         |                 |                   |         |       |        |                   |
| Brousseau 2012    | 18.2    | 14.3    | 43              | 15.4              | 17.1    | 35    | 8.4%   | -1.20 [8.35, 9.89] |
| Eling 1997        | 34.4    | 10.4    | 144             | 40                 | 8.2     | 145   | 0.0%   | -5.60 [17.70, -3.50] |
| Sevasti 2011      | 37.9    | 15.4    | 13              | 31.7              | 15.1    | 13    | 3.4%   | -3.90 [17.26, 9.46] |
| Subtotal (95% CI) | 200     | 197     | 18.4%           | -5.22 [-7.21, -3.23] |

Heterogeneity: Tau² = 0.00; Chi² = 1.38, df = 2 (P = 0.50); I² = 0%
Test for overall effect: Z = 1.14 (P = 0.00001)

Total (95% CI) | 700 | 638 | 100.0% | -7.26 [-10.29, -4.24] |

Heterogeneity: Tau² = 23.67; Chi² = 82.54, df = 17 (P = 0.00001), I² = 73%
Test for overall effect: Z = 4.71 (P = 0.00001)
Test for subgroup differences: Chi² = 2.81, df = 2 (P = 0.27), I² = 20.3%
Supplementary data: Appendix A. Medical Subject Heading (MeSH) terms used for identification of relevant studies

Medline (via Ovid) search strategy:
# Searches
1 motor activity.de.
2 walk$.de.
   lifestyle.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
3 free-living activit$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
4 accelerometer$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
5 pedometer$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
6 activity monitor$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
7 physical fitness.de.
8 exercise therapy.de.
9 aerobic$.mp.
10 exercis$.mp.
11 physical exercise.mp.
12 Musculoskeletal pain.mp.
13 Musculoskeletal diseases.mp.
14 dorsalgia.mp.
15 backache.mp.
16 back pain.mp.
17 Low back pain.de.
18 fibromyalgia.mp.
19 fibromyalgia syndrome.mp.
20 arthritis.mp.
21 osteoarthritis.mp.
22 rehabilitation.de.
23 morbidity.de.
24 mortality.de.
25 randomised controlled trial.mp.
26 controlled clinical trial.mp.
27 double blind method.mp.
28 single-blind method.mp.
29 1 or 2 or 3 or 4 or 5 or 6 or 7
30 8 or 9 or 10 or 11 or 12
31 13 or 14
32 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
33 23 or 24 or 25
34 26 or 27 or 28 or 29
35 30 and 31
36 32 and 36
37 38 and 36
39 30 and 34
40 33 and 35 and 36
41 33 or 35 or 36
42 30 or 33
43 limit 42 to yr="1980-Current"
44 30 or 32
45 limit 44 to yr="1980-Current"
46 35 and 43
47 35 and 45
48 33 and 35 and 36
49 limit 44 to yr="1980-Current"
50 31 and 33
51 limit 50 to yr="1980-Current"
Supplementary data: Appendix B. Quality assessment and adequacy of exercise intervention criteria for individual studies (studies included in the meta-analysis are highlighted in bold)

<table>
<thead>
<tr>
<th>Study</th>
<th>Internal validity</th>
<th>External validity</th>
<th>ACSM criteria met</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bautch et al, 2000 (27)</td>
<td>Poor</td>
<td>Fair</td>
<td>1,3,4,5</td>
</tr>
<tr>
<td>Bautch et al, 1997 (29)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,3,4,5</td>
</tr>
<tr>
<td>Bircan et al, 2008 (44)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Brosseau et al, 2012 (36)</td>
<td>Fair</td>
<td>Good</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Dias et al, 2003 (31)</td>
<td>Poor</td>
<td>Fair</td>
<td>1,3,4,5</td>
</tr>
<tr>
<td>Ettinger et al, 1997 (37)</td>
<td>Good</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Evcik et al, 2002 (38)</td>
<td>Fair</td>
<td>Good</td>
<td>1,4,5</td>
</tr>
<tr>
<td>Ferrell et al, 1997 (52)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,4,5</td>
</tr>
<tr>
<td>Hartvigsen et al, 2010 (48)</td>
<td>Good</td>
<td>Fair</td>
<td>1,3,4,5</td>
</tr>
<tr>
<td>Hiyama et al, 2012 (28)</td>
<td>Poor</td>
<td>Fair</td>
<td>4</td>
</tr>
<tr>
<td>Holtgrafe et al, 2007 (34)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,4,5</td>
</tr>
<tr>
<td>Koldas Dogan et al, 2008</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Kovar et al, 1992 (39)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,3,4,5</td>
</tr>
<tr>
<td>Lemstra et al, 2005 (33)</td>
<td>Good</td>
<td>Good</td>
<td>1,2,4,5</td>
</tr>
<tr>
<td>Martin et al, 1996 (35)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>McDonough et al, 2013</td>
<td>Good</td>
<td>Good</td>
<td>1,4,5</td>
</tr>
<tr>
<td>(51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Messier et al, 2004 (40)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Meyer et al, 2000 (30)</td>
<td>Poor</td>
<td>Fair</td>
<td>1,4,5</td>
</tr>
<tr>
<td>Miller et al, 2006 (41)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Nichols et al, 1994 (45)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Rasmussen-Barr et al, 2009 (32)</td>
<td>Good</td>
<td>Good</td>
<td>1,4,5</td>
</tr>
<tr>
<td>Rooks et al, 2007 (47)</td>
<td>Good</td>
<td>Good</td>
<td>1,3,4,5</td>
</tr>
<tr>
<td>Schlenk et al, 2011 (42)</td>
<td>Fair</td>
<td>Good</td>
<td>4,5</td>
</tr>
<tr>
<td>Shnayderman et al, 2012</td>
<td>Fair</td>
<td>Good</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>(50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Talbot et al, 2003 (43)</td>
<td>Poor</td>
<td>Good</td>
<td>1,4,5</td>
</tr>
<tr>
<td>Valim et al, 2003 (46)</td>
<td>Fair</td>
<td>Fair</td>
<td>1,2,3,4,5</td>
</tr>
</tbody>
</table>

Using the following guideline criteria, internal and external validity of individual studies were judged as “good” “fair” or “poor” based on the following guideline criteria:

For internal validity: (1) Initial assembly of comparable groups: For RCTs: Adequate randomization including concealment and whether potential confounders were distributed equally among groups. (2) Maintenance of comparable groups (includes attrition, crossovers, adherence, and contamination). (3) Important differential loss to follow-up or overall high loss to follow-up. (4) Measurements: equal, reliable, and valid (includes masking of outcome assessment). (5) Clear definition of interventions. (6) All important outcomes considered. (7) Analysis: Intention-to treat analysis used for RCTs.

For external validity: (1) Biologic plausibility. (2) Similarities of the populations studied and primary care patients (in terms of risk factor profile, demographics, ethnicity, gender, clinical presentation, and similar factors). (3) Similarities of the test or intervention studied to those that would be routinely available or feasible in typical practice. (4) Clinical or social
environmental circumstances in the studies that could modify the results from those expected in a primary care setting.

American College of Sports Medicine (ACSM) criteria for assessment of the adequacy of exercise interventions in individual studies: (1) Frequency of exercise of at least three days per week or twice a week for deconditioned individuals. (2) Intensity of exercise sufficient to achieve equal to or greater than 40% of heart rate reserve (min-max: 40-85%) or 64% of predicted maximum heart rate (min-max: 64-94%). (3) Sessions of at least 20 minutes duration (min-max: 20-60 minutes), either as continuous exercise or spread intermittently throughout the day in blocks of 10 minutes or more. (4) A mode of aerobic exercise involving major muscle groups in rhythmic activities. (5) Intervention should last for a minimum of six weeks.
Supplementary data: Appendix C. List of studies excluded from the systematic review where walking was not considered to be the predominant component of the intervention


McCain GA, BELL DA, MAI FM ET AL. A controlled study of the effects of a supervised cardiovascular fitness training program on the manifestations of primary fibromyalgia. Arthritis and Rheumatism, 1988;31(9):1135-1141


Chan CW, Mok NW, Yeung EW. Aerobic exercise training in addition to conventional physiotherapy for chronic low back pain: a randomized controlled trial. Archives of Physical Medicine and Rehabilitation, 2011;92(10):1681-5

Rantonen J. Luoto S. Vehtari A. et al. The effectiveness of two active interventions compared to self-care advice in employees with non-acute low back symptoms: a randomised, controlled trial with a 4-year follow-up in the occupational health setting. Occupational & Environmental Medicine, 2012;69(1):12-20


