Effects of workplace-based physical activity interventions on cardiorespiratory fitness: a systematic review and meta-analysis of controlled trials

Naomi L. Burn[1], Matthew Weston[1] Neil Maguire[1], Greg Atkinson[1] and Kathryn L. Weston[1]

[1]School of Health and Social Care, Teesside University, Middlesbrough, United Kingdom

Corresponding author: Naomi Burn, School of Health and Social Care, Teesside University; Address: Teesside University, Middlesbrough, TS1 3BX.

email: n.burn@tees.ac.uk

ORCID:

Naomi Burn: 0000-0001-6189-4210

Matthew Weston: 0000-0002-9531-3004

Neil Maguire: 000-0002-9793-4314

Greg Atkinson 0000-0002-5459-9042

Kathryn L Weston: 0000-0001-5918-6389

Abstract

Background: Cardiorespiratory fitness (CRF) is a strong predictor of all-cause mortality. Physical activity (PA) of at least moderate intensity can improve CRF. Workplaces may provide a relatively controlled setting in which to improve CRF through PA. Limited work has been conducted to quantify the impact of delivering PA in the workplace on CRF.

Aim: To quantify the effects of workplace PA interventions on VO_{2peak} and explore study and participant characteristics as putative moderators.

Methods: Seven databases were searched up to September 2018. Search terms included "workplace", "physical activity" and "intervention". Inclusion criteria were controlled trials where PA of at least moderate intensity was delivered in the workplace and compared to controls or non-active comparators; and CRF measured by actual or predicted VO_{2peak}. Risk of bias was assessed using the PEDro scale. A random effects meta-analysis was conducted with between study variation quantified and then explored for putative predictors with meta-regression. Pooled estimate uncertainty was expressed as 90% confidence intervals (CI) and assessed against our threshold value for clinical relevance of 1 $mL \cdot kg^{-1} \cdot min^{-1}$.

Results: The final dataset consisted of 25 estimates of VO_{2peak} from 12 trials. The pooled mean differences between intervention and control arms was a beneficial improvement of 2.7 $mL \cdot kg^{-1} \cdot min^{-1}$ (90% CI: 1.6 to 3.8 $mL \cdot kg^{-1} \cdot min^{-1}$). The 95% prediction interval ranged from a reduction in VO_{2peak} of -1.1 to an improvement of 6.5 $mL \cdot kg^{-1} \cdot min^{-1}$. Between-study heterogeneity (τau) was ±1.6 $mL \cdot kg^{-1} \cdot min^{-1}$. Meta regression showed longer interventions (3.2 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI: 1.6 to 3.8 $mL \cdot kg^{-1} \cdot min^{-1}$) to have an additive effect and studies with low risk of bias (-2.5 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI -4.0 to -1.0 $mL \cdot kg^{-1} \cdot min^{-1}$), and participants of greater baseline VO_{2peak} (-1.6 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI -3.6 to 0.4 $mL \cdot kg^{-1} \cdot min^{-1}$), and age (-1.4 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI -3.2 to 0.3 $mL \cdot kg^{-1} \cdot min^{-1}$) having a lesser effect. Participant sex (% female) had an additive effect on VO_{2peak} (0.4 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI: -1.6 to 2.4 $mL \cdot kg^{-1} \cdot min^{-1}$).

Conclusion:

Workplace-based PA interventions consisting of at least moderate intensity activity improve CRF. At the present time, we surmise that no single group of employees (e.g. older employees or less fit individuals) can be definitively identified as standing to benefit more from workplace PA interventions than others. This

demonstrates the potential utility of workplace PA interventions for improving CRF in a broad range of healthy employees.

Protocol registration: PROSPERO (registration number: 42017057498).

Key points

- When compared to control, workplace physical activity (PA) interventions consisting of at least moderate intensity activity resulted in a beneficial improvement in cardiorespiratory fitness (as measured by peak oxygen consumption (VO_{2peak})).
- Meta-regression showed longer interventions and studies with higher risk of bias to have additive effects on VO_{2peak}. Further work is needed to quantify the effects of participant characteristics on the impact of workplace PA interventions on cardiorespiratory fitness.
- Currently, it appears that no single group of employees (e.g. older or less fit individuals) stands to benefit more from workplace PA than others. These types of interventions may therefore be useful for improving cardiorespiratory fitness in a broad range of healthy employees.

1. Introduction

Cardiorespiratory fitness (CRF) is an important component of health-related physical fitness and is defined as "the ability of the circulatory, respiratory, and muscular systems to supply oxygen during sustained physical activity" (PA) [1]. CRF is associated with cardiovascular disease (CVD) and all-cause mortality [2] and is perhaps a stronger correlate than total PA [3], hypertension, diabetes mellitus, smoking and obesity [1]. Recent analyses of population-based longitudinal datasets have emphasised the importance of maintaining good CRF, such that a 1 $mL \cdot kg^{-1} \cdot min^{-1}$ increase in CRF (as measured by peak oxygen uptake [VO_{2peak}]), is associated with a 9% relative risk reduction in all-cause mortality [4]. Nevertheless, it has also been reported that CRF levels have declined by 1.6% per decade internationally since 1967 [5]. Lee et al. [1] maintained that CRF is often overlooked as an outcome measure for intervention strategies aiming to promote or deliver PA compared with other risk factors such as total PA and other cardiometabolic health outcomes.

Evidence from randomised controlled trials (RCT) suggests that activity of at least moderate intensity can maintain or improve CRF [6]. Current guidelines recommend the accumulation of 150 minutes of moderate or 75 minutes of vigorous intensity PA, or a combination thereof per week [7, 8] to promote health. Nevertheless, objectively measured compliance with PA guidelines in adults remains low (5%–47%) [9-11]. One setting in which increasing the compliance with PA guidelines could be targeted is the workplace. Because the workplace provides access to a considerable proportion of the adult population, it has been established as a priority setting for health promotion [12, 13]. Given that lack of time, facilities or social support are longstanding and often cited barriers to PA participation in adults [14-16], delivering PA in the workplace may have the potential to reduce some of these barriers by providing convenient access to PA opportunities.

Although there has been debate surrounding the utility of workplace health promotion strategies [17], a desire to reduce healthcare and absenteeism costs and enhance employee health and productivity may present a strong rationale for organisations to implement such strategies [18], particularly since a physically active workforce has been associated with increased productivity and reduced absenteeism [19, 20]. Recent cross-sectional analyses from both the United States [21] and United Kingdom [22] have found that the onsite provision of facilities to support PA, such as structured exercise programmes, are associated with a greater likelihood of achieving PA guidelines. The relationship between CRF and important workplace outcomes has recently been explored, albeit cross-sectionally in a limited number of studies. Although higher CRF has been associated with higher productivity [23], subsequent studies have found no significant relationships between CRF and productivity or sickness absence [24] and sickness presenteeism [25]. Despite these conflicting outcomes, higher levels of CRF have been linked to higher health-related quality of life [26] and lower risk of depressive symptoms [27], burnout and stress [28] which may be important outcomes from an organisational perspective.

Although previous systematic reviews support the use of workplace PA interventions to increase PA levels [29-34], whether this consistently translates into improvements in CRF is unclear. The last meta-analysis examining the impact of workplace PA programmes on CRF reported an increase of 3.5 $mL \cdot kg^{-1} \cdot min^{-1}$ in VO_{2peak} following workplace PA programmes conducted up to late 2007 [35]. Though encouraging, this finding must be treated with caution due to the inclusion of both controlled and non-controlled trials in their analysis. Due to a lack of comparison group, non-controlled trials can be susceptible to confounding or regression to the mean [36]. Moreover, very broad eligibility criteria led to the inclusion of various intervention designs, including those where increasing PA was the primary outcome (e.g. behaviour change interventions such as the provision of leaflets on the benefits of PA), those where the effect of PA on a primary outcome was examined (e.g. where PA is delivered in the workplace to impact on a health outcome) and those where PA was part of a multicomponent intervention. As these designs and subsequent intervention outcomes are distinctly different, it is questionable whether the effects of such different strategies should be grouped and meta-analysed together [37]. Lastly as the meta-analysis by Conn et al. [35] was conducted a decade ago, new relevant literature is likely to be available [38] therefore an update on the effectiveness of workplace PA programmes on CRF is needed. Accordingly, we aimed to use random effects meta-analysis to quantify the effect of workplace-based PA interventions consisting of at least moderate intensity activity on CRF, and explore the modifying effects of study and participant characteristics.

2. Methods

This review was carried out in accordance with the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA) guidelines [39]. The review methodology was prospectively registered with PROSPERO (registration number: 42017057498).

2.1 Eligibility criteria

We sought to identify peer reviewed journal articles describing workplace PA interventions; defined as interventions where PA is both prescribed and delivered to employees at (or commencing from) their workplace. The inclusion criteria were as follows:

A. Population: working adults (aged 16+ years).

B. Intervention: interventions where PA of at least moderate intensity (e.g. ≥64% of maximum heart rate (HR_{max}) or ≥43% VO_{2peak}, as defined by Garber et al. [40]) was prescribed and delivered to employees in the workplace (e.g. group exercise classes held at the workplace) or commenced from the workplace (i.e. in the case of walking interventions).

C. Comparator: non-treatment control groups (e.g., usual activity), or non-active comparators (e.g. leaflets on the health benefits of PA).

D. Outcome: CRF (as measured by actual or predicted VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)).

E. Study design: randomised and non-randomised controlled trials.

Exclusion criteria were as follows:

F. Studies including patient groups, retirees or unemployed persons.

G. General PA promotion or environmental change interventions where PA was not specifically delivered in the workplace.

H. Interventions designed solely to decrease sedentary behaviour.

I. Multicomponent interventions (e.g. interventions with concurrent PA and dietary components), unless one group received only the PA intervention and data could be extracted for only this group.

J. Articles not in English.

K. Articles were excluded if each of the FITT components (frequency of PA, target intensity, time spent performing PA and type of PA) [41] were not adequately described.

2.2 Deviations from the pre-registered review protocol

During the systematic review process, it was necessary to deviate from the protocol initially registered on PROSPERO on several occasions. Justification for these deviations is provided herein. Following initial protocol registration, it became apparent that we did not have the facilities or funding to translate non-English language papers. As such, 'Articles not in English' was added to the exclusion criteria (criteria J). Due to the substantial biases that can arise in uncontrolled study designs [42], we elected to include only controlled trials in our review (reflected in criteria C and E), thus excluding single arm studies. During early phases of the review screening process, we became acutely aware that the terms 'physical activity' and 'exercise' are operationalised in very different ways across sport and exercise medicine literature. Accordingly, it was necessary to provide a clear definition of what PA constituted for inclusion in our review (e.g. the definition from Garber et al.,[40], described in criteria B). Due to poor reporting of what workplace-based PA interventions actually comprised, we updated our exclusion criteria such that articles which did not adequately describe the frequency of PA sessions, target intensity, time spent performing PA and type of PA (e.g. FITT components [41]) would not be included in our review (criteria K). On initial protocol registration, we intended to meta-analyse a range of physical fitness, cardiometabolic health and wellbeing outcomes. Due to an insufficient number of studies reporting each of the other outcomes and significant heterogeneity in the assessment methods used across different outcomes which precluded meta-analysis [43], we elected to reduce outcome variable selection to CRF only (criteria D). Lastly, although our protocol specified that risk of bias would be assessed, we did not specify *a priori* that the PEDro risk of bias tool [44] would be used.

2.3 Information Sources

An electronic search of seven databases was conducted from the beginning of coverage dates for each individual database (exact coverage detailed below) to 3rd September 2018; MEDLINE (1966 to September 2018), PubMed (1997 to September 2018), PsycINFO (1597 to September 2018), SPORTDiscus (1892 to September 2018), CINAHL (1937 to September 2018), EMBASE (1947 to September 2018) and Scopus (1970 to September 2018).

2.4 Literature Search

The complete search strategy is presented in Table 1 and the strategy was modified according to the indexing systems of each database. Reference lists of all included papers were hand searched for relevant papers by one reviewer (NB).

2.5 Study Selection

Search results were imported from individual databases to reference management software (Mendeley Desktop v1.17.10) and duplicates deleted. The details of each study (authors, date of publication, title, journal title, volume, issue, page numbers and reference type) were then imported into a spreadsheet. Search results were screened for relevance by one reviewer (NB) and all irrelevant articles were removed at this stage. Title and abstracts were retrieved and independently screened against inclusion criteria by two reviewers (NB, NM). Full text articles were independently screened for inclusion by two reviewers (NB, NM). Any disputes were settled by a third reviewer (KW).

2.6 Data extraction

One reviewer (NB) extracted data from all included studies using a specifically developed spreadsheet. Accuracy of all data extraction was confirmed via intra- and inter-individual reassessments of data extraction (NB, NM). The reviewers were not blinded to the authors or journals during data extraction. Details of the PA intervention extracted included the PA modality, session frequency, duration, intensity, intervention length and location. The age, sex, job role and sample sizes for intervention and control groups were also recorded. The primary outcome data taken and used in the analyses included the intervention effect (intervention minus control) on CRF as measured by relative VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$). In the case where a study included two intervention groups and one control group, the outcome data from the two intervention groups were combined to create a single pair-wise comparison [36].

2.7 Risk of bias assessment of individual studies

Two reviewers (NB, NM) independently assessed risk of bias using the PEDro scale [44]. In case of disagreement, consensus was achieved by discussion and disputes settled by a third reviewer (KW). Items in the PEDro scale include; specified eligibility criteria (external validity); demonstrated random allocation; allocation concealment; similar groups at baseline, blinding of participants, facilitators and assessors; follow up of ≥85% of participants; intervention received as intended or intention to treat analysis and reporting of between-group statistical comparisons, point measures and measure of variability. The maximum score for the PEDro scale is 10 points (low scores indicate higher risk of bias), because item 1 (specified eligibility criteria) is not included in the calculation of the total score. Supplementary file 1 details the location (page/ table or figure number) in each of the included studies for each PEDro assessment criteria.

2.8 Small Study Bias

Small study bias was assessed using visual inspection of the funnel plot and Egger’s test to evaluate asymmetry [45].

2.9 Data synthesis

A random effects meta-analysis was conducted using the Comprehensive Meta-analysis Software Package, Version 3 for Windows (Biostat Company, Englewood, NJ, USA). The primary statistic for extraction was the intervention effect (mean change in intervention minus mean change in control). Confidence intervals, for the mean intervention effect were calculated. Between-study heterogeneity was quantified using tau and the prediction interval approach.

2.10 Outcome statistics

We expressed the uncertainty in the estimates of effects on VO_{2peak} as 90% confidence intervals (CI) and then in relation to our threshold value for clinical relevance. The smallest clinically relevant effect on VO_{2peak} was taken from recent work suggesting a 1 $mL \cdot kg^{-1} \cdot min^{-1}$ increase in VO_{2peak} is associated with a 9% relative risk reduction in all-cause mortality [4]. This criterion was applied post-registration of the initial protocol on PROSPERO. Between study heterogeneity was indicated by tau [46-48].

Meta-regression analyses were conducted to explore the influence of putative moderator variables on the intervention effect. Here, we selected study (intervention length and PEDro risk of bias score) and participant characteristics (baseline VO_{2peak}, age and sex (percentage female participants)) that could reasonably influence the overall effect of PA on VO_{2peak}. The modifying effects of these five variables were calculated as the effect of two SDs (i.e. the difference between a typically low and a typically high value) [49]. Meta-regression was performed only when there were ≥10 data sets [36]. We derived a 95% prediction interval [50], providing a plausible range for the expected pooled mean effect (intervention minus control) on VO_{2peak} in a future workplace PA intervention conducted in similar settings.

3. Results

3.1 Study Selection

Figure 1 is the PRISMA diagram representing the flow of studies through the review and reasons for exclusion. Two studies were excluded during data extraction, one because insufficient data was presented in the paper to allow for inclusion, and no response was received from study authors upon request for further information and one because it presented follow up data from another included study.

3.2 Study and Participant Characteristics

The final dataset for changes in VO_{2peak} consisted of 25 estimates from 12 studies [51-62]. Study and participant characteristics from the eligible studies are shown in Table 2. VO_{2peak} estimates from a total of 733 participants were included in the meta-analysis. Participants in the individual studies were aged 25 to 67 years (mean: 41 years) and from a range of both sedentary office based and manual labour intensive occupations including office workers (n=3), "insurance company workers" (n=2), "Westinghouse corporation workers" (n=1), "Ford Motor Company workers" (n=1), pharmaceutical company workers (n=1), construction workers (n=1), care workers (n=1), cleaners (n=1), and poultry processing workers (n=1). Studies were conducted in Denmark (n=4), United Kingdom (n=2), Sweden, (n=1), Norway (n=1), Netherlands (n=1), Turkey (n=1), Canada (n=1) and the United States of America (n=1). The length of interventions ranged from 8 to 52 weeks (median: 12 weeks). Four studies prescribed PA twice per week, six studies prescribed PA thrice weekly and two studies prescribed PA five times per week. PA session length was variable and ranged from ≈4 minutes to up to 60 minutes. The majority of PA interventions involved aerobic activity (n=8), however some studies incorporated resistance training (n=2) and multi-component exercise training (e.g. combination of aerobic and resistance training in one intervention group) (n=2).

3.3 Risk of bias assessment

Risk of bias assessments of all included studies is presented in Table 3. The highest risks of bias were for blinding of participants, facilitators and assessors.

3.4 Effect of workplace PA interventions on VO_{2peak}

In our meta-analysis, we found a beneficial effect on VO_{2peak} following workplace PA interventions when compared to controls (2.7 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI 1.6 to 3.8 $mL \cdot kg^{-1} \cdot min^{-1}$; 733 participants, 12 studies, Figure 2). Between-study heterogeneity (τ) was ±1.6 $mL \cdot kg^{-1} \cdot min^{-1}$ (90% CI 1.1 to 2.7 $mL \cdot kg^{-1} \cdot min^{-1}$). The 95% prediction interval was -1.1 to 6.5 $mL \cdot kg^{-1} \cdot min^{-1}$. Egger's coefficient was 1.51 (95% CI -1.40 to 4.41; $p = 0.28$). The non-significant Eggers test coupled with the balanced nature of the funnel plot presented in Figure 3, suggest there is no evidence of small study bias.

Meta-regressions were undertaken with PEDro risk of bias score, intervention length, mean age, mean baseline VO_{2peak} and sex (percentage female) as selected predictors (Table 4). The relevant scatter plots with fitted meta-regression slopes are shown in Figures 4-8. Meta-regression analysis revealed an additional effect for typically longer interventions (3.2 $mL \cdot kg^{-1} \cdot min^{-1}$; 1.3 to 5.1 $mL \cdot kg^{-1} \cdot min^{-1}$), and lesser effects for studies with lower risk of bias (-2.5 $mL \cdot kg^{-1} \cdot min^{-1}$; -4.0 to -1.0 $mL \cdot kg^{-1} \cdot min^{-1}$), greater baseline VO_{2peak} (-1.6 $mL \cdot kg^{-1} \cdot min^{-1}$; -3.6 to 0.4 $mL \cdot kg^{-1} \cdot min^{-1}$) and older participants (-1.4 $mL \cdot kg^{-1} \cdot min^{-1}$; -3.2 to 0.3 $mL \cdot kg^{-1} \cdot min^{-1}$). Percentage female (sex) had an additive effect on VO_{2peak} (0.4 $mL \cdot kg^{-1} \cdot min^{-1}$; 90% CI: -1.6 to 2.4 $mL \cdot kg^{-1} \cdot min^{-1}$).

4. Discussion

This study presents a quantitative evaluation of the impact of workplace-based PA interventions on CRF in healthy employees. Establishing the effectiveness of workplace PA programmes on CRF is particularly timely in light of recent evidence suggesting that CRF levels have declined by 10.8% (-4.2 $mL \cdot kg^{-1} \cdot min^{-1}$) in working populations in the last two decades [63]. The results of our random effects meta-analysis showed that when compared to control, workplace PA interventions result in a beneficial improvement in VO_{2peak} of 2.7 $mL \cdot kg^{-1} \cdot min^{-1}$. Given that a 1 $mL \cdot kg^{-1} \cdot min^{-1}$ increase in VO_{2peak} has been associated with a 9% relative risk reduction in all-cause mortality [4]; we surmise that the magnitude and precision of improvements in VO_{2peak} induced by workplace-based PA interventions are clinically meaningful. The prediction interval, which describes a plausible range of mean treatment effects in a hypothetical future workplace PA intervention in similar settings, ranged from a reduction in VO_{2peak} of -1.1 to an improvement of 6.5 $mL \cdot kg^{-1} \cdot min^{-1}$. It is possible that this range in plausible effects is explained by the broad range of PA prescriptions included in our review (Table 2). Further work is required to elucidate the optimum prescription of PA for workplace interventions aimed at improving CRF levels.

The results of our meta-analysis are lower than those reported in a previous meta-analysis [35] where an increase of 3.5 $mL \cdot kg^{-1} \cdot min^{-1}$ was found following more broadly defined workplace PA interventions (including the promotion of general PA). However their results should be interpreted with caution due to the inclusion of non-controlled trials and a broad range of intervention formats in the meta-analysis. Despite observing an improvement in VO_{2peak} in our meta-analysis, there was also heterogeneity in the effect (τ = 1.6 $mL \cdot kg^{-1} \cdot min^{-1}$), indicating variation in the intervention effect between studies. To explore this variation further, meta-regression was conducted using intervention (intervention length and PEDro risk of bias) and participant (baseline VO_{2peak}, age and sex) characteristics. Study characteristics (intervention length and PEDro risk of bias) explained some of the heterogeneity of the effect (r^2= 26% and 28%, respectively, Table 4). Although there was an additional improvement in VO_{2peak} for typically longer interventions, one study with a long intervention length and large intervention effect (Figure 4) was an outlier and may have impacted on the slope of the meta-regression. Therefore, this result should be interpreted with caution. There was a lesser effect for studies with lower risk of bias. This result was not unexpected given that methodological characteristics of studies are thought to have a substantial impact on treatment effect estimates in meta-analysis [64]. Given that the modifying effect of participant characteristics explained only a very limited proportion of the heterogeneity of the effect (r^2= 0% for baseline VO_{2peak}, 14% for age and 0% for sex) a conservative approach has been taken in the interpretation of this result. We surmise that given the data available from our included studies, no single group (e.g. older employees or less fit individuals) can be definitively identified as standing to benefit more from workplace PA interventions at this stage. Further work is needed to quantify the effects of participant characteristics on the impact of workplace PA interventions on CRF. One other possible explanation for the heterogeneity in the effect could be the measurement error associated with submaximal prediction of VO_{2peak}. All but two [54, 60] of the included studies used submaximal estimates of VO_{2peak} (either via step tests, submaximal cycle ergometer tests, or treadmill walking tests, Table 2). This is unsurprising given the limited feasibility of performing maximal oxygen consumption tests in the workplace. Alternatively, the heterogeneity may be attributed to the range of PA prescriptions used and/ or the extent to which the fidelity of the intervention was upheld (e.g. whether the intervention was implemented as intended across all participants [65,66]. However, these variables could not be included in a meta-regression due to inconsistencies in reporting practices in the individual studies. Another possible explanation for the heterogeneity in the effect could be the various study designs used in the included studies (e.g. controlled trials vs. RCTs vs. cluster RCTs). In a cluster RCT randomisation is likely to occur at the workplace rather than individual level whereby one workplace would act as an intervention group and a separate workplace would act as the comparator. A cluster RCT design may avoid contamination in the control group, which can be an issue in PA trials when control participants modify their behaviour despite requests to maintain their usual activity pattern [67]. The issue of contamination was discussed by the authors of one of our included studies [51] who employed a traditional RCT design. Here, although participants were blinded to their group allocation at baseline, all were recruited from the same workplace and may therefore have established their group allocation through informal discussions with colleagues.

As demonstrated in Table 2, a range of occupations were targeted in our included studies. Although five of the twelve included studies targeted desk-based employees, the remaining studies included a range of manual labour intensive occupations such as construction workers and cleaners. This suggests that workplace PA interventions could be implemented in a range of occupational environments. All of the studies included in this meta-analysis were from high or upper-middle income countries with the majority of the included studies from European countries and the remaining from North America. It is likely that the implementation of such programmes are highly context dependent [68] due to a variety of factors such as workplace cultural norms or climate. Therefore the generalisability of individual workplace based PA interventions to diverse geographical locations may be limited at this time.

It is important to acknowledge that our study is not without limitations. Although the review protocol was prospectively registered with PROSPERO, it was deemed necessary to deviate from the protocol on occasion. While all deviations are available in the public domain of PROSPERO and the justification and ramifications transparently described in our methods section, this nonetheless could have introduced bias into the review process. Our initial intention was to meta-analyse a broad range of physical fitness, cardiometabolic health and wellbeing outcomes. However due to significant heterogeneity in outcome measures reported in individual studies and low numbers of studies reporting some outcomes, inclusion of such a wide range of outcome variables was not possible. Given that such heterogeneity could violate the assumptions of a meta-analysis [43] a cautious approach was taken in outcome variable selection. A more comprehensive assessment of health, fitness and wellbeing variables in future intervention studies would assist in evidence-based programming decisions. It has also been suggested that a standardised approach to outcome variable selection for health-related effectiveness interventions [69] could aid the selection of pertinent outcomes for workplace PA interventions. Although this approach has merit and could aid future systematic reviews and meta-analyses in the area, in the case of workplace PA interventions it is sometimes necessary that a more pragmatic approach is taken when selecting outcome variables. Indeed, Bauman & Nutbeam [70] recommend that during intervention development, key stakeholders (e.g. employees or management representatives) should be consulted to maximise the reach, uptake and usefulness of such interventions. Relying solely on a core set of outcome variables may therefore be counterproductive as this could neglect the needs of key stakeholders. To be eligible for inclusion in our meta-analysis PA sessions had to be delivered, or at least begin, from the participants workplace (i.e. in the case of walking interventions). This led to the exclusion of studies where participants were recruited from their workplace but PA was conducted elsewhere. The decision was made to exclude such studies because one rationale for delivering PA in the workplace is that there is potential to reduce commonly cited barriers to PA participation, such as lack of time and access to facilities [71]. If participants are required to travel to an external location in order to participate in “workplace” PA, it was deemed that this was no longer a workplace PA intervention because this could result in further exacerbation of such barriers. Only interventions prescribing PA of at least moderate intensity were included in our meta-analysis; which led to the exclusion of studies delivering PA of lower intensities. There is evidence to suggest that PA or exercise training of at least moderate intensity can elicit adaptations in various markers of physical fitness, including CRF [6]. Therefore PA prescriptions of less than moderate intensity may not be expected to have an impact on CRF in healthy populations. The use of the PEDro scale to assess the risk of bias in the included studies could also be considered a limitation of our study. The use of summary scales to assess risk of bias have previously been described as misleading because not all scale elements may be relevant in all contexts, but are still weighted equally [72]. Nonetheless, we believe that the PEDro scale is a useful tool which allows the exploration of the linear negative association between risk of bias and size of treatment effect [73].

We have several recommendations for future workplace-based PA interventions, which in turn will aid further systematic reviews and meta-analyses on the topic. As previously mentioned, a cautious approach was taken in outcome variable selection due to significant heterogeneity in outcome assessment and reporting methods in individual studies. In future intervention studies we suggest that a wide range of physical fitness, cardiometabolic health and wellbeing outcomes are assessed, with consideration of pertinent outcomes for participating organisations. For this to be possible, we suggest that employees and senior management teams are involved in the design and implementation of workplace-based PA interventions (e.g. via a formative evaluation pre-intervention) [70]. Further, a detailed description of the PA prescription (frequency, intensity, time and type) had to be presented for studies to be included in our review. Poor reporting of the PA prescription resulted in the exclusion of a number of studies. This was possibly not surprising, given that inadequate reporting of intervention trials remains commonplace in PA and exercise trials [66]. As such, we support calls for standardised reporting of PA and exercise intervention protocols and intervention fidelity quantification [66, 74-77]. This should include the presentation of PA and exercise training programming variables (e.g. frequency, intensity, time and type of PA) as well as information about the extent to which the fidelity of the intervention was upheld. Adherence to reporting checklists such as the Template for Intervention Description and Replication (TIDieR) checklist [76] or Consensus on Exercise Reporting Template (CERT) guidelines [78] are also recommended. From a practical standpoint, thorough intervention reporting is particularly important when recruiting organisations to workplace PA programmes. Senior management teams will require in depth information regarding the PA prescription involved in a programme, before they can allow their employees to participate. Lastly, the average weekly PA time in the included studies was ~80 minutes. "Lack of time" has been reported as a barrier to workplace PA participation [71], therefore more time efficient PA prescriptions may be well received by organisations and employees alike.

5. Conclusion

Our meta-analysis shows that PA that is delivered in the workplace can result in beneficial improvements in CRF of 2.7 $mL \cdot kg^{-1} \cdot min^{-1}$, when compared to controls. The prediction interval, which describes a plausible range of effects in a hypothetical future workplace PA intervention in similar settings ranged from a reduction in VO_{2peak} of -1.1 to an improvement of 6.5 $mL \cdot kg^{-1} \cdot min^{-1}$. The effect was moderated by risk of bias score and intervention length such that studies with higher risk of bias and longer interventions demonstrated larger intervention effects. However the impact of intervention length should be interpreted with caution due to the presence of an outlier. At this time the effects of age, sex and baseline fitness of participants are not definitive, and further work in the area is needed to fully clarify the impact of these characteristics. Presently therefore, we conclude that no specific group of employees (e.g. older or less fit) stand to gain more, or indeed less, from workplace PA interventions aimed at improving CRF. Future research should explore the effect of workplace PA on outcomes beyond simply CRF, whilst also ensuring that precise and in-depth reporting of the PA prescription and implementation is provided.

6. Funding and conflicts of interest

No sources of funding were used to assist in the preparation of this article. Naomi Burn, Matthew Weston, Neil Maguire, Greg Atkinson and Kathryn Weston declare that they have no conflicts of interest relevant to the content of this review.

7. Data Availability Statement

The dataset is available as a Microsoft excel file in Supplementary File 2.

7. References

1. Lee DC, Artero EG, Sui X, Blair SN. Mortality trends in the general population: the importance of cardiorespiratory fitness. J Psychopharmacol. 2010;24(4 Suppl):27-35.
2. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA. 2009;301(19):2024-35.
3. Lee DC, Sui X, Ortega FB, Kim YS, Church TS, Winett RA, et al. Comparisons of leisure-time physical activity and cardiorespiratory fitness as predictors of all-cause mortality in men and women. Br J Sports Med. 2011;45(6):504-10.
4. Laukkanen JA, Zaccardi F, Khan H, Kurl S, Jae SY, Rauramaa R. Long-term Change in Cardiorespiratory Fitness and All-Cause Mortality: A Population-Based Follow-up Study. Mayo Clin Proc. 2016;91(9):1183-8.
5. Lamoureux NR, Fitzgerald JS, Norton KI, Sabato T, Tremblay MS, Tomkinson GR. Temporal Trends in the Cardiorespiratory Fitness of 2,525,827 Adults Between 1967 and 2016: A Systematic Review. Sports Med. 2018.
6. Lin X, Zhang X, Guo J, Roberts CK, McKenzie S, Wu WC, et al. Effects of Exercise Training on Cardiorespiratory Fitness and Biomarkers of Cardiometabolic Health: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. J Am Heart Assoc. 2015;4(7).
7. World Health Organization. Global recommendations on physical activity for health. Geneva: World Health Organization; 2010.
8. Department of Health. Start active stay active: report on physical

activity in the UK. London: Department of Health; 2011.
9. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008;40(1):181-8.
10. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. Health Rep. 2011;22(1):7-14.
11. Marsaux CF, Celis-Morales C, Hoonhout J, Claassen A, Goris A, Forster H, et al. Objectively Measured Physical Activity in European Adults: Cross-Sectional Findings from the Food4Me Study. PLoS One. 2016;11(3):e0150902.
12. World Health Organization. The workplace as a setting for interventions to

improve diet and promote physical activity. Geneva: World Health Organization; 2007.
13. National Institute for Health and Care Excellence. Physical Activity in the Workplace. London: National Centre for Health and Care Excellence; 2008.
14. Justine M, Azizan A, Hassan V, Salleh Z, Manaf H. Barriers to participation in physical activity and exercise among middle-aged and elderly individuals. Singapore Med J. 2013;54(10):581-6.
15. Reichert FF, Barros AJ, Domingues MR, Hallal PC. The role of perceived personal barriers to engagement in leisure-time physical activity. Am J Public Health. 2007;97(3):515-9.
16. Trost SG, Owen N, Bauman AE, Sallis JF, Brown W. Correlates of adults' participation in physical activity: review and update. Med Sci Sports Exerc. 2002;34(12):1996-2001.
17. Goetzel RZ, Henke RM, Tabrizi M, Pelletier KR, Loeppke R, Ballard DW, et al. Do workplace health promotion (wellness) programs work? J Occup Environ Med. 2014;56(9):927-34.
18. Pescud M, Teal R, Shilton T, Slevin T, Ledger M, Waterworth P, et al. Employers' views on the promotion of workplace health and wellbeing: a qualitative study. BMC Public Health. 2015;15:642.
19. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. Int J Behav Nutr Phys Act. 2009;6:26.
20. Pereira MJ, Coombes BK, Comans TA, Johnston V. The impact of onsite workplace health-enhancing physical activity interventions on worker productivity: a systematic review. Occup Environ Med. 2015;72(6):401-12.
21. Hipp JA, Dodson EA, Lee JA, Marx CM, Yang L, Tabak RG, et al. Mixed methods analysis of eighteen worksite policies, programs, and environments for physical activity. Int J Behav Nutr Phys Act. 2017;14(1):79.
22. Knox ECL, Musson H, Adams EJ. Workplace policies and practices promoting physical activity across England: What is commonly used and what works? Int J Workplace Health Manag. 2017;10(5):391-403.
23. Pronk NP, Martinson B, Kessler RC, Beck AL, Simon GE, Wang P. The association between work performance and physical activity, cardiorespiratory fitness, and obesity. J Occup Environ Med. 2004;46(1):19-25.

24. Bernaards CM, Proper KI, Hildebrandt VH. Physical activity, cardiorespiratory fitness, and body mass index in relationship to work productivity and sickness absence in computer workers with preexisting neck and upper limb symptoms. J Occup Environ Med. 2007;49(6):633-40.
25. Christensen JR, Kongstad MB, Sjøgaard G, Søgaard K. Sickness Presenteeism Among Health Care Workers and the Effect of BMI, Cardiorespiratory Fitness, and Muscle Strength. J Occup Environ Med. 2015;57(12):e146-52.
26. Sloan RA, Sawada SS, Martin CK, Church T, Blair SN. Associations between cardiorespiratory fitness and health-related quality of life. Health Qual Life Outcomes. 2009;7:47.
27. Sui X, Laditka JN, Church TS, Hardin JW, Chase N, Davis K, et al. Prospective study of cardiorespiratory fitness and depressive symptoms in women and men. J Psychiatr Res. 2009;43(5):546-52.
28. Gerber M, Lindwall M, Lindegård A, Börjesson M, Jonsdottir IH. Cardiorespiratory fitness protects against stress-related symptoms of burnout and depression. Patient Educ Couns. 2013;93(1):146-52.
29. Dishman RK, Oldenburg B, O'Neal H, Shephard RJ. Worksite physical activity interventions. American Journal of Preventive Medicine. 1998;15(4):344-61.
30. Proper KI, Koning M, van der Beek AJ, Hildebrandt VH, Bosscher RJ, van Mechelen W. The effectiveness of worksite physical activity programs on physical activity, physical fitness, and health. Clin J Sport Med. 2003;13(2):106-17.
31. Dudgill L, Brettle A, Hulme C. Workplace physical activity interventions: a systematic review. Int J Workplace Health Manag. 2008;1:20-40.
32. Abraham C, Graham-Rowe E. Are worksite interventions effective in increasing physical activity? A systematic review and meta-analysis. Health Psychology Review. 2009;3(1):108-44.
33. To QG, Chen TT, Magnussen CG, To KG. Workplace physical activity interventions: a systematic review. Am J Health Promot. 2013;27(6):e113-23.

34. Reed JL, Prince SA, Elliott CG, Mullen KA, Tulloch HE, Hiremath S, et al. Impact of Workplace Physical Activity Interventions on Physical Activity and Cardiometabolic Health Among Working-Age Women: A Systematic Review and Meta-Analysis. Circ Cardiovasc Qual Outcomes. 2017;10(2).
35. Conn VS, Hafdahl AR, Cooper PS, Brown LM, Lusk SL. Meta-analysis of workplace physical activity interventions. Am J Prev Med. 2009;37(4):330-9.
36. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. 1. Aufl. ed. Oxford: Wiley-Blackwell; 2008.
37. Courneya KS. Efficacy, effectiveness, and behavior change trials in exercise research. Int J Behav Nutr Phys Act. 2010;7:81.
38. Garner P, Hopewell S, Chandler J, MacLehose H, Akl EA, Beyene J et al. When and how to update systematic reviews: consensus and checklist. BMJ. 2016; 354 :i3507.

39. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009;62(10):e1-34.
40. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc. 2011;43(7):1334-59.
41. Heyward VH. Advanced fitness assessment and exercise prescription. 6th ed. ed. Leeds: Human Kinetics; 2010.
42. Hariton E & Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. BJOG. 2018; 125(13): 1716-1716.
43. Jackson D, Turner R. Power analysis for random-effects meta-analysis. Res Synth Methods. 2017; 8:290-302.
44. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther. 2003;83(8):713-21.
45. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.
46. Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to meta-analysis. Chichester, UK: Wiley, 2009.
47. Melsen WG, Bootsma MCJ, Rovers MM, et al. The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses. Clin Microbiol Infect 2014;20:123–9.
48. Rücker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I2 in assessing heterogeneity may mislead. BMC Med Res Methodol 2008;8:79.

49. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc. 2009;41(1):3-13.
50. IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247.
51. Brown DK, Barton JL, Pretty J, Gladwell VF. Walks4Work: assessing the role of the natural environment in a workplace physical activity intervention. Scand J Work Environ Health. 2014;40(4):390-9.
52. Jay K, Frisch D, Hansen K, Zebis MK, Andersen CH, Mortensen OS, et al. Kettlebell training for musculoskeletal and cardiovascular health: a randomized controlled trial. Scand J Work Environ Health. 2011;37(3):196-203.
53. Mulla DM, Wiebenga EG, Chopp-Hurley JN, Kaip L, Jarvis RS, Stephens A, et al. The Effects of Lower Extremity Strengthening Delivered in the Workplace on Physical Function and Work-Related Outcomes Among Desk-Based Workers: A Randomized Controlled Trial. J Occup Environ Med. 2018;60(11):1005-14.
54. Andersen LL, Sundstrup E, Boysen M, Jakobsen MD, Mortensen OS, Persson R. Cardiovascular health effects of internet-based encouragements to do daily workplace stair-walks: randomized controlled trial. J Med Internet Res. 2013;15(6):e127.
55. Korshøj M, Lidegaard M, Skotte JH, Krustrup P, Krause N, Søgaard K, et al. Does aerobic exercise improve or impair cardiorespiratory fitness and health among cleaners? A cluster randomized controlled trial. Scandinavian Journal of Work, Environment & Health. 2015;41(2):140-52.
56. Kennedy RA, Boreham CAG, Murphy MH, Young IS, Mutrie N. Evaluating the effects of a low volume stairclimbing programme on measures of health-related fitness in sedentary office workers. Journal Of Sports Science & Medicine. 2007;6(4):448-54.
57. Sertel M, Üçsular FD, Uğurlu Ü. The effects of worksite exercises on physical capabilities of workers in an industry of a developing country: A randomized controlled study. Isokinetics & Exercise Science. 2016;24(3):247-55.
58. Gram B, Holtermann A, Søgaard K, Sjøgaard G. Effect of individualized worksite exercise training on aerobic capacity and muscle strength among construction workers--a randomized controlled intervention study. Scandinavian Journal Of Work, Environment & Health. 2012;38(5):467-75.
59. Grønningsæter H, Hytten K, Skauli G, Christensen CC, Ursin H. Improved health and coping by physical exercise or cognitive behavioral stress management training in a work environment. Psychology & Health. 1992;7(2):147-63.
60. Oden G, Crouse SF, Reynolds C. Worker productivity, job satisfaction, and work related stress: the infuence of an employee fitness program. Fitness in Business. 1989;3(6):198-203.
61. von Thiele Schwarz U, Lindfors P. Improved fitness after a workbased physical exercise program. International Journal Of Workplace Health Management. 2015;8(1):61-74.
62. de Zeeuw ELEJ, Tak ECPM, Dusseldorp E, Hendriksen IJM. Workplace exercise intervention to prevent depression: A pilot randomized controlled trial. Mental Health and Physical Activity. 2010;3(2):72-7.
63. Ekblom-Bak E, Ekblom Ö, Andersson G, Wallin P, Söderling J, Hemmingsson E, et al. Decline in cardiorespiratory fitness in the Swedish working force between 1995 and 2017. Scand J Med Sci Sports. 2018.
64. Verhagen AP, de Vet HC, de Bie RA, Boers M, van den Brandt PA. The art of quality assessment of RCTs included in systematic reviews. J Clin Epidemiol. 2001;54(7):651-4.
65. Dumas JE, Lynch AM, Laughlin JE, Phillips Smith E, Prinz RJ. Promoting intervention fidelity. Conceptual issues, methods, and preliminary results from the EARLY ALLIANCE prevention trial. Am J Prev Med. 2001;20(1 Suppl):38-47.
66. Taylor KL, Weston M, Batterham AM. Evaluating intervention fidelity: an example from a high-intensity interval training study. PLoS One. 2015;10(4):e0125166.
67. Steins Bisschop CN, Courneya KS, Velthuis MJ, Monninkhof EM, Jones LW, et al. Control Group Design, Contamination and Drop-Out in Exercise Oncology Trials: A Systematic Review. PLOS ONE. 2015; 10(3): e0120996.
68. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions
: the new Medical Research Council guidance. Int J Nurs Stud. 2013;50(5):587-92.
69. Davis K, Gorst SL, Harman N, Smith V, Gargon E, et al. Choosing important health outcomes for comparative effectiveness research: An updated systematic review and involvement of low and middle income countries. PLOS ONE. 2018; 13(2): e0190695.
70. Bauman A, & Nutbeam D, Evaluation in a nutshell: a practical guide to the evaluation of health promotion programs. Sydney AU; McGaw-Hill. 2013.

71. Hunter JR, Gordon BA, Bird SR, Benson AC. Perceived barriers and facilitators to workplace exercise participation. International Journal of Workplace Health Management. 2018;11(5):349-63.

72. Armijo-Olivo S, da Costa BR, Cummings GG, Ha C, Fuentes J, Saltaji H, Egger M. PEDro or Cochrane to assess the quality of physiotherapy trials? A meta-epidemiological study. In: Filtering the information overload for better decisions. Abstracts of the 23rd Cochrane Colloquium; 2015 3-7 Oct; Vienna, Austria. John Wiley & Sons; 2015.
73. Franklin J, Atkinson G, Atkinson J, Batterham A. Peak Oxygen Uptake in Chronic Fatigue Syndrome/Myalgic Encephalomyelitis: A Meta-Analysis. IJSM. 2019; 40(2):77-87.
74. Hurst C, Weston KL, Weston M. The effect of 12 weeks of combined upper- and lower-body high-intensity interval training on muscular and cardiorespiratory fitness in older adults. Aging Clin Exp Res. 2018.
75. Straight CR, Lindheimer JB, Brady AO, Dishman RK, Evans EM. Effects of Resistance Training on Lower-Extremity Muscle Power in Middle-Aged and Older Adults: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Sports Med. 2016;46(3):353-64.
76. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better Reporting of Interventions: Template for Intervention Description and Replication (TIDieR) Checklist and Guide. BMJ. 2014; 348:g1687.
77. Weston M, Taylor KL, Batterham AM, Hopkins WG. Effects of low-volume high-intensity interval training (HIT) on fitness in adults: a meta-analysis of controlled and non-controlled trials. Sports Med. 2014;44(7):1005-17

78. Slade SC, Dionne CE, Underwood M, et al. Consensus on Exercise Reporting Template (CERT): Explanation and Elaboration Statement. Br J Sports Med. 2016;50:1428-1437.

8. Tables and Figures

Table 1: Search Strategy

1. Workplace terms	TI ABS KEY (workplace OR worksite OR “work place” OR “work site” OR “working place” OR “work setting” OR “working environment” OR worker OR workforce OR employe* OR industr* OR corporate)
2. Physical activity terms	TI ABS KEY (exercise OR “physical activity” OR fitness OR sport OR “aerobic training” OR “stair climbing” OR “stair use” OR “resistance training” OR “strength training” OR walk*)
3. Intervention terms	TI ABS KEY (intervention OR program* OR evaluation OR trial OR randomi* OR promotion OR effect*)
1 AND 2 AND 3	

Table 2: Study and subject characteristics for VO_{2peak} estimates included in the meta-analysis

Author, date	Study design	Location	Population, sample size (n), % female, age (year) (mean unless otherwise stated)	Group	Duration (weeks)	Frequency (per week)	Session length (mins)	Physical activity mode	Physical activity intensity	Attendance	Dose Quantification	VO_{2peak} measurement technique	Baseline VO_{2peak} (mL·kg⁻·min⁻¹)	% Δ VO_{2peak} (mL·kg⁻·min⁻¹)
Brown et al. [51]	RCT	United Kingdom	Office workers n=32 (15.6% female, 46y)	I1	8	2	20	Aerobic (walking-nature walk)	2km walk in 20 mins	Twice weekly session attendance achieved by 42% of participants	Average increase in steps/ day: 745 steps/ (target step increase: 600 steps/ day)	Chester step test*	39.0	+0.8
			Office workers n=33 (27.3% female, 39y)	I2	8	2	20	Aerobic (walking-built environment walk)	2km walk in 20 mins	Twice weekly session attendance achieved by 43% of participants	Average increase in steps/ day: 374 steps (target step increase: 600 steps/ day)	Chester step test*	39.8	0.0
			Office workers n=29 (20.6% female, 40y)	C				Wait list control				Chester step test*	39.6	+5.3
Jay et al. [52]	RCT	Denmark	Pharmaceutical company workers	I	8	3	20	Resistance (kettle bells)	Four increasingly difficult	70% average session	Average two-handed kettle bell	Submaximal cycle ergometer	37.0	+7.8

			n=20 (85% female, 44y)						kettlebell exercises. Progression when each exercise could be completed at the highest level of intensity.	attendance	swing weight and number of sets increase from 8.3kg for 23.2 sets in week 1-2 to 12.4kg for 22.1 sets in week 7 and 8.	Astrand protocol*		
			Pharmaceutical company workers n=20 (85% female, 43y)	C				Inactive control				Submaximal cycle ergometer Astrand protocol*	39.0	+12.3
Mulla et al. [53]	RCT	Canada	Ford motor company workers n=21 (57% female, 44y)	I	12	3	45	Resistance (lower body strengthening)	RPE= 5-7 on 10 point RPE scale	Mean session attendance: 2.3 per week	Not reported	Ebbeling Single stage treadmill walking test*	34.9	+2.9
			Ford motor company workers n=21 (68% female, 43y)	C				Wait list control				Ebbeling Single stage treadmill walking test*	34.6	+0.3
Andersen et al. [54]	RCT	Denmark	Office workers n=106 (79.2% female, 42y)	I	10	5	10	Aerobic (Stair climbing)	90% HRR (measured retrospectively)	Mean session attendance: 3.3/ week.	Mean heart rate recorded during one session: 90% HRR	Maximal cycle ergometer	36.0	+6.4

			Office workers n=54 (75.9% female, 43y)	C				Inactive control				Maximal cycle ergometer	38.0	+2.3
Korshøj et al. [55]	Cluster RCT	Denmark	Cleaners n= 57 (75.4 % female, 45y)	I	16	2	30	Aerobic exercise (more specific details not reported)	≥60% VO_{2peak}	Mean session attendance: 51%	Mean heart rate recorded every fourth week: 67% HRR	Step test*	24.8	+7.1
			Cleaners n=59 (76.3% female, 50y)	C				Health lectures				Step test*	25.0	-1.0
Kennedy et al. [56]	RCT	United Kingdom	Office workers n= 35 (43% female, 44y)	I	8	5	1- 3 climbs / sessions (4 mins average)	Aerobic (stair climbing)	8 flights of stairs at 75 steps/ min	Mean session attendance: 88%	Not reported	Submaximal cycle ergometer YMCA protocol*	27.8	+9.4
			Office workers n=17 (65% female, 39y)	C				Inactive control				Submaximal cycle ergometer YMCA protocol*	27.1	+1.9
Sertel et al. [57]	RCT	Turkey	Poultry processing workers n=31 (100% female, 31y)	I1	8	3	30	Resistance (theraband exercises)	50-85% MVC	Session attendance not reported	Not reported	Queens college step test*	38.5	+6.8
			Poultry processing workers n=30 (100%	I2	8	3	30	Aerobic (walking/ running)	40-80% VO_{2peak}	Session attendance not reported	Not reported	Queens college step test*	38.3	+9.4

			female, 33y)											
			Control n=30 (100% female, 35y)	C				Inactive control				Queens college step test*	35.5	0.0
Gram et al. [58]	RCT	Denmark	Construction workers n=35 (100% male, 44y)	I	12	3	20	Multicomponent cycling/ rowing and resistance training	≥70% VO_{2peak} or 60% 1RM	Mean session attendance: 68%	Not reported	Submaximal cycle ergometer Astrand protocol*	27.1	+14.4
			Construction workers n= 32 (100% male, 43y)	C				Health lectures				Submaximal cycle ergometer Astrand protocol*	26.5	+1.1
Grønningsæter et al. [59]	RCT	Norway	Insurance company workers n=30 (43% female, 25-67y [range as mean age not provided])	I	10	3	55	Aerobic (rhythmic aerobics)	70-80% HR_{Max}	Mean session attendance: 76% (men), 80% (women)	Not reported	Submaximal cycle ergometer Astrand protocol*	39.8	+7.0
			Insurance company workers n=31 (48% female, 25-67y)	C				Waitlist control				Submaximal cycle ergometer Astrand protocol*	37.8	-3.3
Oden et al. [60]	RCT	USA	Westinghouse corporation workers n=23 (80%	I	24	3	30	Aerobic (walking, jogging, cycling)	60-80% HRR	Not reported	Not reported	Maximal treadmill Bruce protocol	30.0	+17.9

			female, 29y,)										
			Westinghouse corporation workers n=22 (80% female, 29y)	C				Inactive control				Maximal treadmill Bruce protocol	29.9	+0.6
von Thiele Schwarz & Lindfors [61]	CT	Sweden	Care workers n=15 (100% female, 42y)	I	52	2	60	Aerobic (aerobics, Nordic walking, cross training)	"Middle to high intensity"	Not reported	Not reported	Submaximal cycle ergometer Astrand protocol*	35.8	+11.5
			Care workers n=13 (100% female, 44y)	C				Inactive control				Submaximal cycle ergometer Astrand protocol*	29.5	-9.5
de Zeeuw et al. [62]	RCT	Netherlands	Insurance company workers n= 15 (40% female, 41y)	I	10	2	40-60	Aerobic (cycling, jogging, stair stepping)	80% HR_{Max}	Mean session attendance: 85.6%	Not reported	Submaximal cycle ergometer Astrand protocol*	29.7	+24.9
			Insurance company workers n=15 (53% female, 41y)	C				Inactive control				Submaximal cycle ergometer Astrand protocol*	32.3	-1.9

Studies are sorted from the smallest to largest effects on VO_{2peak}.

Key: CT: controlled trial, RCT: randomised controlled trial, y: years, I: intervention group, C: control group, HR_{max}: maximum heart rate, HRR: heart rate reserve, VO_{2peak}: peak aerobic capacity, MVC: maximum voluntary contraction, 1RM: one repetition maximum, RPE: Rating of Perceived Exertion, *submaximal prediction of VO_{2peak}.

Table 3: PEDro Risk of Bias Assessment for studies included in the meta-analysis

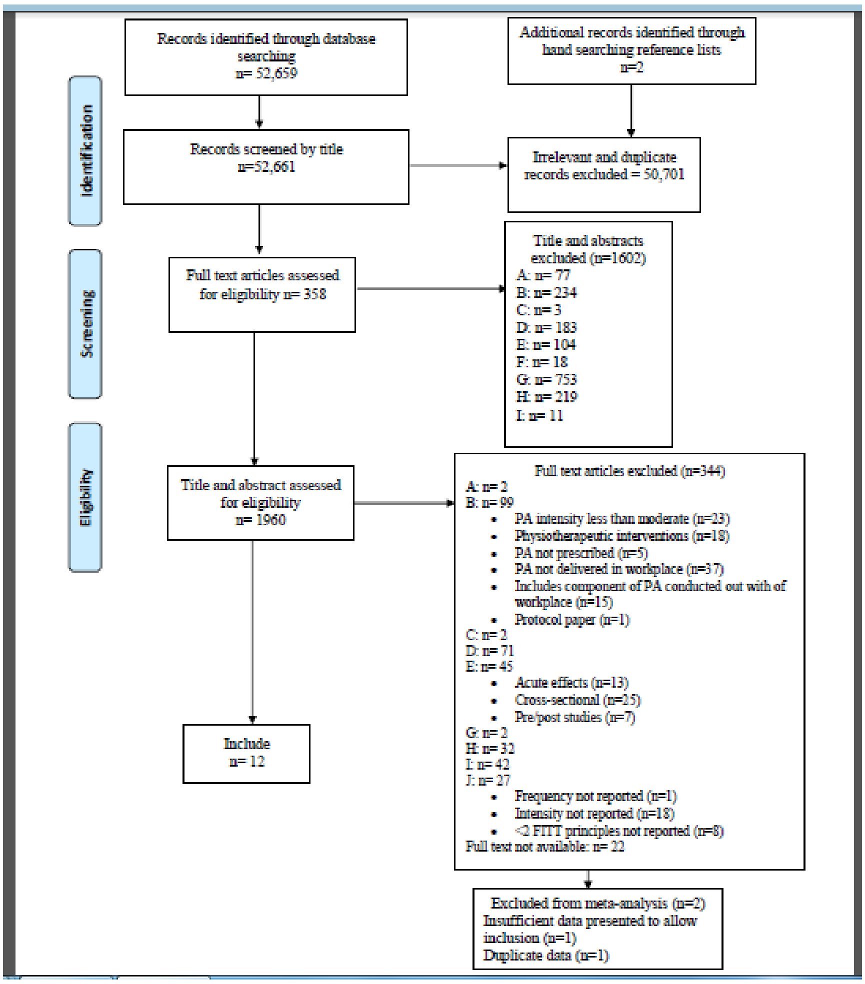
Reference	1) Specified eligibility criteria	2) Demonstrated random allocation	3) Concealed allocation	4) Similar groups at baseline	5) Subjects blinded	6) Facilitators blinded	7) Assessors blinded	8) >85% followed up	9)Intervention received as intended/ ITT analysis	10) Results of between-group statistical comparisons reported for at least one outcome	11)point measures and measures of variability for at least one outcome	Total/10
Brown et al.[51]	Y	+	+	+	-	-	-	-	-	+	+	5
Jay et al. [52]	Y	+	+	+	-	-	+	+	+	+	+	8
Mulla et al. [53]	Y	+	+	+	-	-	+	+	+	+	+	8
Andersen et al.[54]	Y	+	+	+	-	-	+	+	+	+	+	8
Korshøj et al.[55]	Y	+	+	+	-	-	+	-	+	+	+	7
Kennedy et al. [56]	Y	+	-	+	-	-	-	+	-	-	+	4
Sertel et al. [57]	Y	+	-	+	-	-	-	-	-	+	+	4
Gram et al. [58]	Y	+	+	+	-	-	-	+	+	+	+	7
Grønningsæter et al.[59]	Y	+	-	-	-	-	-	+	-	+	+	4
Oden et al. [60]	N	+	-	+	-	-	-	-	-	+	+	4
von Thiele Schwarz & Lindfors [61]	N	-	-	-	-	-	-	+	-	+	+	3
de Zeeuw et al. [62]	Y	+	+	+	-	-	-	+	-	+	+	6
	Total/ 12	11	7	10	0	0	4	8	5	11	12	

Key: Y= yes, N=no. += demonstrated evidence, - = no demonstrated evidence. Scores of ≥6 considered of low risk of bias, scores of <6 considered high risk of bias.

Table 4: Meta-regression for the modifying effects for study and participant characteristics

Model	**Coefficient (90% CI)**	τ^2	R^2	I^2	**p-value**
Intervention length (weeks)	0.17 (1.3 to 5.1)	1.79	26%	73%	0.01
Pedro risk of bias score	-0.66 (-4.0 to -1.0)	1.74	28%	78%	0.01
Baseline VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)	-0.16 (-3.6 to 0.4)	2.49	0%	78%	0.18
Age (years)	-0.15 (-3.2 to 0.3)	2.42	14%	74%	0.16
Sex (% female)	0.01 (-1.6 to 2.4)	2.77	0%	80%	0.73

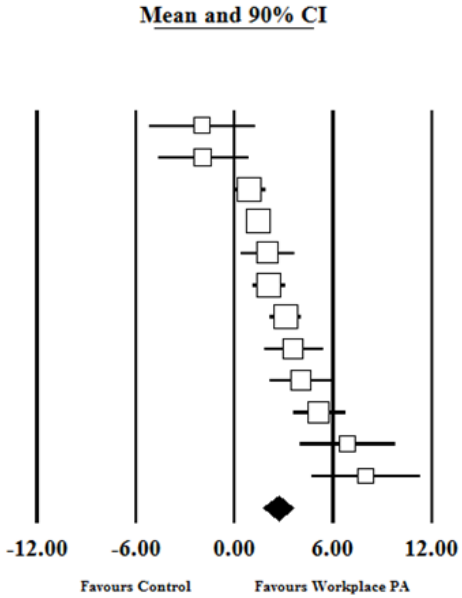
Figure 1: PRISMA Flow Diagram of study selection



Key: Inclusion/ exclusion criteria:
Inclusion: A. population: working adults (aged 16+ years), B. intervention: exercise of ≥moderate intensity is prescribed and delivered in the workplace, C. comparator: no intervention control groups, or non-active comparators, D. outcome: Cardiorespiratory fitness (actual or predicted VO_{2peak}), E. study design: RCTs or controlled trials. Exclusion: F. Population: specific patient groups, disabled populations, retirees or unemployed, G. PA promotion/ environmental change/ sedentary behavior interventions when exercise was not specifically prescribed and delivered in the workplace, H. multi-component interventions (e.g. interventions with concurrent exercise and dietary components). I: Not in English J: All FITT (frequency, intensity, time and type of exercise) principles adequately reported.

Figure 2: Forest Plot for the effect of workplace PA interventions on VO_{2peak}

Study name	Mean	Standard error	Lower limit	Upper limit	p-Value
Brown et al. [51]	-2.0	2.0	-5.2	1.3	0.3
Jay et al. [52]	-1.9	1.7	-4.7	0.9	0.3
Mulla et al. [53]	0.9	0.6	-0.1	1.9	0.1
Andersen et al. [54]	1.5	0.4	0.8	2.1	0.0
Korshøj et al. [55]	2.0	1.0	0.4	3.6	0.0
Kennedy et al. [56]	2.1	0.6	1.1	3.1	0.0
Sertel et al. [57]	3.1	0.6	2.2	4.1	0.0
Gram et al. [58]	3.6	1.1	1.8	5.4	0.0
Grønningsæter et al. [59]	4.0	1.2	2.1	6.0	0.0
Oden et al. [60]	5.2	1.0	3.6	6.8	0.0
von Thiele Schwarz & Lindfors [61]	6.9	1.8	4.0	9.8	0.0
de Zeeuw et al. [62]	8.0	2.0	4.7	11.3	0.0
	2.7	0.6	1.8	3.6	0.0



Legend: The size of the boxes represents the weight given to each individual study in the meta-analysis. Larger boxes indicate a higher weighting.

Figure 3: Funnel Plot

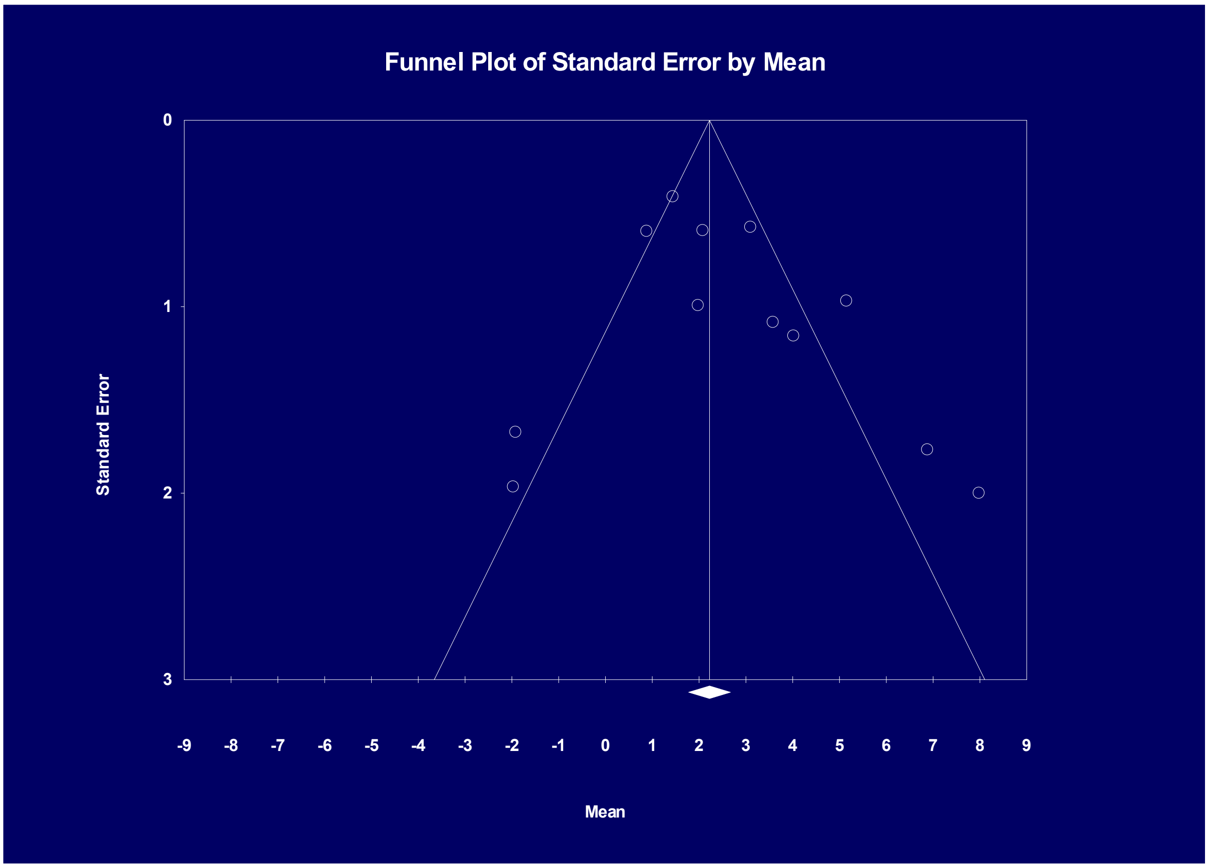
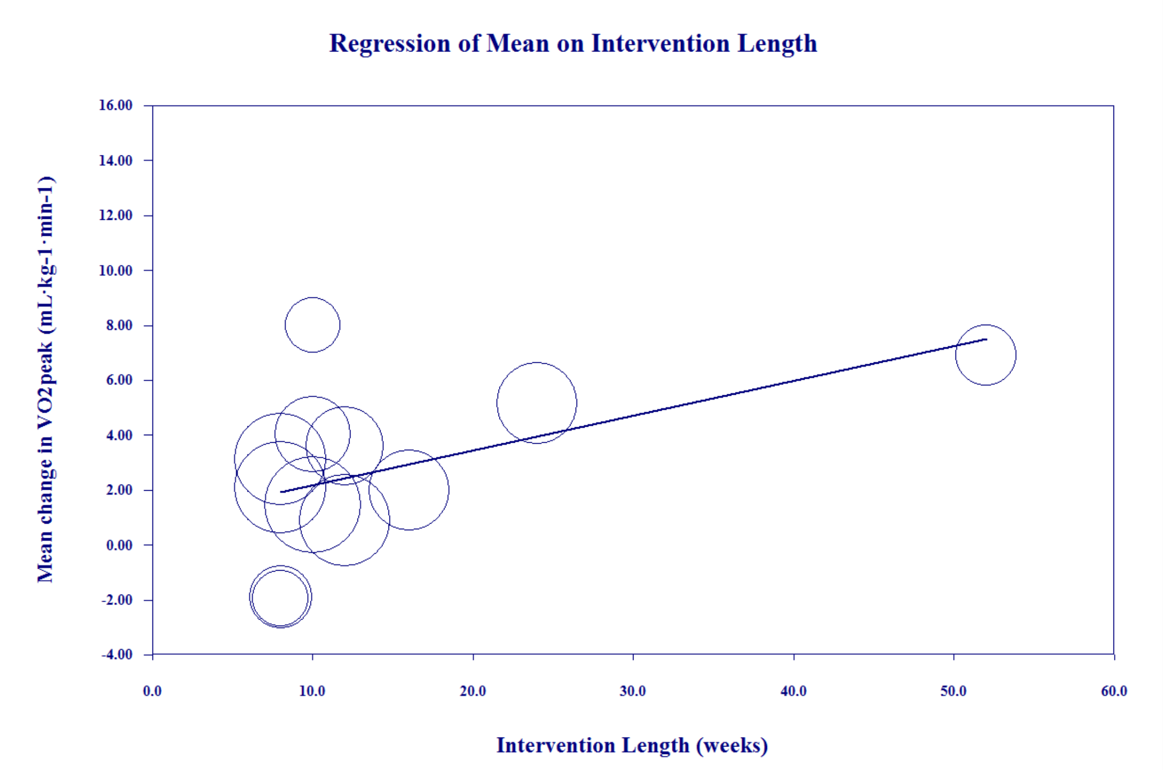
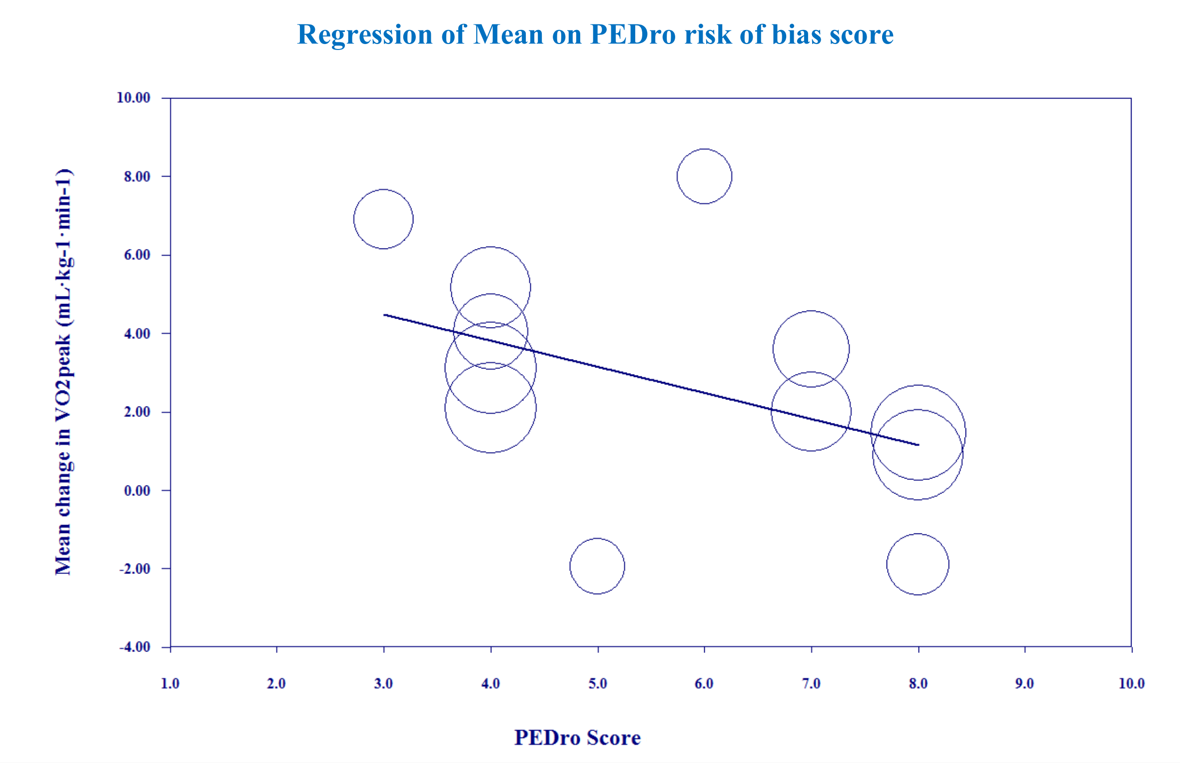


Figure 4: Scatter plot of meta-regression of intervention length.



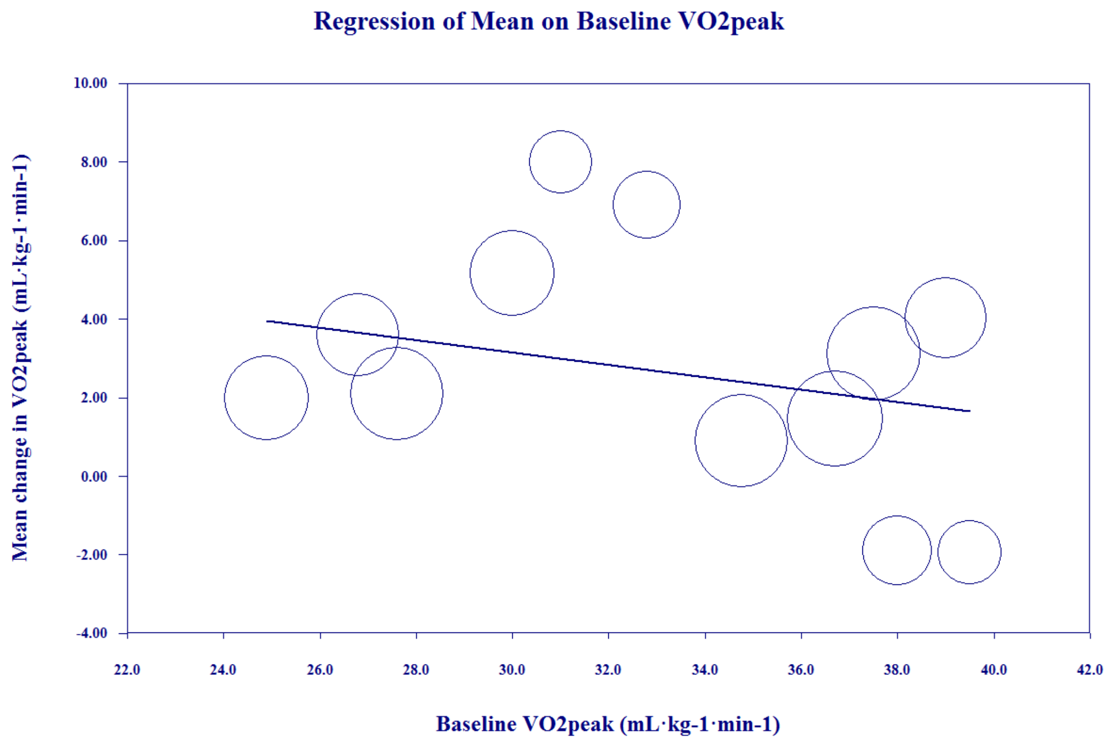
Legend: The size of the circles represents the weight given to each individual study in the meta-analysis. Larger circles indicate a higher weighting.

Figure 5: Scatter plot of meta-regression of PEDro risk of bias assessment score..



Legend: The size of the circles represents the weight given to each individual study in the meta-analysis. Larger circles indicate a higher weighting.

Figure 6: Scatter plot of meta-regression of baseline VO_{2peak}.

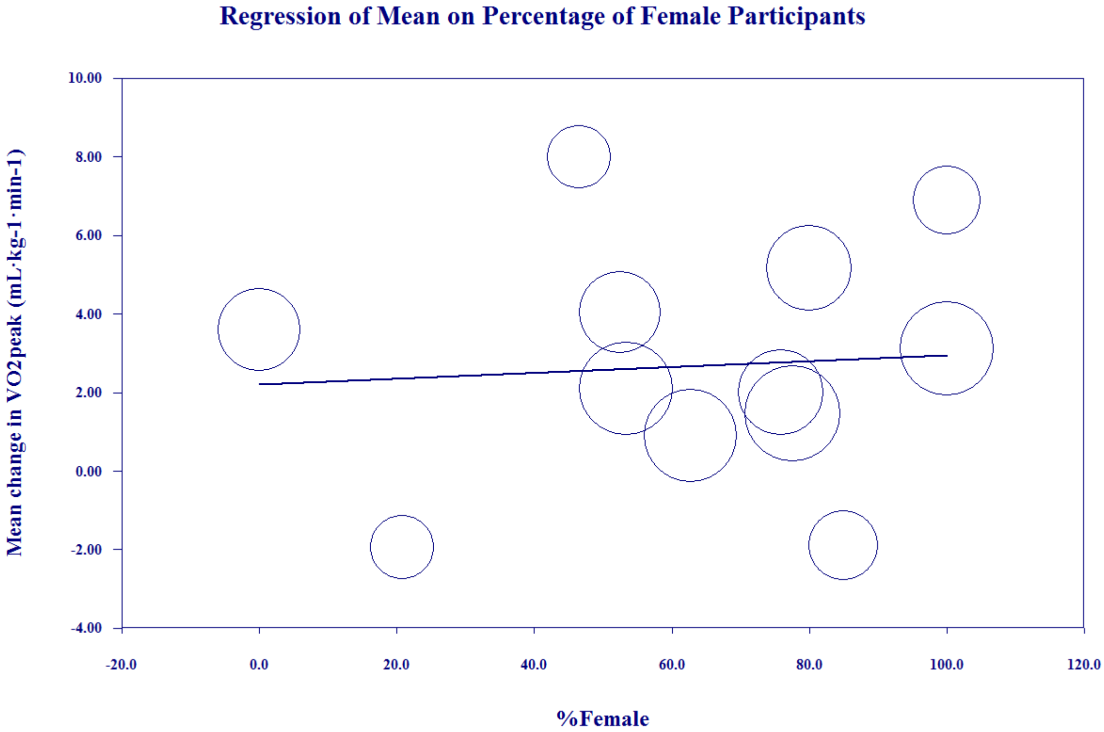


Legend: The size of the circles represents the weight given to each individual study in the meta-analysis. Larger circles indicate a higher weighting.

Figure 7: Scatter plot of meta-regression of mean age.

Legend: The size of the circles represents the weight given to each individual study in the meta-analysis. Larger circles indicate a higher weighting.

Figure 8: Scatter plot of meta-regression of percentage of female participants.



Legend: The size of the circles represents the weight given to each individual study in the meta-analysis. Larger circles indicate a higher weighting.