Critical Review

The Effectiveness of Worksite Physical Activity Programs on Physical Activity, Physical Fitness, and Health

Karin I. Proper, MSc, Marjan Koning, MSc, Allard J. van der Beek, PhD, Vincent H. Hildebrandt, MD, PhD, Ruud J. Bosscher, PhD, and Willem van Mechelen, MD, PhD

Objective: To critically review the literature with respect to the effectiveness of worksite physical activity programs on physical activity, physical fitness, and health.

Data Sources: A search for relevant English-written papers published between 1980 and 2000 was conducted using MEDLINE, EMBASE, Sportdiscus, CINAHL, and Psychlit. The key words used involved a combination of concepts regarding type of study, study population, intervention, and outcome measure. In addition, a search was performed in our personal databases, as well as a reference search of the studies retrieved.

Study Selection: The following criteria for inclusion were used: 1) randomized, controlled trial or nonrandomized, controlled trial; 2) working population; 3) worksite intervention program to promote employees’ physical activity or physical fitness; and 4) physical activity, physical fitness, or health-related outcomes.

Data Extraction: Two reviewers independently evaluated the quality of relevant studies using a predefined set of nine methodological criteria. Conclusions regarding the effectiveness of a worksite physical activity programs were based on a rating system consisting of five levels of evidence.

Data Synthesis: Fifteen randomized, controlled trials and 11 nonrandomized, controlled trials met the criteria for inclusion and were reviewed. Six randomized, controlled trials and none of the nonrandomized, controlled trials were of high methodological quality. Strong evidence was found for a positive effect of a worksite physical activity program on physical activity and musculoskeletal disorders. Limited evidence was found for a positive effect on fatigue. For physical fitness, general health, blood serum lipids, and blood pressure, inconclusive or no evidence was found for a positive effect.

Conclusions: To increase the level of physical activity and to reduce the risk of musculoskeletal disorders, we support implementation of worksite physical activity programs. For the other outcome measures, scientific evidence of the effectiveness of such a program is still limited or inconclusive, which is mainly the result of the small number of high-quality trials. Therefore, we recommend performing more randomized, controlled trials of high methodological quality, taking into account criteria such as randomization, blinding, and compliance.

Key Words: Physical activity—Physical fitness—Exercise—Intervention studies—Health—Workplace.

INTRODUCTION

The positive associations between physical activity and health are no longer subject to debate. People who are physically active at a sufficient level obtain a wide array of physical and mental health benefits compared with those who are not active. However, because the majority of adults in developed countries are not physically active to a satisfactory degree, promoting physical activity is of great relevance. Based on the fact that most adults spend about 8 hours a day at their workplace, offering physical activity programs at the workplace can be an efficient way to enhance adults’ levels of activity. In the past few decades, programs aimed at increasing employees’ physical activity or fitness have become popular in a wide variety of work settings. In the last 20 years, many studies were performed regarding the effectiveness of programs enhancing workers’ physical activity or fitness levels. Also, some reviews were conducted addressing the effectiveness of such programs. However, no systematic review on the effectiveness of worksite physical activity programs (WPAPs) on health-related benefits has been performed, with the exception of the meta-analysis by Dishman et al. Dishman et al. conducted a quantitative synthesis of the literature and concluded that WPAPs have a small,
nonsignificant positive effect on physical activity or fitness. However, no other health-related components were evaluated in that review. The purpose of the present review is to systematically assess the effectiveness of WPAPs on physical activity, physical fitness, and health.

METHODS

Study Selection
The following three steps were taken to identify relevant studies published in the English language between 1980 and 2000: 1) a computerized search of MEDLINE, EMBASE, Sportdiscus, CINAHL, and Psychlit; 2) a reference search of studies retrieved; and 3) a search in our personal databases. Table I presents the search strategy used. A study was included if 1) the study was a randomized, controlled trial (RCT) or a nonrandomized, controlled trial (NCT); 2) the study population involved a healthy working population; 3) the intervention was a worksite program aimed at enhancing levels of physical activity, exercise, and/or fitness; and 4) the outcome measure included physical activity, health-related fitness, or health.

Methodological Quality and Best Evidence Synthesis
Two reviewers (K.I.P. and M.K.) independently evaluated the methodological quality of the studies by using a criteria list based on the list of the Cochrane Collaboration Back Review Group. An item was scored positive in case of a satisfactory description and the use of adequate methods. Disagreements between the two reviewers were discussed during a meeting to achieve consensus. If they could not reach agreement, a third reviewer (A.J.B.) was consulted to achieve final judgment. To draw conclusions regarding the effectiveness, a rating system consisting of five levels of evidence (i.e., strong, moderate, limited, inconclusive, or no evidence) was applied, which was derived from several existing best evidence syntheses. For the description of the methodological quality criteria and the best evidence synthesis, we refer to Proper et al., who performed a systematic review of the effectiveness of WPAPs on work-related outcomes. Studies were considered to be of (relatively) high quality if more than 50% of the methodological criteria were scored positively. A WPAP was considered to have a positive effect in case of statistically significant results or a relevant effect size (i.e., ≥20% difference between study groups). By taking a relevant effect size into account next to statistical significance, the problem of a lack of statistical significant results due to a lack of statistical power is overcome. By categorizing studies according to a level of evidence, a hierarchical order of design and quality of the studies was created to draw conclusions as to the effectiveness of a WPAP on each outcome measure. For example, in case of two or more high-quality RCTs, conclusions were based on these RCTs only, leading to a conclusion of “strong evidence” in case of consistent results or leading to “inconclusive evidence” in case of inconsistent results, regardless of the results of any other study. However, in the case of only one or no high-quality RCTs, the conclusion of “strong evidence” was not possible, and the conclusion had to be drawn on the basis of this single high-quality RCT, eventually in combination with the results of available low-quality RCTs.

RESULTS

Selected Studies
The search identified a total of 772 publications. After reading the title or abstract, 693 publications were excluded. Common reasons for exclusion were the lack of comparison groups, the lack of relevant outcome measures, or an intervention not having its main focus on physical activity or fitness. In addition, 50 publications were excluded for not having met one of the inclusion criteria or because they could not be retrieved. Finally, 29 publications concerning 26 studies were selected. Tables 2 and 3 present characteristics of the relevant RCTs (n = 15) and NCTs (n = 11), respectively. Initially, the two reviewers disagreed on the methodological quality of 30 of the 223 items (13%). Cohen’s κ was 0.72. Disagreement was mainly due to differences in the interpretation of the definition of the methodological quality items and to reading errors. As the two reviewers could reach complete agreement after a discussion session, the third reviewer was not consulted. Table 4 shows the methodological quality score of the studies reviewed. In general, the quality of the studies was poor: six RCTs were of high quality, and none of the NCTs was of high quality. Common methodological limitations of the studies included a poor description of the randomization procedure or an inadequate randomization procedure, inappropriate blinding of the outcome assessor, and/or the absence of an analysis according to the intention-to-treat principle. For the NCTs, either the description or the rate of compliance with the program was also insufficient.

**TABLE 1. Search strategy (databases, hits per database, and key words) used for literature search**

<table>
<thead>
<tr>
<th>Databases (hits)</th>
<th>Key words used</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE (n = 266)</td>
<td>(intervention or evaluation or randomis* or randomiz* or control* or effect) and (workplace or worksite or work site or working place or worker or occupation or employer or employee or corporate or company or business or industry or industrial) and ([exercise or fitness or physical activity or sport] and [program or programme]) and (physical activity or fitness or health or musculoskeletal or back or neck or shoulder or elbow or wrist or knee or upper extremity or lower extremity)</td>
</tr>
<tr>
<td>EMBASE (n = 89)</td>
<td>(intervention or evaluation or randomis* or randomiz* or control* or effect) and (workplace or worksite or work site or working place or worker or occupation or employer or employee or corporate or company or business or industry or industrial) and ([exercise or fitness or physical activity or sport] and [program or programme]) and (physical activity or fitness or health or musculoskeletal or back or neck or shoulder or elbow or wrist or knee or upper extremity or lower extremity)</td>
</tr>
<tr>
<td>Sportdiscus (n = 263)</td>
<td>(intervention or evaluation or randomis* or randomiz* or control* or effect) and (workplace or worksite or work site or working place or worker or occupation or employer or employee or corporate or company or business or industry or industrial) and ([exercise or fitness or physical activity or sport] and [program or programme]) and (physical activity or fitness or health or musculoskeletal or back or neck or shoulder or elbow or wrist or knee or upper extremity or lower extremity)</td>
</tr>
<tr>
<td>CINAHL (n = 44)</td>
<td>(intervention or evaluation or randomis* or randomiz* or control* or effect) and (workplace or worksite or work site or working place or worker or occupation or employer or employee or corporate or company or business or industry or industrial) and ([exercise or fitness or physical activity or sport] and [program or programme]) and (physical activity or fitness or health or musculoskeletal or back or neck or shoulder or elbow or wrist or knee or upper extremity or lower extremity)</td>
</tr>
<tr>
<td>Psychlit (n = 75)</td>
<td>(intervention or evaluation or randomis* or randomiz* or control* or effect) and (workplace or worksite or work site or working place or worker or occupation or employer or employee or corporate or company or business or industry or industrial) and ([exercise or fitness or physical activity or sport] and [program or programme]) and (physical activity or fitness or health or musculoskeletal or back or neck or shoulder or elbow or wrist or knee or upper extremity or lower extremity)</td>
</tr>
<tr>
<td>Personal databases (n = 35)</td>
<td>1/1/80 to 5/26/00; English language; human</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Emmons et al.24 (M = 5)</td>
<td>1. Individually focused activities (2.5 years)</td>
</tr>
<tr>
<td>Gundewall et al.39 (M = 5)</td>
<td>1. Back muscle strength, endurance, and coordination exercises (13 months; 2-6 times per week, 20 min)</td>
</tr>
<tr>
<td>Pritchard et al.26 (M = 5)</td>
<td>1. Exercise program: moderate, unsupervised, aerobic exercise; diet unchanged (12 months; 3 times per week, 30 min)</td>
</tr>
<tr>
<td>Oden et al.17 (M = 2)</td>
<td>1. Aerobics, walking, jogging, bicycle ergometer (24 weeks, 3 times per week)</td>
</tr>
<tr>
<td>Kerr and Vos16 (M = 5)</td>
<td>1. Endurance, strength, and flexibility exercises (12 months, 1 time per week, 60 min)</td>
</tr>
<tr>
<td>Grönningsäter et al.15 (M = 6)</td>
<td>1. Acrobics, strength, flexibility, and relaxation exercises (10 weeks, 2 times per day, 55 min, 3 times per week)</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
</tr>
<tr>
<td>Härma et al.</td>
<td>Jogging, running, swimming, skiing, walking, and gymnastics (4 months, 2–6 times per week)</td>
</tr>
<tr>
<td>Grandjean et al.</td>
<td>Walking, jogging, cycling (24 weeks, 3 times per week, 20–60 min)</td>
</tr>
<tr>
<td>Oja et al.</td>
<td>Walking and cycling to and from work at self-selected speed (10 weeks, daily)</td>
</tr>
<tr>
<td>Helfon et al., Rosenfeld et al., Raskin et al.</td>
<td>Stretching, relaxation, muscular strength, aerobic exercise (7 months, 5 times per week, 15 min)</td>
</tr>
<tr>
<td>Hilyer et al.</td>
<td>Flexibility program (6 months, daily, 30 min)</td>
</tr>
</tbody>
</table>
TABLE 2—(Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Posttest/ follow-up</th>
<th>Study population (number used for analysis)</th>
<th>Outcome measures and method of measurement</th>
<th>Adherence</th>
<th>Results</th>
</tr>
</thead>
</table>
| Lee and White (M = 4) | 1. Self-administered program of low-impact aerobic exercise and education (12 weeks, 2 or 3 sessions between weekly classes)  
2. Reference: wait-list group, invited to participate in a second 12-week exercise program | 12, 24, and 48 weeks | 37 female university workers; 35 included in analysis | 1. Physical activity: questionnaire of 1-week activity recall  
2. Fitness: maximum power output, resting HR, 5-minute recovery (bicycle test until HR reached 130 beats/min)  
3. Flexibility: sit and reach test  
4. % body fat: 3 skin folds  
5. BMI  
6. Serum lipids: total cholesterol, HDL, triglycerides, total cholesterol/HDL ratio  
7. Blood pressure | Average attendance of 96 of the 12 sessions | Significant effect on maximum power output (F[1.30] = 8.29, p < 0.001); relevant effect sizes in flexibility in favor of the reference group; no significant effect on physical activity, 5-minute recovery, resting HR, percent of body fat, BMI, serum lipids, blood pressure |
| Gerdle et al. (M = 6) | 1. Coordination, strength, aerobic activities, stretching (1 year, 2 times per week, 60 min)  
2. Reference: no intervention | 1 year | 160 female home care workers; 77 included in analysis | 1. Fitness: VO₂ max (submaximal bicycle test; Astrand protocol)  
2. General health: questionnaire about somatic and psychosomatic complaints  
3. MSD: questionnaire including nine anatomical regions  
4. Body weight | Participation rate was 75 ± 12% | No significant effects on VO₂ max, health, body weight; relevant effect sizes in percentage of low back disorders (no change in the intervention group but an increase in the reference group [from 22 to 27%]) |
| Gamble et al. (M = 3) | 1. Flexibility, soccer, aerobic, and strength exercises (10 weeks, 2 times per week, 1 hour [?])  
2. Reference: no intervention | 10 weeks | 24 ambulance workers; 14 included in analysis | 1. Strength: sit-up, standing broad jump  
2. Flexibility: back and hamstring  
3. Body weight  
4. % body fat: 4 skin folds | Attendance was at least 90% of the organized sessions | No significant effects [strength: 25% and 15% improvement (I) versus 6% and 1% (R); flexibility: 27% (I) versus 10% (R); weight and % body fat: 0–2% change in both groups] |
| Bassey et al. (M = 6) | 1. Unsupervised walking program (12 weeks, 5 times per week, 20–40 min)  
2. Reference: no intervention | 12 and 24 weeks | 580 factory workers; 59 included in analysis (pedometers), 56 included in analysis (HR records) | 1. Physical activity: mechanical pedometers, 24-h body-borne tape recorders  
2. Fitness: HR (self-paced walking test) | One third had nearly or fully achieved the protocol prescribed | No significant effects |

?, unclear or not described specifically; BMI, body mass index; HDL, high-density lipoprotein; HR, heart rate; (I), intervention group; LDL, low-density lipoprotein; M, methodological quality score; MSD, musculoskeletal disorders; (R) reference group; VLDL, very low density lipoprotein; VO₂ max, maximum oxygen consumption.
### TABLE 3. Characteristics of the nonrandomized, controlled trials in terms of the effectiveness on physical activity, fitness, and health

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Posttest/follow-up</th>
<th>Study population (number used for analysis)</th>
<th>Outcome measures and method of measurement</th>
<th>Adherence</th>
<th>Results</th>
</tr>
</thead>
</table>
| Wier et al. 50 (M = 4) | 1. Physical exercise, individualized exercise prescription, short lecture (12 weeks, 3 times per week) | 2–3 years          | All (n = 7) NASA–Johnson Space Center workers; 320 enrolled in intervention; 258 included in analyses | 1. Physical activity: exercise habits questionnaire  
2. Fitness: VO\(_{2}\)\(_{\text{max}}\) (treadmill test; Bruce protocol)  
3. Body weight  
4. % body fat: skin folds  
5. Serum lipids: total cholesterol, HDL, triglycerides | Average compliance was 79% | Significant effects (physical activity: F = 16.66, df = 3.22, p < 0.001; fitness: +2.5 mL/kg/min in comply group versus −2.4 mL/kg/min in noncomply group; weight, % body fat, and serum lipids: groups changed at different rates over time: F = 2.65, df = 15.707, p < 0.001) |
| Blair et al. 35 (M = 3)| 1. Encouragement to initiate/maintain exercise, programs in exercise, smoking cessation, stress management, nutrition, weight control, and blood pressure control (2 years) | 2 years            | 4,300 Johnson & Johnson Company workers (4 intervention companies (n = 2,600), 3 reference companies (n = 1,700); 2,147 included in analysis | 1. Physical activity: 3 methods of self-reported exercise: a) global self-estimate; b) vigorous exercise; c) (moderate, hard, and very hard) activity of past 7 days  
2. Fitness: VO\(_{2}\)\(_{\text{max}}\) (submaximal bicycle test) | ? | Significant effects (energy expenditure: increase of 104% (I) versus 33% (R); VO\(_{2}\)\(_{\text{max}}\) increased 10.4% (I)) |
| Hartig and Henderson 43 (M = 2)| 1. Addition of 3 hamstring stretching sessions  
2. Reference: health screen | 13 weeks           | 298 military infantry basic trainees; 262 included in analysis | 1. Flexibility: knee extension  
2. MSD (lower extremity injuries): weekly reviews | ? | Significant effects (flexibility: change of 7° (I) versus 5° (R); injuries: 25 injuries (I) versus 43 (R)) |
| Genaidy et al. 37 (M = 1)| PRE: strength training progressive resistance exercise (4 weeks, 4 times per week, 20 min)  
TFF: flexibility, strength training: progressive resistance exercise and trunk flexibility exercises (4 weeks, 4 times per week, 35 min) | 5 weeks            | 28 manual handling operations workers; 28 included in analysis | 1. Strength: dynamic and static (procedure of Berger and Ayoub)  
2. Flexibility: sit and reach test, trunk rotation test | ? | Significant effects of the combined program (strength improvement of 86% [dynamic] 59% [static back] 25% [static arm], 23% [static shoulder] flexibility: 11% [low back] 48% [trunk]) |
| Norms et al. 21 (M = 1)| 1. Aerobic training: road running (10 weeks, 3 times per week approximately 30–40 min)  
2. Anaerobic training: weight training (10 weeks, 3 times per week approximately 30 min)  
3. Reference: no formal training | 10 weeks           | 150 male police officers; 77 included in analysis | 1. Fitness: HR (semiautomatic sphygmomanometer) and timed run (1.5-mile run)  
2. General health: questionnaire regarding psychological components of ill health  
3. Blood pressure | ? | Significant effect of the aerobic program on general health; significant effect of the aerobic and anaerobic program on HR and diastolic and systolic blood pressure (in general for all outcome measures: both the aerobic and anaerobic group showed an improvement, whereas the reference group remained stable) |
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Posttest/ follow-up</th>
<th>Study population (number used for analysis)</th>
<th>Outcome measures and method of measurement</th>
<th>Adherence</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cox et al., 20 Shephard and Cox 49 (M = 3)</td>
<td>1. Rhythmic calisthenics, jogging, ball games, lectures (6 months, 3 times per week, 30 min) 2. Reference: no intervention</td>
<td>6 months</td>
<td>1,858 insurance workers (1 test company [n = 1281], 1 reference company [n = 577]); 614 included in analysis (VO2max, HR, test duration), 534 included in analysis (handgrip, flexibility, body composition)</td>
<td>1. Fitness: VO2max, HR, test duration (Canadian Home Fitness Test) 2. Strength: handgrip dynamometer 3. Flexibility: sit and reach test 4. Body weight 5. % body fat: 4 skin folds</td>
<td>Attendance of 2 or more classes per week (regular participants)</td>
<td>Significant differences in VO2max in high adherents; significant decrease in strength in the reference group; flexibility and percentage of body fat improved in all categories but most in adherents; no significant effect on body weight</td>
</tr>
<tr>
<td>Harrell et al. 41 (M = 1)</td>
<td>1. Stretching and strengthening exercises, aerobic training (9 weeks, 3 times per week, 60 min) 2. Reference: continued usual training</td>
<td>9 weeks</td>
<td>1,504 police trainees (25 sites); 1,504 included in analysis</td>
<td>1. Fitness: VO2max (submaximal bicycle test; Åstrand protocol) 2. Strength: maximal bench press, sit ups 3. Flexibility: sit and reach test 4. % body fat: 3 skin folds</td>
<td>?</td>
<td>Significant effect on fitness, general strength, flexibility and percentage of body fat; no significant effect on upper body strength [fitness change of 21.5% (I) versus 13.4% (R); sit ups: 26.0% (I) versus 16.4% (R); flexibility: 8.3% (I) versus 7.4% (R); % fat: −5.6% (I) versus −1.2% (R)]</td>
</tr>
<tr>
<td>Fisher and Fisher 36 (M = 2)</td>
<td>1a. Group activities: aerobic dance, exercise, weight training, swimming (6 months, 3 times per week, 45 min) 1b. Individual activities: walking, swimming, cycling, jogging, tennis, racquetball (6 months, 5 times per week, 1–2 hours) 2. Reference: no intervention</td>
<td>6 months</td>
<td>65 college faculty, staff, and administrative workers; 65 included in analysis</td>
<td>1. Body weight 2. % body fat 3. Serum lipids, total cholesterol, HDL, LDL, cholesterol/HDL ratio, triglycerides 4. Blood pressure</td>
<td>?</td>
<td>Significant effect on body weight, HDL, total cholesterol/HDL ratio, and triglycerides; no significant effect on percentage of body fat, total cholesterol, LDL, and blood pressure</td>
</tr>
<tr>
<td>Ostwald 46 (M = 1)</td>
<td>1. Mild: seminar, monthly newsletter on exercise and nutrition (3 months) 2. Moderate: further interpretation of test results, physical examination, maximal treadmill exercise test, access to exercise facility (3 months) 3. Intensive: individual explanation of test results; individual exercise prescription; organized, supervised, aerobic exercises (12 weeks, 3 times per week) 4. Reference: no intervention</td>
<td>5 months</td>
<td>Experimental company: n = 261 responded to survey; n = 167 volunteered to participate in intervention; reference company: n = 343 of n = 536 responded to survey; 421 included in analysis</td>
<td>1. Physical activity: self-reported (vigorous) exercise practices 2. Body weight 3. Blood pressure 4. Serum lipids: cholesterol, HDL, HDL/cholesterol ratio, triglycerides</td>
<td>?</td>
<td>Significant differences in body weight in favor of the intensive intervention program; no significant differences among the three intervention groups in physical activity, total cholesterol, HDL, HDL/cholesterol ratio, triglyceride levels, and blood pressure</td>
</tr>
</tbody>
</table>
### TABLE 3. — (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Posttest/ follow-up</th>
<th>Study population (number used for analysis)</th>
<th>Outcome measures and method of measurement</th>
<th>Adherence</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skargren and Oberg⁹ (M = 4)</td>
<td>1. Strength and cardiovascular exercises (8 weeks, 2 times per week, 45 min) 2. Reference: no intervention</td>
<td>Before and after the two exercise periods</td>
<td>106 nurses and nursing aides; included in analysis during exercise periods: 86 (questionnaire), 74 (VO(_2) max), 70 (strength); included in analysis during control periods: 78 (questionnaire), 58 (VO(_2) max), 55 (strength)</td>
<td>1. Fitness: VO(_2) max (submaximal bicycle test; Astrand protocol) 2. Strength: isokinetic dynamometer, knee flexion 3. General health: questionnaire about psychosomatic complaints 4. MSD: questionnaire about symptoms in 7 areas</td>
<td>Regular participants attended at least 8 times</td>
<td>No significant effects in VO(_2) max or number of MSD or psychosomatic symptoms (VO(_2) max: change of 1.3 ml/kg/min [exercise periods] versus 0.7 ml/kg/min [reference periods]; MSD [n]: -0.4 [exercise periods] versus -0.1 [reference periods]; a higher increase in strength during exercise periods [+1 nm] than during reference periods [-3 nm])</td>
</tr>
<tr>
<td>Pavet et al.⁴⁷ (M = 1)</td>
<td>1. Circuit weight training exercise (12 weeks, 3 times per week) 2. Reference: no intervention</td>
<td>12 weeks</td>
<td>350 navy and marine corps men; 245 included in analysis</td>
<td>1. General health: questionnaire consisting of a physical symptoms and a psychological distress scale</td>
<td>?</td>
<td>No significant effect (change from 1.79 to 1.84 [L]) versus a change from 19.5 to 1.84 [H]</td>
</tr>
</tbody>
</table>

<unlabeled data>

M, methodological quality score; MSD, musculoskeletal disorders; (I), intervention group; (R), reference group; VO\(_2\) max, maximum oxygen consumption.

### Conclusion

There is strong evidence.

### Physical Activity Programs

#### Cardiorespiratory Fitness

Three high-quality RCTs (total eight NCTs, five RCTs, three NCTs) were identified. The first high-quality RCTs evaluated the effect of WPAP at both the midpoint and at the end of the intervention and reported that participants had significantly increased their exercise behavior compared with the reference condition. The other high-quality RCTs evaluated the effect of WPAP at the end of the intervention, compared with the reference and the diet condition. The other high-quality RCTs evaluated the effect of WPAP at the end of the intervention, compared with the reference and the diet condition.

#### Muscle Strength

One high-quality RCT and three low-quality RCTs, all of high quality, and four NCTs, all of low quality, were identified. Results of the RCs were inconclusive. In contrast, Lee et al. found significant muscle strength increases in the experimental group. No significant differences between experimental and control groups were found.

#### Muscle Flexibility

Muscle flexibility was assessed using the sit-and-reach test. The remaining RCTs did not find an effect on muscle flexibility. Conclusion: there is inconclusive evidence.

#### Physical Activity

Five RCTs and three NCTs, all of high quality, and four low-quality RCTs and six NCTs, of low quality, were identified. The outcomes of the high-quality RCTs showed a positive effect. The outcomes of the low-quality RCTs showed a significant increase in maximum oxygen consumption. Conclusion: there is inconclusive evidence.

### Effectiveness of Worksite Physical Activity Programs

The studies evaluated the following outcome measures: physical activity, physical fitness, muscle strength and endurance, muscle flexibility, muscle strength and endurance, muscle flexibility, muscle strength, and muscle flexibility. Physical fitness was defined as health-related fitness, including cardiorespiratory fitness, musculoskeletal fitness, muscle strength and endurance, muscle flexibility, muscle strength, and muscle flexibility. Physical activity was measured using self-report questionnaires and objective measures such as accelerometers and pedometers. The outcomes of the high-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the low-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the low-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance. The outcomes of the mixed-quality RCTs showed a significant increase in muscle strength and endurance.
the intervention and reference group were found. In addition, Ruskin et al.\(^\text{48}\) reported no effect of their physical activity program on handgrip strength. **Conclusion**: there is inconclusive evidence.

### Body Weight and Body Composition

Six RCTs,\(^\text{15,22,25,26,38,45}\) three of which were of high quality,\(^\text{15,26,38}\) and four NCTs,\(^\text{20,36,46,49,50}\) evaluated the effectiveness on body weight. Two high-quality

**TABLE 5.** Studies demonstrating a positive effect, no effect, or a negative effect of the worksite physical activity program per outcome measure

<table>
<thead>
<tr>
<th>Outcome measure (number of studies identified)</th>
<th>Positive effect</th>
<th>No effect</th>
<th>Negative effect</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical activity (n = 8)</td>
<td>18, 24, 26, 35, 50</td>
<td>23, 34, 46</td>
<td>—</td>
<td>Strong evidence</td>
</tr>
<tr>
<td>Cardiorespiratory fitness (n = 16)</td>
<td>15, 16, 17, 20/49, 21, 23, 25, 35, 41, 42, 45, 50</td>
<td>19, 23, 34, 38, 42, 48</td>
<td>—</td>
<td>Inconclusive evidence</td>
</tr>
<tr>
<td>Muscle flexibility (n = 8)</td>
<td>20, 37, 41, 43, 44</td>
<td>22, 48</td>
<td>23</td>
<td>Inconclusive evidence</td>
</tr>
<tr>
<td>Muscle strength (n = 8)</td>
<td>19, 20, 37, 39, 41, 42</td>
<td>22, 41, 48</td>
<td>—</td>
<td>Inconclusive evidence</td>
</tr>
<tr>
<td>Body weight (n = 10)</td>
<td>25, 26, 36, 46, 50</td>
<td>15, 20/49, 22, 38, 45</td>
<td>—</td>
<td>Inconclusive evidence</td>
</tr>
<tr>
<td>Body composition (n = 10)</td>
<td>17, 26, 20/49, 41, 50</td>
<td>22, 23, 25, 36, 42</td>
<td>—</td>
<td>Inconclusive evidence</td>
</tr>
<tr>
<td>General health (n = 7)</td>
<td>15, 21</td>
<td>16, 19, 38, 42, 47</td>
<td>—</td>
<td>Limited evidence</td>
</tr>
<tr>
<td>Fatigue (n = 2)</td>
<td>40, 42</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Musculoskeletal disorders (n = 7)</td>
<td>15, 38, 39, 42, 43</td>
<td>19, 44</td>
<td>—</td>
<td>Strong evidence</td>
</tr>
<tr>
<td>Blood serum lipids (n = 7)</td>
<td>36, 50</td>
<td>15, 23, 25, 36, 45, 46</td>
<td>—</td>
<td>No evidence</td>
</tr>
<tr>
<td>Blood pressure (n = 5)</td>
<td>21</td>
<td>15, 23, 36, 46</td>
<td>—</td>
<td>No evidence</td>
</tr>
</tbody>
</table>

Numbers listed refer to the reference numbers. High-quality randomized, controlled trials are indicated by boldface type.
RCTs, 15, 38 did not find an effect of the program on body weight, whereas Pritchard et al. 26 found a significant difference between groups in change of body weight. With the exception of the study of Grandjean et al., 25 the remaining low-quality RCTs 22, 45 did not find an effect on body weight. Conclusion: there is inconclusive evidence.

Body composition was defined as the percentage of body fat and/or body mass index. One high-quality RCT, 26 five low-quality RCTs, 17, 23, 25, 42 and four NCTs, 20, 36, 41, 49, 50 were identified evaluating body composition. Pritchard et al. 26 found significant changes in body mass index and total fat mass in favor of both the diet and the exercise group. Oden et al. 17 reported a significant favorable effect of the exercise program, whereas no effect was found in the remaining four trials. 22, 23, 25, 42 Conclusion: there is inconclusive evidence.

General Health

Three high-quality RCTs, 15, 16, 38 one low-quality RCT, 42 and three NCTs, 19, 21, 47 were identified. Although Grönningssäter et al. 15 found a nonsignificant tendency toward a reduction in general health complaints in the intervention group compared with the reference group, effect sizes between the groups appeared to be relevant. In contrast, the remaining two high-quality RCTs showed no influence of the intervention on general health. 16, 38 Conclusion: there is inconclusive evidence.

Fatigue

Two RCTs, both of low quality, 40, 42 were identified. Härma et al. 42 showed relevant effect sizes in fatigue between the study groups. In addition, Halfon et al. 40 reported a significantly greater increase of mental and physical fatigue in the reference group compared with the intervention group. Conclusion: there is limited evidence.

Musculoskeletal Disorders

Five RCTs, 15, 38, 39, 42, 44 and two NCTs, 19, 43 were identified. Three RCTs were of high quality. 15, 38, 39 At least Gerdl et al. 38 could not find statistically significant changes in prevalence or intensity of musculoskeletal complaints, effect sizes between the two groups in prevalence of low back pain were considered to be relevant. Grönningssäter et al. 15 found a significant effect of the exercise intervention on both neck and back pain. Finally, Gundewall et al. 30 also observed a positive effect of the intervention on back pain. Conclusion: there is strong evidence.

Blood Serum Lipids

Four RCTs, 15, 23, 25, 45 one of them of high quality, 15 and three NCTs, 26, 46, 50 were identified. None of the RCTs found either significant or relevant effect sizes in serum lipids between the study groups. Conclusion: there is no evidence.

Blood Pressure

One RCT of high quality, 15 one RCT of low quality, 23 and three NCTs, 21, 36, 46 were identified. Grönningssäter et al. 15 showed no significant changes in systolic blood pressure. In addition, with the exception of a significant change in systolic blood pressure in favor of the exercise group after 24 weeks, the study of Lee and White 23 did not show any significant changes in diastolic or systolic blood pressure between pretest and any of the follow-up measurements (12, 24, and 48 weeks). Conclusion: there is no evidence.

DISCUSSION

Effectiveness

The purpose of this review was to draw conclusions regarding the effectiveness of WPAPs on physical activity, physical fitness, and health. Our results indicate that the primary goal of such programs (i.e., enhancing general physical activity levels) is achieved. According to the model of Bouchard et al., 7 which describes the relationship between physical activity, fitness, and health, one would expect that this enhancement of physical activity would result in an improvement of cardiorespiratory fitness. However, no such evidence was found. One plausible explanation might be the fact that enhancement of cardiorespiratory fitness requires quite intensive physical activity (at least three times a week at 40 or 50 to 85% of maximum oxygen uptake reserve for at least 20 minutes), 51 and it is likely that participants in WPAPs do not reach this frequency, intensity, and duration. Unfortunately, adherence to the programs was generally poorly reported in the studies, which made it impossible to verify this assumption.

Compared with the literature, our conclusions do not seem to be in line with those drawn by Dishman et al., 13 who concluded that WPAPs have a small, nonsignificant effect on physical activity. However, differences in conclusions are, in our opinion, the result of the different methods used for reviewing. For example, Dishman et al. 13 performed a quantitative analysis, taking into account the methodological quality of the studies included, whereas we used a qualitative method. Also, the criteria used by Dishman et al. 13 to evaluate the methodological quality of the studies were different from ours. Moreover, the types of interventions evaluated in the Dishman et al. 13 review differed somewhat from those evaluated in the present review. In the present review, only worksite interventions with a primary focus on stimulating the level of physical activity or fitness were included, whereas Dishman et al. 13 included programs with a more comprehensive training regimen as well. Thus, the methods used are of relevance in interpreting the conclusions.

Another important finding of this review is the strong evidence for the effectiveness of WPAPs on reducing musculoskeletal disorders. The literature regarding the associations between physical activity, physical fitness, and low back pain, for example, shows contradictory results. Videman et al. 52, 53 reported that physical activity seems to have a dual role, imposing a positive and negative influence on the spine. In addition, a recent review of epidemiological literature on the relationship between physical activity and musculoskeletal morbidity showed inconsistent results, 54 leaving the question of whether the...
promotion of physical activity could be an attractive additional preventive strategy in reducing musculoskeletal morbidity at the workplace unanswered. Although it is unclear how the structural changes and (musculoskeletal) symptoms are related, this review indicates that the implementation of a WPAP may be a promising component of a strategy aimed at reducing or preventing musculoskeletal disorders.

With the exception of fatigue, we found no (conclusive) evidence that a WPAP has positive effects on other health-related outcomes. With respect to body weight or body composition, our findings seem to be in contrast with the pertinent literature. This contradiction may be explained by the fact that the populations in the studies we reviewed were generally healthy, nonobese employees; therefore, benefits on body weight or body composition would presumably be small. Another plausible explanation for this contradiction can be attributed to the significant increases in physical activity due to WPAPs not being of sufficient magnitude to affect body weight and body composition.

**Methodological Quality of the Studies**

This review shows that the majority of the studies on the effectiveness of WPAPs had methodological shortcomings. Major problems included the lack of a sufficient description of the randomization procedure, blinding of the person performing the measurements, and absence of an intention-to-treat analysis. As several studies have provided empirical evidence that trials with inadequate methodological approaches or incomplete descriptions of procedures, particularly concerning concealment of treatment allocation and blinding, are associated with bias, future investigators should pay attention to these aspects. Finally, it is worth mentioning that almost all of the studies applied self-reported data for the measurement of physical activity and health outcomes and therefore lacked the use of objective, more valid measures. If more objective instruments had been used, results regarding the effectiveness might have been different. However, because there was only one study that applied an objective physical activity measure, we were not able to investigate a possible influence of such measure. Finally, particularly among the NCTs, the description of, or the rate of compliance with, the program was insufficient. In cases in which there was a lower compliance rate with the program than was prescribed, an underestimation of the results might have occurred. Thus, both from the employee’s and the researcher’s perspective, adherence to the intervention should be supported.

**CONCLUSION**

There is strong evidence for a positive effect of a WPAP on physical activity and musculoskeletal disorders, limited evidence for a positive effect on fatigue, and inconclusive or no evidence for a positive effect on cardiorespiratory fitness, muscle flexibility, muscle strength, body weight, body composition, general health, blood serum lipids, and blood pressure. The methodological quality of most studies evaluating the effectiveness of WPAPs is generally poor. Future studies should pay attention to the description and performance of the randomization, blinding of the outcome assessor, compliance, and intention-to-treat analysis.

**REFERENCES**

19. Skargren E, Öberg B. Effect of an exercise program on musculo-


