IJE-2018-05-0566: Revision #1

Examining associations between physical activity and cardiovascular mortality using negative control outcomes

Mark Hamer,[1,2] Adrian Bauman,[3,4] Joshua A. Bell,[5,6] Emmanuel Stamatakis[3,4]

[1]School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough LE11 3TU, United Kingdom.

[2]Department of Epidemiology and Public Health, University College London, London WC1E 6BT, United Kingdom.

[3]Charles Perkins Centre, University of Sydney, Sydney, Australia.

[4]Prevention Research Collaboration, School of Public Health, University of Sydney, Sydney, Australia.

[5]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

[6]Population Health Sciences, Bristol Medical School, Bristol, UK

*Correspondence: Prof Mark Hamer, Loughborough University, Loughborough LE11 3TU, United Kingdom. Phone: +44 (0) 1509 228473. Email: m.hamer@lboro.ac.uk

Word count: 1,738

Disclosures: None of the authors have any competing interests to declare.

Funding: Stamatakis is funded by the National Health and Medical Research Council (NHMRC) through a Senior Research Fellowship. Hamer is funded through the NIHR Leicester BRC; Bell through CRUK (C18281/A19169)

Contributor and guarantor information: Hamer conceptualized and designed the study, performed analyses, drafted the initial manuscript, and approved the final manuscript as submitted. He is the manuscript's guarantor; Stamatakis/Bauman/Bell conceptualized and designed the study, provided statistical input and critical revision of the manuscript, and approved the final manuscript as submitted

Abstract

Background: The purpose of a negative control is to reproduce a condition that cannot involve the hypothesized causal mechanism, but does involve the same sources of bias and confounding that may distort the primary association of interest. Observational studies suggest physical inactivity is a major risk factor for cardiovascular disease (CVD) although potential sources of bias, including reverse causation and residual confounding, make it difficult to infer causality. The aim was to employ a negative control outcome to explore the extent to which the association between physical activity and CVD mortality is explained by confounding.

Methods: The sample comprised 104,851 participants (aged 47 ± 17 yrs; 45.4% male) followed up over mean [SD] 9.4±4.5 years, recruited from The Health Survey for England and the Scottish Health Surveys.

Results: There were 10,309 deaths, of which 3,109 were attributed to CVD, and 157 to accidents (negative control outcome). Accidental death was related to age, male sex, smoking, longstanding illness and psychological distress, with some evidence of social patterning. This confounding structure was similar to that seen with CVD mortality, suggesting that our negative control outcome was appropriate. Physical activity (per SD unit increase in MET-hr-wk) was inversely associated with CVD (HR=0.75; 95% CI, 0.70, 0.80); the point estimate between physical activity and accidental death was in the same direction but of lesser magnitude (HR=0.86; 0.69, 1.07). A linear dose-response pattern was observed for physical activity and CVD but not with the negative control.

Conclusions: Inverse associations between physical activity and risk of CVD mortality are likely causal but of a smaller magnitude than commonly observed. Negative control studies have potential to improve causal inference within the physical activity field.

Key words: Physical activity; Cardiovascular disease; Negative control; Confounding

Key messages

- A negative control outcome was used to explore if associations between physical activity and CVD are explained by confounding.
- Our negative control (accidental death) had no plausible links with physical activity but displayed a similar confounding structure to CVD.
- The point estimate between physical activity and accidental death was in the same direction but of a lesser magnitude and gradient than physical activity and CVD.
- Inverse associations between physical activity and risk of CVD mortality are likely causal but of a smaller magnitude than commonly observed.

Introduction

Extensive epidemiological evidence has documented associations between moderate-to-vigorous physical activity (MVPA) and reduced risk of mortality from cardiovascular diseases (CVD).[1-4] It remains difficult, however, to firmly establish the causal nature of these associations since reverse causation, residual confounding, and selection biases are inherent limitations of observational data.[5,6] Randomised controlled exercise trials are a better test of causation, but these studies cannot realistically follow initially healthy participants for long periods of time in order to track incident disease risk. Trials also tend to be highly structured and thus unreflective of usual physical activity patterns. In addition participant characteristics can be poorly distributed between control and experimental groups and blinding is not possible in exercise trials.[7]

Common techniques of dealing with potential bias in observational epidemiology include statistical adjustment for measured confounding variables, removal of deaths occurring the first few years of follow up, and removal of participants with known disease at baseline. Although rarely used in the physical activity field, negative control comparisons are increasingly recognised as a means of strengthening causal inference.[8] The purpose of a negative control is to reproduce a condition that cannot involve the hypothesized causal mechanism, but does involve the same sources of bias and confounding that may distort the primary association of interest. If physical activity is similarly associated with a negative control (implausible) outcome as with the primary (plausible) outcome, this suggests that the association with the primary outcome is generated through pervasive confounding (e.g. is non-causal).

In the present study we aimed to employ a negative control outcome to explore the extent to which the association between MVPA and CVD mortality is explained by confounding. We chose accidental (external causes) deaths as our negative control outcome since these have no plausible links with physical activity but do have similar confounding structures to CVD mortality (Figure 1).

Materials and methods

Participants

The Health Survey for England (HSE) and the Scottish Health Survey (SHS) are household-based surveillance studies[9] consisting largely (91.2%) of a Caucasian sample. Participants in the present study took part in one of the following surveys: 1994 (HSE only), 1995 (SHS only), 1997 (HSE only), 1998 (HSE and SHS), 1999 (HSE only), 2003 (HSE and SHS), 2004, 2006, 2008 (HSE only). The proportion of eligible households that took part ranged from 66% to 81%, and 91.4% of participants gave consent to prospective linkage with mortality records. A multistage, stratified probability design was used to select participants to ensure representativeness of the target populations in their corresponding countries. Stratification was based on geographical areas and not on individual characteristics; postcode (zip code) sectors were selected at the first stage and household addresses selected at the second stage. Local research ethics committees approved all aspects of each survey and all participants gave written informed consent.

Physical activity assessment

The questionnaires used to assess physical activity along with their validity and reliability are detailed elsewhere.[10,11] In brief, trained interviewers enquired about the frequency (number of days in the last four weeks) and duration (of an average episode) of participation in: domestic physical activity; light-intensity (slow/average pace) and moderate-intensity (fairly brisk/fast pace) walking; and type-specific sports and exercises. For sports and exercises, there was a follow-up question about relative intensity: 'Was the effort of [activity] usually enough to make you out of breath or sweaty?' A compendium[12] was used to identify MVPA in the present study: moderate activities were of 3.0-5.9 metabolic equivalents (METs) and vigorous activities were of ≥6.0 METs, where one MET is considered to represent resting energy expenditure. Based on current United States guidelines[3] participants were

categorised as inactive (not reporting any MVPA), insufficiently active (>0<150 min/week of moderate intensity PA (MPA) and < 75 min/wk of vigorous PA (VPA)), or sufficiently active (at least 150 min/week MPA or 75 min/wk of VPA, or equivalent combinations of MVPA).

Covariates

Age and sex were self-reported. Participants were asked if they had ever regularly smoked a cigarette, cigar, or pipe, and if so, whether they presently smoked. Health status was assessed by asking participants whether they had 'any longstanding illness, disability of infirmity.' Socioeconomic position was assessed using the Registrar General's classification: professional and managerial occupations; skilled (non-manual/manual occupations); and, semi/unskilled (routine and manual) occupations. The 12-item version of the General Health Questionnaire (GHQ-12) was used to measure psychological distress, higher scores (range 0-12) indicating greater distress.

Mortality follow-up

Participants were flagged by the British National Health Service Central Registry. For participants who survived, the data were censored up to the end of 2009 (SHS) or the first quarter of 2011 (HSE). Diagnoses for the primary cause of death were based on the International Classification of Diseases, Ninth (ICD-9) and Tenth (ICD-10) Revisions. Codes for CVD mortality were 390-459 for ICD-9 and I01-I99 for ICD-10. The negative control outcome, accidental death, was identified from codes 800 – 999 (ICD-9) and codes V01-X59 (ICD-10), excluding cycling accidents (V10-19) and intentional self-harm (X60-84), which could have plausible links with physical activity.

Statistical analysis

Cox proportional hazards regression was used to estimate associations of physical activity with CVD mortality and negative controls, and to compare the confounding structures of the negative controls and CVD. Physical activity was modelled both as a categorical variable (no MVPA; insufficiently active; sufficiently active) and as a continuous variable (total MET-hr-wk). The proportional hazards assumption was examined by comparing the cumulative hazard plots grouped on exposure, although no appreciable violations were noted. The timescale was calendar time (months). We selected covariates a priori based on existing physical activity literature.[1,2,4] Models were adjusted for age, sex, social occupational class, longstanding illness, smoking, and psychological distress. We chose not to include body mass index as a covariate since it is unclear if obesity acts as a true confounder or more so as an intermediate mechanism. Analyses were performed using SPSS version 22 (IBM Inc.).

Results

The sample comprised 104,851 participants (aged 47 ± 17 yrs; 45.4% male). During 919,949 person years of follow up (mean [SD] 9.4±4.5 yr; range 0 – 17 years) there were 10,309 deaths, of which 3,109 were attributed to CVD, 157 to accidents.

Accidental death was related to age, male sex, smoking, longstanding illness and psychological distress, with some evidence of social patterning (Table 1). These patterns were similar to those seen with CVD mortality, suggesting that our negative control outcome was appropriate.

Relative to inactivity, we observed the expected associations of insufficient activity and sufficient activity with lower risk of CVD mortality (Table 2). An estimate of similar magnitude

but lower precision was observed for insufficient activity and risk of accidental death (HR=0.65; 95% CI, 0.40, 1.05). Weaker evidence of association was observed for sufficient activity and accidental death (HR=0.86; 0.54, 1.38), albeit the point estimate was in the same direction but of lesser magnitude compared with CVD (HR =0.58; 0.49, 0.68). When modelled as a continuous variable (per SD unit increase in MET-hr-wk) physical activity remained inversely associated with CVD (HR=0.75; 0.70, 0.80) and to a lesser extent with accidental death (HR=0.86; 0.69, 1.07).

In sensitivity analyses we removed deaths occurring in the first 2 years of follow up (n= 447 CVD; n=33 accidental deaths excluded) although results were not changed; physical activity (per SD unit increase in MET-hr-wk) was associated with CVD (HR=0.77; 0.71, 0.83) and to a lesser extent with accidental death (HR=0.86; 0.68, 1.10).

In further sensitivity analyses we considered other potential negative outcome controls, in particular certain cancers. Although there are very few cancers that consistently demonstrate null associations with physical activity[13] the relationship with prostate cancer seem less plausible; thus we conducted an additional analysis in men (n=47,586) using prostate cancer death (n=144 events) as the negative control. Prostate cancer had a less comparable confounding structure, although was strongly associated with age, longstanding illnesses, and smoking history (Table S1). The point estimate for sufficient activity and risk of prostate cancer death (HR=0.85; 95% CI, 0.49 – 1.48) was in the same direction but of lesser magnitude compared with CVD (Table S2).

Discussion

This study made use of negative control outcomes to examine the extent to which the association between MVPA and CVD mortality is likely generated by bias and confounding. Based on existing evidence physically active participants tend to be male, younger, have better general health, less likely to smoke, and come from more socially advantaged backgrounds.[14] These confounding factors are therefore commonly included as covariates in models that explore associations between physical activity and longevity. The negative control employed here appeared to be a suitable instrument since it demonstrated a similar confounding structure to our outcome of interest (CVD mortality). The point estimate for sufficient physical activity and both negative control outcomes was in the same direction but of lesser magnitude than that of CVD; thus a possible interpretation of our results is that confounding explains part of the association between physical activity and CVD. Previous data suggested around 59% of the inverse association between physical activity and CVD could be explained by plausible biological mechanisms, particularly inflammatory/hemostatic factors and blood pressure.[15] An association was found between insufficient activity and the negative control with a point estimate similar to that observed for CVD mortality. This would suggest that the potential health effects of lower levels of MVPA (not meeting guideline) are more strongly accounted for by pervasive confounding.

The number of events in the negative control outcome were limited (110, 25, and 22 accidental deaths in the inactive, insufficiently and sufficiently active groups, respectively) rendering these estimates imprecise. The models were adjusted for a specific set of covariates commonly cited in the literature but we cannot exclude the possibility of residual confounding. Our negative control outcome, accidental death, was largely chosen because we believed that accidents would be randomly distributed between physically active and inactive participants (for example, an accident when the participant was the occupant of a motor vehicle). We cannot exclude the possibility that some accidents could be more likely in frail/inactive participants, equally other events, such as those related to cycling, could be

more common in active people. However there were only two deaths related to cycling accidents and inclusion of these events did not alter the results.

Negative control studies have potential to improve causal inference within the physical activity field by illustrating pervasive confounding, but they are prone to the same degree of measurement error of confounders being examined as primary outcome associations.[16] Results from such studies should therefore be used in combination with existing[1-4] and emerging evidence[17] to improve causal estimation /inference of physical activity for longevity.

References

1. Wen CP, Wai JP, Tsai MK, Yang YC, Cheng TY, Lee MC, Chan HT, Tsao CK, Tsai SP, Wu X. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. Lancet. 2011;378(9798):1244-53.
2. Hupin D, Roche F, Gremeaux V, Chatard JC, Oriol M, Gaspoz JM, Barthélémy JC, Edouard P. Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged ≥60 years: a systematic review and meta-analysis. Br J Sports Med. 2015 Oct;49(19):1262-7.
3. Physical Activity Guidelines Advisory Committee. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Washington, DC: U.S. Department of Health and Human Services, 2018. Accessed March 2018. https://health.gov/paguidelines/second-edition/report/pdf/PAG_Advisory_Committee_Report.pdf
4. 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 Intern Med. 2015;175(6):959-67.
5. Wade KH, Richmond RC, Davey Smith G. Physical activity and longevity: how to move closer to causal inference. Br J Sports Med. 2018 Mar 15. pii: bjsports-2017-098995. doi: 10.1136/bjsports-2017-098995. [Epub ahead of print]
6. Kujala UM. Is physical activity a cause of longevity? It is not as straightforward as some would believe. A critical analysis. Br J Sports Med Published Online First: 15 March 2018. doi: 10.1136/bjsports-2017-098639
7. Krauss A. Why all randomised controlled trials produce biased results. Ann Med. 2018 Apr 4:1-11. doi: 10.1080/07853890.2018.1453233. [Epub ahead of print]
8. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010 May;21(3):383-8.

9. Mindell J, Biddulph JP, Hirani V, Stamatakis E, Craig R, Nunn S, Shelton N. Cohort profile: the health survey for England. Int J Epidemiol. 2012;41(6):1585-93.
10. Stamatakis E, Hillsdon M, Primatesta P. Domestic physical activity in relationship to multiple CVD risk factors. Am J Prev Med. 2007;32(4):320-327.
11. Scholes S, Coombs N, Pedisic Z, et al. Age- and sex-specific criterion validity of the health survey for England Physical Activity and Sedentary Behavior Assessment Questionnaire as compared with accelerometry. Am J Epidemiol. 2014;179(12):1493-1502.
12. Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011;43(8):1575-1581.
13. Moore SC, Lee IM, Weiderpass E, Campbell PT, Sampson JN, Kitahara CM, Keadle SK, et al. Association of Leisure-Time Physical Activity With Risk of 26 Types of Cancer in 1.44 Million Adults. JAMA Intern Med. 2016 Jun 1;176(6):816-25
14. Smith L, Gardner B, Fisher A, Hamer M. Patterns and correlates of physical activity behaviour over 10 years in older adults: prospective analyses from the English Longitudinal Study of Ageing. BMJ Open. 2015 Apr 15;5(4):e007423.
15. Mora S, Cook N, Buring JE, Ridker PM, Lee IM. Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. Circulation. 2007;116(19):2110-8.
16. Sanderson E, Macdonald-Wallis C, Davey Smith G. Negative control exposure studies in the presence of measurement error: implications for attempted effect estimate calibration. Int J Epidemiol. 2017 Oct 27. doi: 10.1093/ije/dyx213. [Epub ahead of print]
17. Doherty A, Smith-Bryne K, Ferreira T, Holmes MV, Holmes C, Pulit SL, Lindgren CM. GWAS identifies 10 loci for objectively-measured physical activity and sleep with causal roles in cardiometabolic disease. BioRxiv doi: https://doi.org/10.1101/261719

Figure caption

Figure 1. Depiction of an ideal negative control outcome (N) to evaluate the causal effect of physical activity on CVD in an observational study setting. N should ideally have the same incoming arrows as CVD, but a causal link between physical activity and N is implausible. Unmeasured confounding variables are designated as U.

Table 1. Comparison of confounding structure between negative outcome controls and cardiovascular disease mortality.

Measured confounders	Accidental death HR (95 % CI)	Cardiovascular disease HR (95 % CI)
Age (per year)	1.04 (1.03, 1.05)	1.11 (1.11, 1.12)
Sex		
Female	1.0 (Ref)	1.0 (Ref)
Male	2.08 (1.50, 2.88)	1.37 (1.27, 1.47)
Smoking		
Never	1.0 (Ref)	1.0 (Ref)
Ex-smoker	1.84 (1.24, 1.74)	1.31 (1.21, 1.42)
Current smoker	1.99 (1.36, 2.89)	2.15 (1.96, 2.37)
Social occupational class		
Professional /Managerial	1.0 (Ref)	1.0 (Ref)
Skilled, non-manual/manual	1.30 (0.88, 1.90)	1.32 (1.21, 1.44)
Semi-skilled / Unskilled	1.31 (0.85, 2.03)	1.52 (1.38, 1.67)
Longstanding illnesses		
None	1.0 (Ref)	1.0 (Ref)
Yes	1.80 (1.32, 2.48)	3.86 (3.57, 4.18)
Psychological distress (GHQ-12 score)		
0	1.0 (Ref)	1.0 (Ref)
1 – 3	1.20 (0.83, 1.73)	1.24 (1.14, 1.35)
4 – 7	0.96 (0.51, 1.81)	1.28 (1.13, 1.45)
8 – 12	1.97 (1.21, 3.22)	1.47 (1.30, 1.66)

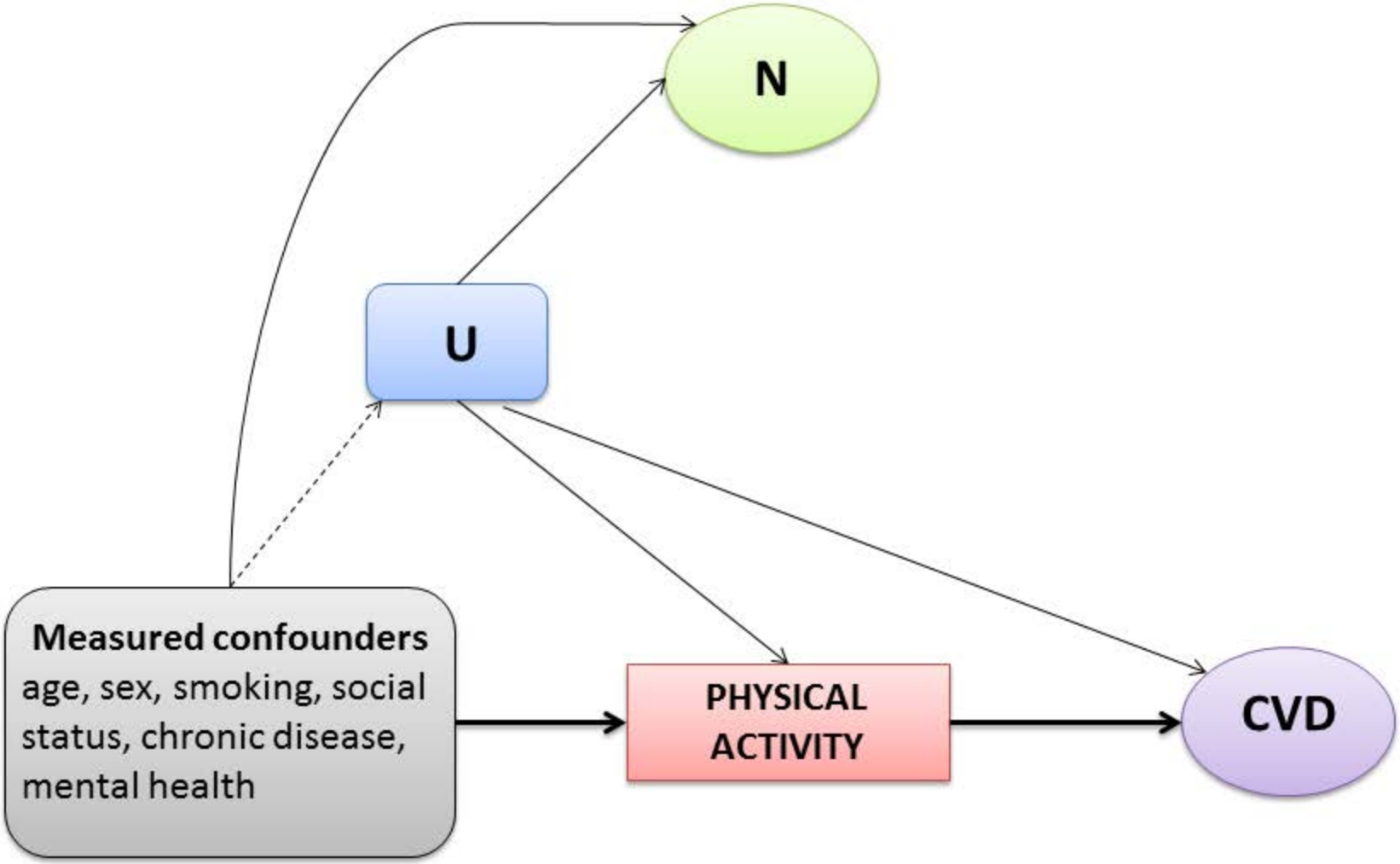
Hazard ratios (HR) are unadjusted

Table 2. Association between physical activity, CVD mortality and negative outcome control (n=104,851)

Moderate-to-vigorous physical activity	N	Accidental death HR (95 % CI)[1]	Cardiovascular disease death HR (95 % CI)[1]
Inactive	54,441	1.0 (Ref)	1.0 (Ref)
Insufficiently active	24,287	0.65 (0.40, 1.05)	0.62 (0.55, 0.71)
Sufficiently active	25,916	0.86 (0.54, 1.38)	0.58 (0.49, 0.68)

[1] Model adjusted for age, sex, smoking, social occupational class, longstanding illness, psychological distress.

Figure 1.


N
U
Measured confounders
age, sex, smoking, social status, chronic disease, mental health
PHYSICAL ACTIVITY
CVD