QUANTIFYING THE ASSOCIATION BETWEEN PHYSICAL ACTIVITY AND CARDIOVASCULAR DISEASE & DIABETES: A SYSTEMATIC REVIEW AND META-ANALYSIS

Abdul Ahad Wahid[1], BSc, MBBS; Nishma Manek[2], BSc, MBBS; Melanie Nichols[3,4], BSc, PhD; Paul Kelly[3,5], BSc, MSc, PhD; Charlie Foster[3], BSc, PhD; Premila Webster[6], MBBS; Asha Kaur[3], BSc; Claire Friedemann Smith[3,7] , BSc, MSc, PhD; ElizabethWilkins[3], BA; Mike Rayner[3], BSc, PhD; Nia Roberts[8]; Peter Scarborough[3*], BSc, PhD

[1] John Radcliffe Hospital, Oxford & University of Oxford Clinical Academic Graduate School (OUCAGS)

[2] Watford General Hospital, Imperial College Healthcare NHS Trust, London

[3] The British Heart Foundation Centre on Population Approaches for Non-Communicable Disease Prevention, Nuffield Department of Population Health, University of Oxford

[4] WHO Collaborating Centre for Obesity Prevention, Deakin University, Geelong, Australia

[5]Sport, Physical Education and Health Sciences Department (SPEHS), University of Edinburgh

[6] Nuffield Department of Population Health, University of Oxford

[7] Health Behaviour Research Centre, Department of Epidemiology and Public Health, University College London

[8] Bodleian Health Care Libraries, Knowledge Centre, University of Oxford Old Road Campus

* Corresponding author: Peter Scarborough; email peter.scarborough@dph.ox.ac.uk; Tel: +44 (0)1865 289245, Fax: +44 (0) 1865 617789
The BHF Centre on Population Health, Old Road Campus, Roosevelt Drive, Headington, Oxford OX3 7LF, UK

Total word count: 7658
Journal subject code: [8] Epidemiology, [26] Exercise, [110] Health policy and outcome research and [121] Primary prevention

Abstract

Background: The relationships between physical activity (PA) and both cardiovascular disease (CVD) and Type II Diabetes (T2DM) have predominantly been estimated using categorical measures of PA, masking the shape of the dose response relationship.

In this systematic review and meta-analysis, for the very first time we are able to derive a single continuous PA metric to compare the association between PA and CVD/T2DM, both before and after adjustment for a measure of body weight.

Methods and Results: The search was applied to MEDLINE and EMBASE electronic databases for all studies published from January 1981 to March 2014. A total of thirty-six studies (3,439,874 participants and 179,393 events, during an average follow up period of 12.3 years) were included in the analysis (33 pertaining to CVD and 3 to T2DM).

An increase from being inactive to achieving recommended PA levels 150 minutes of moderate-intensity aerobic activity per week).was associated with lower risk of CVD by 23%, CVD incidence by 17% and T2DM incidence by 26% (RR 0.77 (0.71-0.84)), (RR 0.83 (0.77-0.89)) and (RR 0.74 (0.72-0.77)) respectively, after adjustment for body weight.

Conclusions: By using a single continuous metric for PA levels we were able to make a comparison of the effect of physical activity on CVD incidence and mortality including myocardial infarct (MI), stroke and heart failure, as well as type II diabetes. Effect sizes were generally similar for CVD and type II diabetes, and suggested that the greatest gain in health is associated with moving from inactivity to small amounts of PA.

Keywords: Physical activity; cardiovascular diseases; systematic review; meta-analysis

Review registration: The protocol was registered with the PROSPERO database: CRD42014009655

INTRODUCTION

Insufficient physical activity (PA) is a key risk factor for non-communicable diseases (NCDs) such as cardiovascular diseases (CVD), cancer and diabetes.[1] CVD is the number one cause of death globally, with 17.5 million deaths from CVD and 1.5 million deaths from diabetes in 2012, representing 31% and 2.7% of global deaths respectively[2, 3]

It has been over half a century since the pioneering studies of bus drivers and then longshoremen which first established the beneficial impact of physical activity (PA) upon cardiovascular disease (CVD) risk[2, 3]. More recently, in 2010, the UK Chief Medical Officer Sir Liam Donaldson declared that the benefits of regular PA on health, longevity, and wellbeing "easily surpass the effectiveness of any drugs or other medical treatment"[4].

Current international recommendations for PA are to achieve at least 150 minutes per week of moderate intensity aerobic PA, or 75 minutes per week of vigorous PA[5, 6], and a higher level of 300 minutes per week of moderate intensity PA has been recommended to reduce the risk of cancer [7]. These recommendations were typically constructed on systematic reviews and meta-analysis of epidemiological studies however these reviews were not without limitations. Many only assessed the benefits of a single domain or facet of PA, such as Lee et al (2000)[8] which examines only walking and recreational activity. Most had pooled results based on different ranges of categorical exposure measures (e.g. 'high' or 'low'), for example Fujita et al (2004)[9] which compares the sole domain of walking on mortality, or Woodcock (2011)[10] which focuses on 'non-vigorous' physical activity. Further, existing reviews such as Jeon et al (2006)[11] do not explore the risk reductions after adjusting for body weight and therefore do not allow for an assessment of the independent effect of PA on health outcomes, and at least part of the observed effect is likely to be mediated by maintenance of healthy weight status.

The aim of this study was to conduct a systematic review and meta-analysis to draw together the epidemiological studies that assesses the independent association between PA levels and both CVD and T2DM outcomes, using a single continuous metric and adjusting for body weight. In order to make comparable results across disease domains, exposure data for physical activity was converted to a common continuous metric of MET hours per week. Although subject to limitations from original studies, this metric allowed us to study results in order to reflect international PA guidelines. More importantly, it also allowed us to estimate the relative risk associated with a unit increase in PA at any activity level, and to explore the dose-response relationship between PA and CVD or T2DM outcomes.

METHODS

Eligibility criteria

Studies considered for inclusion were prospective cohort studies that measured PA levels where at least two of the following domains were measured: leisure, household, active travel and occupational activity.

Estimates of the relative risk for incidence of or mortality from cardiovascular diseases or type two diabetes in participants free of disease at baseline had to be reported. Studies that provided estimates for total cardiovascular disease were included under the main CVD incidence/mortality results. Otherwise, any studies that reported individual CVD outcomes such as heart failure or myocardial infarct were included in separate meta-analyses.

In addition, the relative risk had to be adjusted for a measure of body weight (e.g. continuous measurements of body weight, BMI, waist circumference etc. or binary measures of overweight, obesity etc.). For consistency, all such measures are referred to as ‘body weight’. If multiple studies measuring the same outcome were published from the same cohort, preference was given to that

with the most years of follow up in the analysis. Studies were excluded if the PA measure was one of fitness, as opposed to a measure of time or volume of PA. Only studies published in English were included.

Search Strategy

Relevant studies were identified by searching electronic databases and supplemented by scanning the reference lists of included studies and relevant systematic reviews. The search was applied to MEDLINE and EMBASE electronic databases for all studies published from January 1981 to March 2014. Key terms, amongst others, included “physical activity”, “cardiovascular diseases”, “stroke”, “heart diseases”, and “mortality”. The complete search strategy is attached as supplemental online material.

Study selection and data extraction

Identified titles and abstracts were initially obtained through application of the search strategy, and these were divided equally amongst three authors to be screened for inclusion on the basis of title, with a fourth reviewer cross-checking a 10% sample of exclusion decisions. The abstracts of studies included after the initial screening were then independently assessed by two reviewers for inclusion in the review, with discrepancies referred to a third reviewer and resolved through discussion. A single reviewer reviewed full text articles, with a second reviewer cross-checking undecided cases and a random 10% sample of excluded papers. A single reviewer performed data extraction, with a 10% sample check conducted by a second reviewer. If any discrepancies arose between reviewers these were referred to a third reviewer for discussion.

Assessment of study quality

A quality criteria scale at the study level was developed using applicable elements from the Newcastle-Ottawa scale for cohort studies [12]. The developed scale was piloted on ten studies and

refined accordingly, and is displayed in Supplementary Table S1, along with the assessments of quality for each included study. This scale was previously used in the meta-analysis of walking and cycling by Kelly et al (2014)[13].

Conversion to standard PA units

Relative risks (RRs) with standard errors were obtained from each study for the various reported levels of PA in relation to the reported health outcomes. The PA exposure in each study was converted to MET hours per day (METhr/d) above that expended by the baseline group (referred to throughout as the inactive group) in order to quantify the association on a comparable scale.

One MET (Metabolic Equivalent) is defined as 1 kcal/kg and is approximately equivalent to the energy cost of sitting quietly [14]. This unit was chosen as it provided a continuous variable which could be converted to a single RR estimate for a selected change in volume of PA. Conversion to this scale was performed with reference to the Physical Activity Compendium [14] which was developed for use in epidemiologic studies to standardise the assignment of MET intensities in PA questionnaires.

If PA exposure estimates were not directly reported in the form of MET hours, the PA Compendium was used to convert the information provided about the amount of PA conducted to our standard unit to reflect the time and intensity of activity. Any reference to 'moderate' PA was assigned a value of 4.5 METs and vigorous activity a value of 6.5 METs, based on WHO recommendations [5]. Using the compendium, it was decided to assign any references to 'inactive behaviour' a value of 1.5 METs, and 'light activity' a value of 2.5 METs, and average body weights of 84kg and 71kg [15]were assigned to men and women respectively, where no other information was provided by the studies.

The rules for converting physical activity measures to a standard metric of 'Additional MET hours/day' are listed in Table S2 of the supplementary material.

Statistical analysis

The relationship between relative risk and PA was assessed using regression analysis for each study individually. In all cases, it was assumed that the relationship between relative risk and MET hours per week followed a 0.25 power transformation (i.e. RR = 1 + b*MET^0.25, where b is a regression coefficient). Additional analyses, namely both linear and loglinear associations, and power transformations of 0.375, 0.50 and 0.75 were also conducted to assess the sensitivity of the findings to the method of parameterising the dose-response relationship. The 0.25 power transformation was chosen for the primary outcome measures as it closely models the observed findings from previous studies that the relative risk for chronic disease falls quickly with a small addition of PA, but further increases of PA produce rapidly diminishing returns. This transformation has previously been used in a meta-analysis of the effect of PA on all-cause mortality[10] using the Alaike's Information Criterion, in order to select the transformation that best fits the data[16.]

Additionally, a categorical analysis was conducted, where the dose-response relationship was assessed non-parametrically, as previously conducted by Kelly et al (2014)[13]. For this analysis, data lines were grouped into three categories (0.1 – 12.0 MET hrs / week; 12.1 – 29.5 MET hrs / week; 29.6+ MET hrs / week) and meta-analyses within these three PA categories were conducted. The three categories refer to tertiles of data points collected from the studies.

For each meta-analysis where the dose-response relationship is described parametrically, the same dose-response relationship was assumed for each individual study, and the dose-response parameter was estimated separately for each study. The meta-analysis was then conducted on the dose-response parameter, with an average value of standard error from the PA groups included in each study. Using this method, results could be reported at any value of additional PA, and we have chosen to report results for an increase of 11.25MET hours per week – equivalent to moving from inactive behaviour to achieving international PA recommendations. The reported results using the

primary outcome measures can also be used to determine the modelled association between PA and the health outcome at any level of PA using a method described in Table 1.

Random effects meta-analyses were performed, due to the high degree of heterogeneity in the populations under investigation and follow-up time[17]. The heterogeneity between studies in the meta-analyses was assessed using the I^2 statistic. For each health outcome, two sets of meta-analyses were conducted using Stata version 12 (StataCorp LP, College Station, Texas, USA): 1) with RRs adjusted for body weight, including all identified studies (primary outcome); 2) with RRs NOT adjusted for body weight, including only studies where RRs were also available without adjustment for body weight. Here, 'RRs adjusted for body weight' were taken from the fully adjusted models reported in the included studies, and 'RRs NOT adjusted for body weight' were taken from the models adjusting for the most covariates without inclusion of a measure of body weight. Funnel plots were analysed to assess the possibility of small-study or publication bias. A sensitivity analysis was conducted by restricting results to studies that achieved at least six of the eight quality criteria. Meta-regression was performed to explore whether heterogeneity in the results could be explained by study-level variables including achievement of each of the eight quality criteria, the gender of participants, mean age of participants, presence of active travel, recreational, occupational and household domains in the PA measurement, mean follow-up years and geography.

RESULTS

The initial literature searches produced 16,628 titles. After a review of the titles and abstracts, 329 papers were reviewed in full, as detailed in the study flow chart (Figure 1). An additional eight studies were identified by scanning reference lists of the included studies. The majority of the excluded studies only measured one domain of PA or did not adjust for any measure of obesity. Thirty-six studies were included in the final review (listed in Table S3 of the Supplementary material), of which 33 contributed to CVD meta-analyses and 3 contributed to the type II diabetes meta-analyses.

Table 2 provides a summary of the results of the meta-analyses for estimates both with and without adjustment for body weight and also displays the total number of data points and incidents for each disease outcome. Five data points had to be excluded from meta-analyses as they pertained to conditions for which we only had one result, such as Wattanakit et al (2013)[18] which was the only study to report the PA effect on Venous Thromboembolism (VTE) incidence. These studies are listed in Table S4 in the supplementary material. In addition, three studies were removed in the review stage due to overlapping cohorts.

The various power transformations, namely both linear and loglinear associations, as well as power transformations of 0.25, 0.375, 0.50 and 0.75 are presented in Table 3 for the various outcomes. More specifically, Figure 2 graphically outlines relative risk results for the CVD mortality studies plotted against the amount of PA in our standardised metric. The 0.25 power transformation is chosen as the preferred transform in this paper as it demonstrates a better fit to the data; the r^2 value for this transformation was 0.75 compared to 0.29 for the loglinear transformation. In terms of the dose-response relationship, the incremental change in risk for three differing intensities of PA (low, medium and high are listed in Table 4. In general, we noted that the greatest rate of reduction occurs in the first category, i.e. as you move from low to medium amounts of physical activity rather than from medium to high.

Cardiovascular disease meta-analyses

The 33 studies included in the cardiovascular disease meta-analyses were conducted in Europe (n = 13), USA (n = 13) and the rest of the world (n = 7). They included a total of 1,683,693 participants, with 89,493 events occurring during an average follow up period of 12.8 years. The number of data points pooled for each health outcome was as follows: 5 for CVD incidence, 14 for CVD mortality, 9 for stroke incidence; 6 for CHD incidence, 2 for CHD mortality, 5 for heart failure incidence and 2 for MI incidence.

Increasing PA by 11.25 MET hours per week was associated with a significant decrease in risk for all of the cardiovascular outcomes. The protective association for CVD mortality (RR 0.77) was greater than for CVD incidence (RR 0.83), a result that was also shown when restricting both meta-analyses to studies that reported results for both CVD mortality and incidence. Figure 3 displays the meta-analysis of the effect of PA on CVD mortality after adjustment for body weight for the included studies. The meta-analyses for the remaining conditions are shown in figures S1-S7 in the Supplementary material.

The results that were adjusted for body weight were only slightly attenuated in comparison with the results that were not adjusted for body weight. The greatest effect of body weight adjustment was noted in the relative risk of CVD mortality, which reduced from 0.66 (unadjusted) to 0.77 (adjusted for body weight).

The meta-analysis results for the various sub-categories of CVD outcomes (such as MI incidence) showed that the estimated relative risks were very similar to CVD as a whole. The 0.25 power transformation for overall CVD incidence was 0.83, versus 0.82 for stroke incidence, 0.80 for CHD incidence, 0.81 for heart failure incidence and 0.75 for myocardial infarction incidence. The dose-response graphs (Supplementary Figures S8-S9) show that the greatest risk reduction was observed when moving from inactive to moderate physical activity. Therefore one can extrapolate that the

greatest benefit may be derived from an additional 6 METS hours/week, with a risk reduction of approximately 4.3% per METhr/week for CVD mortality and 1.7% for CVD incidence respectively.

Type II diabetes

Three studies were identified with incidence of type II diabetes as an outcome variable. These three studies were conducted in the UK, US and China respectively. They included a total of 261,618 participants and 19,417 events occurring during an average follow up period of 7.5 years.

The meta-analysis found a 0.74 relative risk (95% CI: 0.77- 0.72) for type II diabetes after adjustment for body weight and a marginally greater reduction in risk in models not adjusted for body weight.

The dose response curve shown in Supplementary Figure S10 demonstrates that the largest benefit may be derived when moving from inactive (0 METS) to 6 METs (RR 0.77), compared with 0.74 in response to 11.25 METhr/week of PA as per current guidelines.

Sensitivity analysis

For type II diabetes, CVD mortality and CVD incidence, the results of the meta-analysis were robust to the method of parameterising the dose-response relationship (Table 3). In general, the results obtained using the selected 0.25 transformation were similar to those obtained using other transformation for each of the disease outcomes, including both linear and loglinear relationships. For CVD incidence, CVD mortality, stroke incidence, CHD incidence, CHD mortality, heart failure incidence and myocardial infarction incidence, the 0.25 power transformation produced the result furthest from the null hypothesis. For instance, in estimating the relative risk of CVD incidence for an 11.25 METhr/week increase in PA ranged from 0.83 (0.25 power transformation), vs RR 0.89 (0.75 power transformation) and 0.92 (linear transformation).

Uniquely however, for T2DM incidence, the 0.25 power transformation produced a lower relative risk estimate as compared with the other transformations. For 0.25, the relative risk was 0.74, compared with a relative risk of 0.70 for the 0.75 transform and 0.68 for the linear relationship.

Estimates of relative risk for CHD incidence and heart failure incidence were more sensitive to the choice of parameterisation, with results for CHD incidence ranging from RR 0.80 (0.25 power transformation) to 0.90 (linear relationship).

Supplementary Table S5 demonstrates the effects of limiting the meta-analyses to those studies that achieved at least six out of the eight quality criteria. This restriction made very little difference to the results, with a slight increase in the observed relative risk reduction associated with an 11.25 METhr/week increase in PA for the vast majority of disease outcomes. Interestingly, for CHD incidence, the relative risk reduction attenuated slightly from 0.23 to 0.19 when including only the higher quality studies. As expected, however, there was some sensitivity when the quality criterion was used as a continuous measure, in that the very low quality studies over-estimated the impact of PA on health.

A further sensitivity analysis was conducted to exclude studies with PA levels that were deemed implausible i.e. PA levels that were excessively high or low. The thresholds for implausibility were PA levels exceeding 10 times the recommendations, or lower than 30 minutes of PA per week respectively (a level too low to be accurately measured by questionnaire). For CVD mortality, only three out of seventeen studies exceeded the maximum threshold, and only one study was under the minimum threshold (i.e. recommended PA levels of 11.25METs (Supplementary Figure A1). However, Supplementary Table S5 demonstrates that repeating the analysis without these implausible studies made negligible difference to the overall results. The highest level of PA in most CVD studies was around two to three times the recommended 11.25 METhr/week (Supplementary Figure A2). The T2DM studies all had plausible PA level ranges, as demonstrated in Figure A3. The

highest exposure category had PA levels varying from 8.25 to 13.8 Mets, corresponding well to the recommended PA levels of 11.25 Mets/week.

Heterogeneity and assessment of bias

Heterogeneity was demonstrated through the I^2 statistics and examination of both the forest and funnel plots. Table 2 demonstrated significant heterogeneity in the study results for CVD and CHD mortality, which may well be a consequence of the varying populations with differing baseline measures of PA and varying methods of measurement. Of note however, inclusion of body-weight adjusted studies did reduce the I^2 statistics throughout.

Examination of the forest plot for CVD mortality, Figure 3 revealed that both the Framingham [19] and the Health Survey for England[20] studies showed considerably larger effect sizes than the remaining studies. Meta-regression was conducted to explore possible explanations of this heterogeneity.

Meta-regression

Supplementary Table S6 shows the effects of moderator variables on the relative risk of CVD mortality. Firstly, there was an association between quality of study and effect size, with a meta-regression coefficient of 0.07 per quality score point. In addition, it was noted that studies that measure bodyweight using subjective measurements estimate a 0.24 lower relative risk as compared with those that use objective measurements. In assessing the effects of geographical location, US cohort studies estimated a relative risk further from the null hypothesis, with a difference of approximately 0.15, as compared with studies conducted elsewhere.

Small study bias or publication bias

Small study effects, or publication bias, was assessed through visual inspection of the funnel plots, as shown in Figure 4 for CVD mortality (the remaining conditions displayed in Supplementary Figures S12-S18). In general, the plots were highly symmetrical, showing little signs of publication bias. In certain disease outcomes, the smaller, less powerful studies often overestimated the effects of PA. For example, in CVD incidence, the Framingham male cohort, which contributed a mere 0.4% weighting provided a relative risk estimate of 0.53 as compared with the overall estimated relative of CVD incidence of 0.83.

In contrast however, for CHD incidence, the smallest study reported an increase in disease outcome for increasing PA levels. The Belstress men's cohort[21], which was afforded a weighting of 0.50% reported a relative risk of 1.12 compared with the overall relative risk estimate of 0.80 for CHD incidence. These data points highlighted the effect of small study bias, which could be attributed to clinical or methodological diversity.

DISCUSSION

We found a decrease in the risk of all cardiovascular outcomes and diabetes incidence with increasing levels of PA. These relative risks were only marginally attenuated when adjusting for a measure of body weight, suggesting that the majority of the health benefit that accrues from increasing PA is mediated by mechanisms beyond weight maintenance. Our findings suggest that an increase in 11.25 MET hours per week for an inactive individual is associated with a reduction of risk for cardiovascular mortality by 23% and diabetes incidence by 26%, independent of body weight. This may provide greater effect sizes compared to more recent studies such as Arem et al (2015)[22,] who used a value of 7.5 MET hours per week, however the key benefit of this paper is that the results can be generated for an MET value, including 7.5 to provide a direct comparison.

Strengths and weaknesses

This is the first meta-analysis to assess the effect of PA on CVD and diabetes using a continuous index of PA, thereby allowing for direct comparison of results across studies with heterogeneous data collection methods. Previous meta-analyses that have considered the effect of PA on health outcomes have mostly used categorical measures of PA (e.g. high vs moderate vs low) [23-25]. Those that used a single continuous comparable index of PA have not considered multiple health outcomes [26-29]. This study is the first to consider a range of cardiovascular diseases simultaneously, allowing for comparable results across these disease boundaries.

The results of these meta-analyses, which can be used to estimate the risk reduction associated with a unit increase of PA at any PA level, are vital for an accurate assessment of the population burden of physical inactivity. A recent study estimated the global burden of physical inactivity and concluded that the population health burden associated with physical inactivity was of an equivalent size to that associated with smoking [30]. However, this analysis has been criticised for over-estimating the burden of PA [31], as the authors derived their population impact fractions (PIFs) for physical inactivity from RRs drawn from meta-analyses which either compare pooled estimates of 'low' PA with 'high' PA, or compare inactive behaviour with meeting PA recommendations (although other assumptions by the authors are likely to result in under-estimation) [30]. The appropriate RRs for such PIFs should compare PA at the average level within the inactive population with the level required to meet recommendations. Such RRs can only be derived from analyses similar to those reported here, which account for the continuous nature of the PA variable.

Our results supported the assumption of other meta-analysis that the inactive have most to gain by any increase in PA. The relationship between PA and health outcomes is such that a small increase from inactive behaviour provides most of the benefit, and subsequent increases produce diminishing

returns. This is supported by two previous meta-analyses that assessed the dose-response relationship between PA and coronary heart disease [27] and all-cause mortality [26]. The Woodcock meta-analysis suggested that a first degree fractional polynomial with 0.25 power (as has been used in this study) provides the best fit to the data [26]. However, a drawback of using such a transformation is that it is necessary to set a somewhat arbitrary 'zero' level of PA. We assumed that the reference group from each of the studies achieved zero PA (i.e. they were considered to be inactive), but the actual PA level in these groups will vary between studies. The fact that the included studies tended to show the same 'diminishing returns' relationship with PA suggests that the levels of PA in the reference groups were fairly similar and approximately sedentary. However, this 'diminishing returns' relationship could be evidence of a general bias present in the studies inflating the risk of disease in the PA reference groups. We explored the effect on meta-analysis results of method of parameterising the dose-response relationship and also compared results with non-parametric categorical analyses and in most cases found that the results were robust to choice of method.

A limitation of using the comparable index of PA is that it does not distinguish between sustained periods of moderate activity and short periods of vigorous activity, which may have differing effects on health. Also, the PA recommendations [32] advise either 150 minutes of moderate to vigorous activity (150MVPA) or 75 minutes of vigorous physical activity (75VPA), whereas in this analysis we have focused on the more commonly used 150MVPA. Therefore our results pertain to 11.25 METhrs/week whereas for 75VPA (with 6.5METS for vigorous activity) the results would be presented for 9.75 METhrs/week, therefore this would attenuate the relative risk estimates. With respect to the search strategy, the potential of introducing bias into the review process by use of a single reviewer for reviewing full text articles and for the data extraction process cannot be

excluded, although a second reviewer cross-checked undecided cases and a 10% sample was checked by another reviewer at each stage of assessment.

The aim of this paper was to explore the effect of total PA, expressed as metabolic expenditure, as opposed to domain specific investigation. In reality, although many studies use a combination of domains to estimate total PA, they rarely truly include measurement of total PA exposure. We decided to include studies that had assessed two or more (of the four main) domains as this is likely to represent a reasonable minimum to be able to assess a meaningful proportion of overall PA exposure.

The studies included in this systematic review were mostly of high quality, notwithstanding the variability in PA measures. The prospective cohort study design provides some protection against recall or selection bias. There were few issues with loss to follow-up, or with detection of events. However, the measures of PA used in the studies were heterogeneous, both in terms of the measurement tools and in the aspects of PA that were being measured. As a result, the derivation of the comparable index of PA was challenging and is likely to have included some misclassification bias, as we were forced to make assumptions about behaviour and activity levels that were not reported in the studies. Usually misclassification bias would result in an underestimate of effect size, but here it may have led to a reduction in variance in PA levels, which will have led to a general bias away from the null hypothesis for the studies included in the meta-analysis.

It has previously been noted that studies of the effect of PA on health would benefit from improved measures of non-leisure time PA, particularly for studying the effect in women where levels of recreational activity are lower [24]. The use of self-reported PA questionnaires in all of the included

studies is problematic, as self-reported PA has been shown to have a low to moderate correlation with objective measures [33]. Furthermore, whilst all of the studies used healthy participants, only half included a 'burn-out' period to reduce the risk of reverse causation. The lack of availability of studies conducted in low and middle income countries precludes our ability to assess the impact of increasing PA levels in these settings, where the burden of physical inactivity and non-communicable diseases is also high. As all of the studies included in the meta-analysis were observational, the potential for residual confounding from imprecise or unmeasured factors cannot be excluded from pooling. Whilst all of the studies adjusted for multiple potential confounding variables, not all confounders were adjusted for in every study. Similarly, all of the results included in the meta-analysis were based on analyses using measures of PA and body weight made at only one time point, which is likely to result in under-estimates of the effect of PA on health.

Comparison with other studies

Three other reviews investigating the effects of PA on cardiovascular disease are summarised in table 5. Each review had different objectives and used different methods both to identify studies and to combine results. The majority have chosen to combine risk estimates based on categorical measures of the exposure variable. We believe that our point estimates for the RR associated with cardiovascular disease outcomes (increasing PA by 11.25 MET hours per week) correspond most closely with comparisons between 'moderate' and 'low' levels of PA. Our estimated RR of 0.83 for CHD incidence is similar to estimates from meta-analyses by Sofi et al [34] and Oguma et al [23] whereas our estimate for stroke incidence is of greater magnitude than that by Diep et al [24]. Our estimate for type 2 diabetes incidence is similar to that produced by Jeon et al[11].

Our study results for CVD mortality also proved similar to Arem et al[22] , which fell outside our search dates but was added later at the review stage. Arem et al studied the effects of leisure time physical

activity on all-cause and also CVD mortality, and found a CVD mortality HR of 0.80 between 0.1-7.5 METs and 0.67 corresponding to 7.5-15 METS of leisure-time PA. These estimates correspond extremely well with our estimates for CVD mortality at 11.25 METs, although the main focus of the Arem paper was to explore the potentially harmful effects of extremely high PA levels, up to 75 MET h/wk.

Conclusions

This meta-analysis has provided an assessment of the health benefits for a unit increase in PA levels, both before and after adjustment for body weight. The methods used here enable direct comparison of the effect of PA on a range of cardiovascular diseases and diabetes.

Future studies should investigate the effect of increasing PA levels in low and middle-income countries. Further analysis of the effect of increasing PA on other health behaviours such as diet and smoking, and an analysis of the potential differences that may arise with age, ethnicity and socioeconomic status for example, is also warranted. This meta-analysis provides clear direction for policy makers that there may be greater gains for population health by targeting those who do very little PA. However, population-level approaches to improving PA are still likely to be effective at reducing the health burden due to physical inactivity if the protective effect of PA on health is observed at high as well as low PA levels.

FOOTNOTES

Funding: This review was funded by the British Heart Foundation (grant number 021/P&C/Core/2010/HPRG). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Melanie Nichols is supported by funding from the National Heart Foundation of Australia and Deakin University for the Australian Heart Disease Statistics Project.

Ethical approval: Not required.

Data sharing: No additional data available.

Competing interests: The authors have declared that no competing interests exist.

References:

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, Anderson HR, Andrews KG, Aryee M, Atkinson C, Bacchus LJ, Bahalim AN, Balakrishnan K, Balmes J, Barker-Collo S, Baxter A, Bell ML, Blore JD, Blyth F, Bonner C, Borges G, Bourne R, Boussinesq M, Brauer M, Brooks P, Bruce NG, Brunekreef B, Bryan-Hancock C, Bucello C, Buchbinder R, Bull F, Burnett RT, Byers TE, Calabria B, Carapetis J, Carnahan E, Chafe Z, Charlson F, Chen H, Chen JS, Cheng AT, Child JC, Cohen A, Colson KE, Cowie BC, Darby S, Darling S, Davis A, Degenhardt L, Dentener F, Des Jarlais DC, Devries K, Dherani M, Ding EL, Dorsey ER, Driscoll T, Edmond K, Ali SE, Engell RE, Erwin PJ, Fahimi S, Falder G, Farzadfar F, Ferrari A, Finucane MM, Flaxman S, Fowkes FG, Freedman G, Freeman MK, Gakidou E, Ghosh S, Giovannucci E, Gmel G, Graham K, Grainger R, Grant B, Gunnell D, Gutierrez HR, Hall W, Hoek HW, Hogan A, Hosgood HD, 3rd, Hoy D, Hu H, Hubbell BJ, Hutchings SJ, Ibeanusi SE, Jacklyn GL, Jasrasaria R, Jonas JB, Kan H, Kanis JA, Kassebaum N, Kawakami N, Khang YH, Khatibzadeh S, Khoo JP, Kok C, Laden F, Lalloo R, Lan Q, Lathlean T, Leasher JL, Leigh J, Li Y, Lin JK, Lipshultz SE, London S, Lozano R, Lu Y, Mak J, Malekzadeh R, Mallinger L, Marcenes W, March L, Marks R, Martin R, McGale P, McGrath J, Mehta S, Mensah GA, Merriman TR, Micha R, Michaud C, Mishra V, Mohd Hanafiah K, Mokdad AA, Morawska L, Mozaffarian D, Murphy T, Naghavi M, Neal B, Nelson PK, Nolla JM, Norman R, Olives C, Omer SB, Orchard J, Osborne R, Ostro B, Page A, Pandey KD, Parry CD, Passmore E, Patra J, Pearce N, Pelizzari PM, Petzold M, Phillips MR, Pope D, Pope CA, 3rd, Powles J, Rao M, Razavi H, Rehfuess EA, Rehm JT, Ritz B, Rivara FP, Roberts T, Robinson C, Rodriguez-Portales JA, Romieu I, Room R, Rosenfeld LC, Roy A, Rushton L, Salomon JA, Sampson U, Sanchez-Riera L, Sanman E, Sapkota A, Seedat S, Shi P, Shield K, Shivakoti R, Singh GM, Sleet DA, Smith E, Smith KR, Stapelberg NJ, Steenland K, Stockl H, Stovner LJ, Straif K, Straney L, Thurston GD, Tran JH, Van Dingenen R, van Donkelaar A, Veerman JL, Vijayakumar L, Weintraub R, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams W, Wilson N, Woolf AD, Yip P, Zielinski JM, Lopez AD, Murray CJ, Ezzati M, AlMazroa MA, Memish ZA. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the global burden of disease study 2010. *Lancet*. 2012;380:2224-2260.

2. Paffenbarger RS, Hale WE. Work activity and coronary heart mortality. *N Engl J Med*. 1975;292:545-550.

3. Morris JN. Exercise versus heart attack: History of a hypothesis. 2005.

4. Health Do. On the state of public health: Annual report of the chief medical officer 2009. 15 March 2010.

5. Organisation WH. Global recommendations on physical activity for health. 2010.

6. Chief Medical Officers of England S, Wales, and Northern Ireland. Start active, stay active: A report on physical activity from the four home countries' chief medical officers. 2011.

7. US Department of Health and Human Services. 2008 physical activity guidelines for americans. 2008.

8. Lee IM, Paffenbarger RS, Jr. Associations of light, moderate, and vigorous intensity physical activity with longevity. The harvard alumni health study. *American journal of epidemiology*. 2000;151:293-299.

9. Fujita K, Takahashi H, Miura C, Ohkubo T, Sato Y, Ugajin T, Kurashima K, Tsubono Y, Tsuji I, Fukao A, Hisamichi S. Walking and mortality in japan: The miyagi cohort study. *Journal of epidemiology / Japan Epidemiological Association*. 2004;14 Suppl 1:S26-32.

10. Woodcock J, Franco OH, Orsini N, Roberts I. Non-vigorous physical activity and all-cause mortality: Systematic review and meta-analysis of cohort studies. *International journal of epidemiology*. 2011;40:121-138.

11. Jeon CY, Lokken RP, Hu FB, van Dam RM. Physical activity of moderate intensity and risk of type 2 diabetes: A systematic review. *Diabetes care*. 2007;30:744-752.

12. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The newcastle-ottawa scale (nos) for assessing the quality of nonrandomised studies in meta-analyses.

13. Kelly P, Kahlmeier S, Gotschi T, Orsini N, Richards J, Roberts N, Scarborough P, Foster C. Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship. *The international journal of behavioral nutrition and physical activity*. 2014;11:132.

14. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr., Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. 2011 compendium of physical activities: A second update of codes and met values. *Med Sci Sports Exerc*. 2011;43:1575-1581.

15. Moody A. Adult anthropometric measures, overweight and obesity. *Health Survey for England*. 2012.

16. Alaike H MI. A new look at the statistical model identification. *IEE Transactions on Automatic Control*. 1974;19:716-723.

17. Egger M, Smith GD, Phillips AN. Meta-analysis: Principles and procedures. *Bmj*. 1997;315:1533-1537.

18. Wattanakit K, Lutsey PL, Bell EJ, Gornik H, Cushman M, Heckbert SR, Rosamond WD, Folsom AR. Association between cardiovascular disease risk factors and occurrence of venous thromboembolism. A time-dependent analysis. *Thrombosis and haemostasis*. 2012;108:508-515.

19. Sherman SE, D'Agostino RB, Cobb JL, Kannel WB. Does exercise reduce mortality rates in the elderly? Experience from the framingham heart study. *American heart journal*. 1994;128:965-972.

20. Soedamah-Muthu SS, De Neve M, Shelton NJ, Tielemans SM, Stamatakis E. Joint associations of alcohol consumption and physical activity with all-cause and cardiovascular mortality. *The American journal of cardiology*. 2013;112:380-386.

21. Clays E, De Bacquer D, Janssens H, De Clercq B, Casini A, Braeckman L, Kittel F, De Backer G, Holtermann A. The association between leisure time physical activity and coronary heart disease among men with different physical work demands: A prospective cohort study. *European journal of epidemiology*. 2013;28:241-247.

22. Arem H, Moore SC, Patel A, Hartge P, Berrington de Gonzalez A, Visvanathan K, Campbell PT, Freedman M, Weiderpass E, Adami HO, Linet MS, Lee IM, Matthews CE. Leisure time physical activity and mortality: A detailed pooled analysis of the dose-response relationship. *JAMA internal medicine*. 2015;175:959-967.

23. Oguma Y, Shinoda-Tagawa T. Physical activity decreases cardiovascular disease risk in women: Review and meta-analysis. *Am J Prev Med*. 2004;26:407-418.

24. Diep L, Kwagyan J, Kurantsin-Mills J, Weir R, Jayam-Trouth A. Association of physical activity level and stroke outcomes in men and women: A meta-analysis. *J Womens Health (Larchmt)*. 2010;19:1815-1822.

25. Nocon M, Hiemann T, Muller-Riemenschneider F, Thalau F, Roll S, Willich SN. Association of physical activity with all-cause and cardiovascular mortality: A systematic review and metaanalysis. *Eur J Cardiovasc Prev Rehabil*. 2008;15:239-246.

26. Woodcock J, Franco OH, Orsini N, Roberts I. Non-vigorous physical activity and all-cause mortality: Systematic review and meta-analysis of cohort studies. *Int J Epidemiol*. 2011;40:121-138.

27. Sattelmair J, Pertman J, Ding EL, Kohl HW, 3rd, Haskell W, Lee IM. Dose response between physical activity and risk of coronary heart disease: A meta-analysis. *Circulation*. 2011;124:789-795.

28. Bao Y, Michaud DS. Physical activity and pancreatic cancer risk: A systematic review. *Cancer Epidemiol Biomarkers Prev*. 2008;17:2671-2682.

29. O'Rorke MA, Cantwell MM, Cardwell CR, Mulholland HG, Murray LJ. Can physical activity modulate pancreatic cancer risk? A systematic review and meta-analysis. *Int J Cancer*. 2010;126:2957-2968.

30. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *Lancet*. 2012;380:219-229.

31. Scowcroft H. Is being inactive really 'as bad as smoking'? *Science Update Blog*. 2012;2012.

32. Chief Medical Officers of England S W, and Northern Ireland Start active, stay active: A report on physical activity from the four home countries 2011.

33. Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *The international journal of behavioral nutrition and physical activity*. 2008;5:56.

34. Sofi F, Capalbo A, Cesari F, Abbate R, Gensini GF. Physical activity during leisure time and primary prevention of coronary heart disease: An updated meta-analysis of cohort studies. *Eur J Cardiovasc Prev Rehabil*. 2008;15:247-257.

35. Sofi F, Valecchi D, Bacci D, Abbate R, Gensini GF, Casini A, Macchi C. Physical activity and risk of cognitive decline: A meta-analysis of prospective studies. *J Intern Med*. 2011;269:107-117.

36. Diep L, Kwagyan J, Kurantsin-Mills J, Weir R, Jayam-Trouth A. Association of physical activity level and stroke outcomes in men and women: A meta-analysis. *J Womens Health (Larchmt)*. 2010;19:1815-1822.

Table 1. Deriving a modelled estimate of the association between PA and health outcomes at any level of PA using the results from these meta-analyses

The results reported in this paper (displayed in table 3) collapse the modelled relationship between PA and health outcomes across a continuous PA metric to a single parameter in order to provide comparable results across disease outcomes. These parameters can be used to estimate the modelled association between PA and health outcome for the difference between any two levels of PA. For example, the meta-analysis of CVD incidence using an *Additional* or *Marginal* METs approach suggests that the Relative Risk (RR) for a change from 0 METhr/d to 1.61METhr/d (equivalent to 11.25MET/hr/week) is 0.83 These values can be put into equation (1) to obtain the estimate b=-0.15.

(1) $RR = 1 + bMEThr/d^{0.25}$

This can then be used to estimate the RR for a unit increase of METhr/d at any PA level. For example, the RR of CVD incidence associated with a change from 2METhr/d to 5METhr/d is estimated as follows:

$RR_2 = 1 - 0.15.2^{0.25} = 0.82$

$RR_5 = 1 - 0.15.5^{0.25} = 0.77$

$$RR_{2\to5} = \frac{RR_5}{RR_2} = 0.94$$

Table 2. Meta-analysis results for effect of increase in physical activity equivalent to moving from inactivity to achieving current recommendations (11.25 MET hours/week, for cardiovascular disease incidence and mortality and type II diabetes incidence – assuming a 0.25 power transformation

			Adjusted for body weight		*Not adjusted for body weight*	
Condition (ICD-10 code)	Number of contributing studies	Total number of events	RR (95% CI)	I^2	RR (95% CI)	I^2
CVD incidence (I00-I99)	5	6,945	0.83 (0.77, 0.89)	0.0%	0.79 (0.72, 0.87)	33.3%
CVD mortality (I00-I99)	14	39,708	0.77 (0.71, 0.84)	73.6%	0.66 (0.52, 0.84)	93.6%
Stroke incidence (I60-I69)	9	13,599	0.82 (0.77, 0.87)	0.0%	0.78 (0.69, 0.88)	3.1%
CHD incidence (I20-I25)	6	12,655	0.80 (0.75, 0.86)	0.0%	0.77 (0.71, 0.83)	0.0%
CHD mortality (I20-25)	2	1,022	0.80 (0.58, 1.09)	59.1%	n/a	n/a
Heart failure incidence (I50)	5	9,457	0.81 (0.76, 0.86)	0.0%	0.75 (0.69, 0.82)	0.0%
Myocardial infarction incidence (I21-22)	2	6,445	0.75 (0.62, 0.89)	0.0%	n/a	n/a
Type 2 diabetes incidence (E11)	3	19,417	0.74 (0.72, 0.77)	0.0%	0.73 (0.68, 0.79)	56.0%

n/a - too few studies for a meta-anal

Table 3. Meta-analysis results for 11.25 MET hours /week increase on cardiovascular diseases and type 2 diabetes: sensitivity to transformation assumptions

	RR for 11.25 METh/w increase in PA, with 95% confidence intervals					
Health outcome (ICD-10 code)	0.25 power	0.375 power	0.5 power	0.75 power	Linear	Loglinear
CVD incidence (I00-I99)	0.83 (0.77, 0.89)	0.85 (0.79, 0.91)	0.86 (0.80, 0.92)	0.89 (0.83, 0.96)	0.92 (0.86, 0.98)	0.91 (0.85, 0.97)
CVD mortality (I00-I99)	0.77 (0.71, 0.84)	0.78 (0.70, 0.87)	0.79 (0.70, 0.90)	0.80 (0.67, 0.95)	0.80 (0.63, 1.01)	0.82 (0.70, 0.95)
Stroke incidence (I60-I69)	0.82 (0.77, 0.87)	0.82 (0.77, 0.87)	0.82 (0.77, 0.88)	0.84 (0.77, 0.91)	0.85 (0.77, 0.94)	0.85 (0.77, 0.93)
CHD incidence (I20-I25)	0.80 (0.75, 0.86)	0.82 (0.77, 0.88)	0.84 (0.79, 0.90)	0.88 (0.82, 0.93)	0.90 (0.85, 0.96)	0.89 (0.84, 0.95)
CHD mortality (I20-25)	0.80 (0.58, 1.09)	0.81 (0.61, 1.08)	0.82 (0.63, 1.07)	0.84 (0.67, 1.05)	0.86 (0.70, 1.05)	0.85 (0.68, 1.06)
Heart failure incidence (I50)	0.81 (0.76, 0.86)	0.83 (0.79, 0.89)	0.86 (0.81, 0.91)	0.89 (0.84, 0.95)	0.92 (0.87, 0.98)	0.91 (0.86, 0.97)
Myocardial infarction incidence (I21-22)	0.75 (0.62, 0.89)	0.76 (0.63, 0.91)	0.77 (0.65, 0.93)	0.81 (0.67, 0.97)	0.84 (0.70, 1.00)	0.83 (0.69, 0.99)
Type 2 diabetes incidence (E11)	0.74 (0.72, 0.77)	0.73 (0.71, 0.76)	0.72 (0.70, 0.75)	0.70 (0.68, 0.72)	0.68 (0.66, 0.70)	0.69 (0.67, 0.71)

Table 4. Categorical analyses of dose response relationship of physical activity on cardiovascular diseases and type 2 diabetes, compared to baseline of inactive behaviour

Health outcome	Low physical activity (0.1 – 11.5 MET hrs / week)	Medium physical activity (11.5 – 29.5 MET hrs / week)	High physical activity (29.5+ MET hrs / week)
CVD incidence (I00-I99)	0.89 (0.82, 0.98)	0.79 (0.69, 0.89)	0.75 (0.64, 0.87)
CVD mortality (I00-I99)	0.72 (0.67, 0.77)	0.72 (0.66, 0.78)	0.73 (0.67, 0.79)
Stroke incidence (I60-I69)	0.85 (0.80, 0.91)	0.81 (0.74, 0.88)	0.76 (0.68, 0.85)
CHD incidence (I20-I25)	0.87 (0.80, 0.95)	0.78 (0.74, 0.82)	0.70 (0.66, 0.75)
CHD mortality (I20-25)	*N/A*	0.76 (0.63, 0.93)	*N/A*
Heart failure incidence (I50)	*N/A*	0.79 (0.72, 0.85)	0.74 (0.68, 0.79)
Myocardial infarction incidence (I21-22)	*N/A*	0.76 (0.66, 0.87)	*N/A*
Type 2 diabetes incidence (E11)	0.77 (0.74, 0.80)	0.70 (0.54, 0.90)	*N/A*

N/A - Too few data points for meta-analysis

Table 5. Comparison with results from other meta-analyses of the effect of physical activity on cardiovascular disease outcomes reported in peer-reviewed journals

Citation	Health outcome	Physical activity comparison	Relative risk	95% confidence intervals	Studies included in meta-analysis	Additional Notes
Oguma et al., (2004)[23]	CHD (incidence or mortality)	Categorical: Moderate vs low	0.77	(0.64-0.92)	2	Includes case-control studies and retrospective cohort studies
		Categorical: High vs low	0.57	(0.41-0.79)		
Sofi et al., (2008)[35]	CHD (incidence or mortality)	Categorical: Moderate vs low	0.88	(0.83-0.93)	22	Leisure time PA only
		Categorical: High vs low	0.73	(0.66-0.80)		
Diep et al 2010 [36]	Stroke (incidence or mortality)	Categorical: High vs moderate vs low	0.81 0.89	(0.75 – 0.87) (0.86-0.93)	13	
Jeon et al 2006[11]	Type 2 Diabetes Incidence	Categorical: moderate vs sedentary	0.69	(0.58 – 83)	5	Not adjusted for body weight
Arem et al (2015)[22]	CVD Mortality	Categorial: low (0.1-7.5 METs)	0.80	0.77-0.84	6	Adjusted for body weight
		Categorial: low (0.1-7.5 METs)	0.67	0.65-0.80		

Figure 1. Flow chart for inclusion of studies

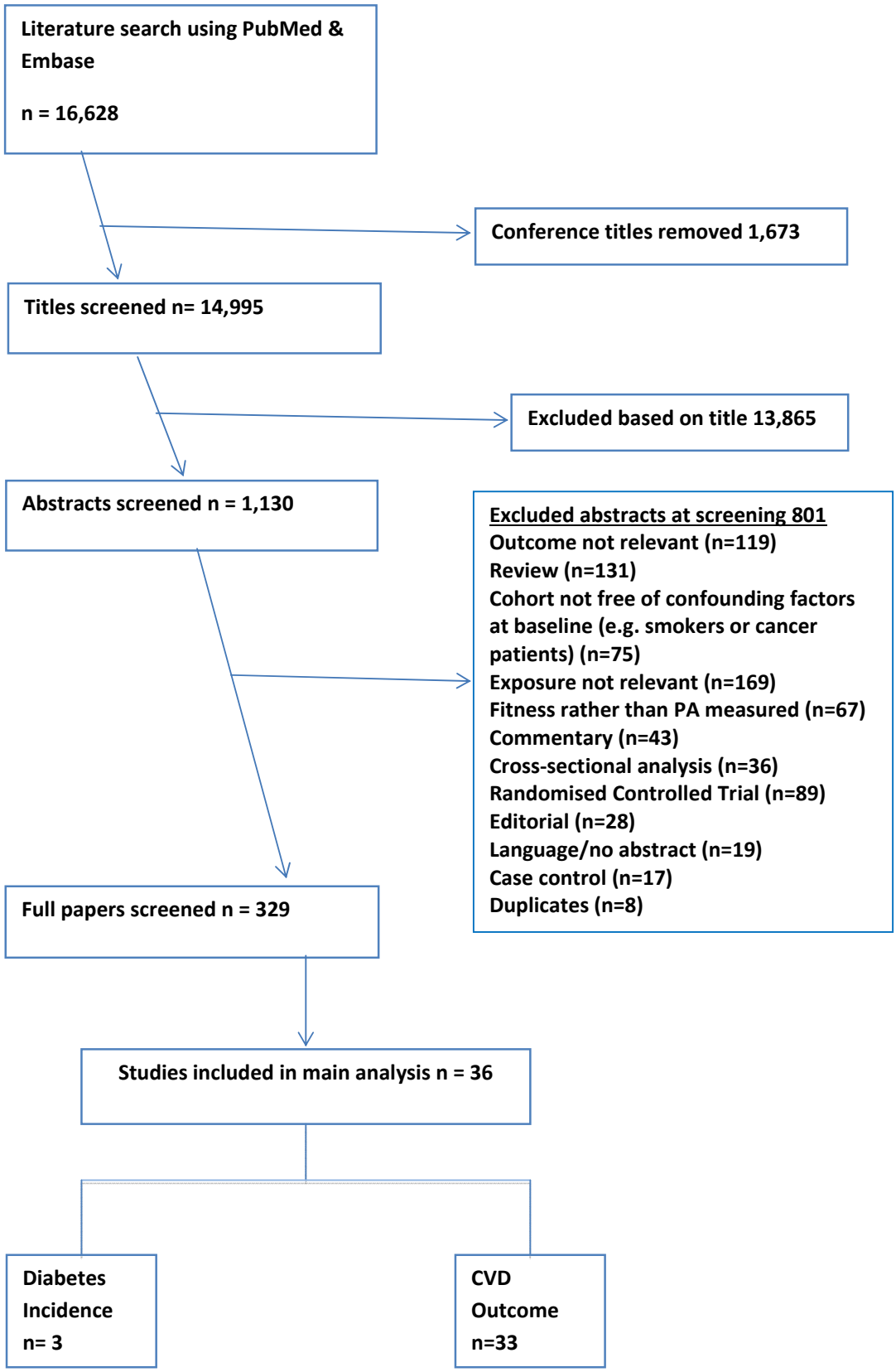
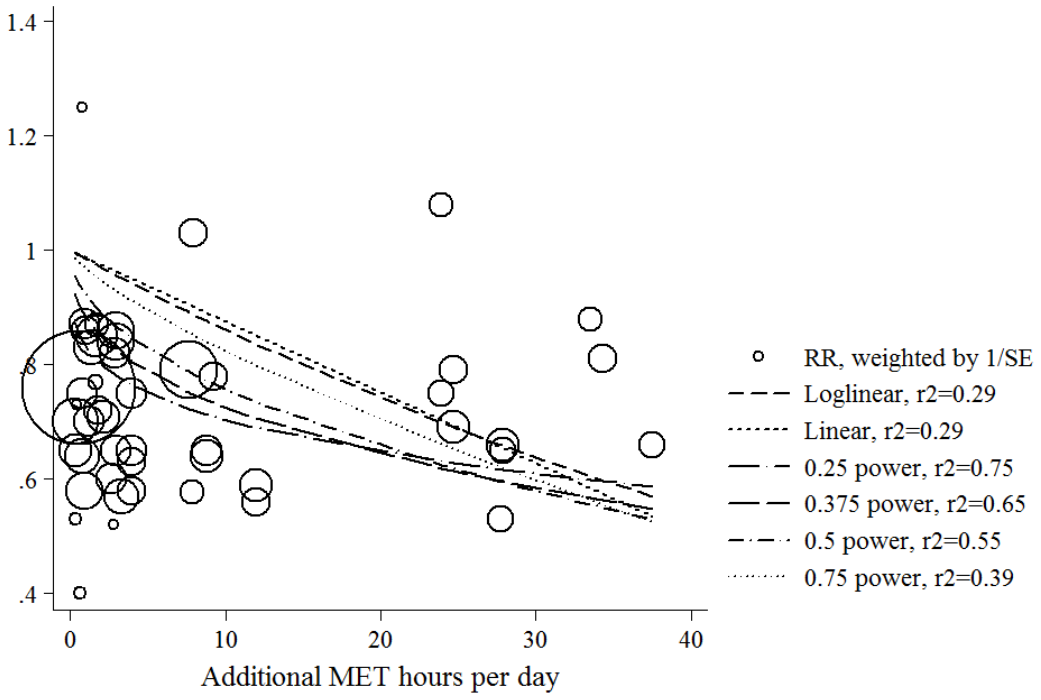
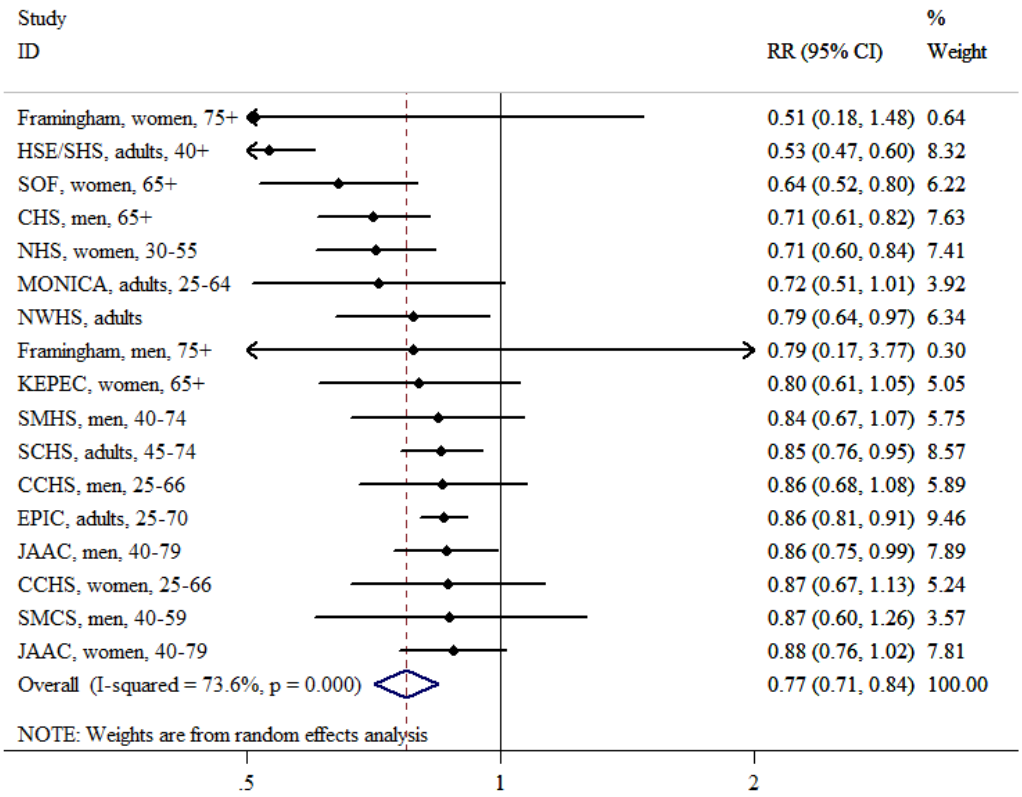


Figure 2. Relative risk for CVD mortality against MET hours per day. Results from 14 studies, including the 0.25 power transformation fit line as well as linear, log-linear, 0.375. 0.5 and 0.750 power transformations



Relative risk estimates are weighted by the inverse of the reported standard error, with larger circles for results with greater weighting. The red line represents a log-linear transformation, and the orange line represents a 0.25 power transformation.

Figure 3. Meta-analysis of 11.25 METhr/week increase in physical activity on CVD mortality, with a 0.25 power transformation, adjusted for body weight



Figures 4. Funnel plot for meta-analyses of 11.25 METhr/week increase in physical activity, with a 0.25 power transformation, for CVD mortality, adjusted for body weight