Customary physical activity and odds of depression: a systematic review and meta-analysis of 111 prospective cohort studies

Rodney K Dishman,[1] Cillian P McDowell ,[2] Matthew Payton Herring [3]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bjsports-2020-103140).

[1]Biomedical & Health Sciences Institute and Department of Kinesiology, University of Georgia, Athens, Georgia, USA
[2]The Irish Longitudinal Study on Ageing, Trinity College Dublin, Dublin, Ireland
[3]Health Research Institute and Department of Physical Education and Sports Sciences, University of Limerick, Limerick, Ireland

Correspondence to
Professor Rodney K Dishman, University of Georgia, Athens, GA 30602, USA; rdishman@uga.edu

Accepted 16 December 2020

Check for updates

© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Dishman RK, McDowell CP, Herring MP. *Br J Sports Med* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjsports-2020-103140

ABSTRACT

Objective To explore whether physical activity is inversely associated with the onset of depression, we quantified the cumulative association of customary physical activity with incident depression and with an increase in subclinical depressive symptoms over time as reported from prospective observational studies.

Design Systematic review and meta-analysis.

Data sources MEDLINE, PsycINFO, PsycARTICLES and CINAHL Complete databases, supplemented by Google Scholar.

Eligibility criteria Prospective cohort studies in adults, published prior to January 2020, reporting associations between physical activity and depression.

Study appraisal and synthesis Multilevel random-effects meta-analysis was performed adjusting for study and cohort or region. Mixed-model meta-regression of putative modifiers.

Results Searches yielded 111 reports including over 3 million adults sampled from 11 nations in five continents. Odds of incident cases of depression or an increase in subclinical depressive symptoms were reduced after exposure to physical activity (OR, 95% CI) in crude (0.69, 0.63 to 0.75; I^2=93.7) and adjusted (0.79, 0.75 to 0.82; I^2=87.6) analyses. Results were materially the same for incident depression and subclinical symptoms. Odds were lower after moderate or vigorous physical activity that met public health guidelines than after light physical activity. These odds were also lower when exposure to physical activity increased over time during a study period compared with the odds when physical activity was captured as a single baseline measure of exposure.

Conclusion Customary and increasing levels of moderate-to-vigorous physical activity in observational studies are inversely associated with incident depression and the onset of subclinical depressive symptoms among adults regardless of global region, gender, age or follow-up period.

INTRODUCTION

Depression is the leading cause of disability worldwide, affecting approximately 322 million people.[1] Scaled-up treatment of depression could yield a net global economic benefit of US$230 billion by 2030.[2] Major depression is prevalent and is a leading risk factor for cardiovascular morbidity and mortality.[3 4] Exercise has been recommended as a low-risk augmentation therapy for depression.[5–8]

English scholar Robert Burton recognised 'want of exercise' as 'the bane of body and minde … and the sole cause of melancholy' nearly 400 years ago.[9] In modern days, exercise training trials have shown reduced depressive symptoms among depressed[10–12] and chronically ill[13] patients, with effect sizes ranging from moderately large to small depending on the presumed rigour used to judge method bias in the studies.[10 14 15] More than 40 meta-analyses or otherwise systematic reviews confirm the benefits of exercise as therapy for depression.[16]

Whether exercise confers protection against the onset of depression, as Burton surmised, remains a key question in 2021. Two authoritative reviews[17 18] and two small meta-analyses[19 20] of a limited number of prospective, observational studies concluded that physical activity might aid primary prevention of depression. However, the reviews did not reconcile whether the reduced odds of depression depended on the amount of physical activity exposure or change in exposure, which are important for clinical practice and for satisfying minimal criteria (eg, dose–response, temporal sequence, accurate classification) in observational studies for a possibly causal association between physical activity exposure and depression.[21]

Therefore, we performed a systematic review and meta-analysis to extend the scope and elevate the methods of the few prior reviews. We expanded the focus on physical activity exposure and putative modifiers of exposure. We hypothesised that odds of incident depression (defined by clinical diagnosis or cut-points on depression screening tests) or participants having greater number of subclinical depressive symptoms over time would be inversely associated with either baseline physical activity dose or increased physical activity across time. We also examined a priori whether the expected lower odds varied according to gender, age, follow-up period, and the type and timing of measures of physical activity exposure and depressive outcome.

METHODS

Data sources and searches

A systematic review and meta-regression analysis was conducted in accordance with established practice for conduct and reporting.[22–24] The protocol satisfied contemporary standards.[25] Articles published from database inception to January 2020 were located by all authors using MEDLINE, PsycINFO, PsycARTICLES and CINAHL Complete. Keywords included combinations of 'physical activity' or 'leisure time' or 'exercise' or 'sport*' and 'depress*' or 'mood' or 'dysthymia' and 'association' or 'follow-up' or 'risk factor' or 'protect*' or 'causal*' or 'onset' or 'prospective' or 'cohort' or

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

'longitudinal'. Supplemental searches of articles citing, and cited by, included studies and relevant review articles were performed manually in Google Scholar. Inclusion criteria were: (1) a cohort of adults that was not depressed at baseline, according to criteria of clinical diagnosis or cut-points on depression screening test or when change in depressive symptoms was assessed in a population-based cohort; (2) physical activity exposure assessed at baseline or at baseline and one or more follow-up assessments; (3) a defined follow-up assessment when a depression outcome was assessed; (4) crude or adjusted ORs reported with SEs or CIs, or frequencies of exposed and non-exposed cases and non-cases, or test statistics that could be converted to ORs.[26] In one report, we received additional data from the corresponding author.[27]

We excluded investigations that (1) sampled patients with pre-existing clinically diagnosed depression or who were taking anti-depressants at baseline; (2) assessed peripartum or postpartum depression; (3) included physical activity as part of a multicomponent exposure; (4) used an adjective checklist rather than a validated scale of depression symptoms; (5) failed to adjust symptom outcomes for baseline symptoms; (6) reported associations that could not be converted to ORs (eg, HR or risk ratio). Online supplemental figure 1 illustrates a flowchart of study selection.

Physical activity was the exposure measured by single or multiple questions assessing participation in exercise, sports or physical activity (defined as bodily movement produced by skeletal muscles that requires energy expenditure) specified as: (1) frequency (typically per week or month), (2) time spent or distance covered, (3) as meeting public health guidelines for moderate or vigorous physical activity,[20 28] (4) a validated measure that estimated total volume (ie, frequency×time×intensity)[17] or metabolic equivalents (METs) expended per unit of time (eg, MET-hours) (see online supplemental table 1). Physical activity domains were categorised as leisure time (excluding chores, occupational work and active transport) and as total physical activity.

Depression was the outcome assessed as: (1) incident depression based on a diagnosis using a standardised, structured diagnostic interview, a self-reported physician diagnosis of depression, established cut-scores or otherwise elevated scores on validated screening tests for depression; or (2) increased subclinical depressive symptoms assessed by a validated questionnaire (see online supplemental table 1).

Online supplemental table 2 presents study characteristics. Studies were community-based or population-based, but only 11 studies reported crude[29] or adjusted[30–39] results separately for men and women and only 1 study reported results stratified by age.[34] Race and ethnicity were poorly represented (number of studies, median % of sample): African American or black (18, 16%), Mexican American, Latino, Hispanic (16, 15%), Asian (19, 5%), Native American (3, 2%) and were not described in two-thirds of the studies. A single study[40] compared black and white Americans. Nineteen studies sampled patients with chronic conditions other than depression, and 34 studies reported the proportion of the cohort with a comorbid chronic disease. Finally, just three studies included participants solely from low-to-middle income countries (each from Mexico). The baseline depression rate reported in 26 studies of symptom change was 13% (0%–20%), consistent with 12-month rates in US adults.[5 41]

Modifiers identified a priori were global region, gender, age, exposure dose, follow-up period, timing and measures of exposure and outcome (online supplemental table 1). Study quality was evaluated according to selection bias, confounding, exposure measures and outcome measures.[42] Consistent with the Grade Working Group (http://www.gradeworkinggroup.org/),[43] the quality of the summary evidence was judged as good, acceptable or poor according to: risk of bias, directness of evidence, consistency and precision of results, risk of publication bias, magnitude of the effect, dose–response gradient and influence of residual confounding (online supplemental table 3).

Data analysis

ORs were retrieved as published values, computed from 2×2 frequency tables of exposure cases and non-cases or converted from standardised mean differences, Pearson correlation coefficients or standardised linear regression coefficients using standard conversion procedures.[26 44 45] Reports from the same cohort were included when they added novel results based on different putative modifiers of the association between physical activity exposure and depression. Online supplemental table 4 describes reports from the same cohorts. Redundant or ancillary reports were excluded.[46 47]

ORs were log transformed, weighted by their inverse variance prior to random effects aggregation or mixed effects regression modelling and then back transformed to ORs for summary reports. Precision is reported consistent with EQUATOR guidelines.[48] Inter-rater reliability for ORs was examined with intraclass correlation coefficients (ICCs) for absolute agreement among three raters. Rater agreement on modifier levels and study quality was judged with Kappa.[49] Initial rater agreements exceeded 0.95 for ORs and ranged from 90% to 100% for three rater agreements on modifiers (kappa exceeded 0.75) and study quality (kappa was 0.64–0.89 for five facets and 0.80–0.95 for overall quality). Discrepancies were resolved by adjudication by three raters after recalculation and/or recoding.

A multilevel, random-effects model estimated parameters and their SEs after adjustment for correlated effects within random nesting factors of study, cohort or global region. Overall mean effect size was calculated and mixed-model multiple-linear regression modifier analysis with restricted information maximum likelihood estimation[50 51] was conducted using Metafor[52 53] in R V.3.5.[54] Knapp and Hartung adjustment[55 56] provided estimates of summary ORs and their variation that are precise and robust to bias.[57] Results were corroborated by Bayes estimation in Mplus V.8.2,[58] using standard procedures.[59–61] Each modifier variable was coded according to planned contrasts among its levels (see tables 1 and 2) and centred for meta-regression. The reference for each OR for each type of physical activity measure reported in tables 1 and 2 was the lowest grouping level, not unit of measurement, for each type of measure (ie, lowest frequency; lowest time or distance; lowest MET or volume; and not meeting guidelines). Odds from each study's most adjusted model were used for adjusted summary ORs, which were compared according to the extent of adjustment for confounders (see table 2 and online supplemental table 1). Tests of the regression model and its residual error (Q_E) are reported. Type I error was constrained by testing each main effect at $p<0.01$. Heterogeneity of mean effects was tested with Q and I^2 statistics.[62] Funnel plots with Egger's test[63] and the rank correlation test[64] examined risk of publication bias. Fail-safe N_1 estimated the size of a cohort with null effects that would overturn the observed effects.[65]

RESULTS

Sixty-eight of 104 (65%) crude odds and 88 of 179 (49%) adjusted odds were statistically significant. Sample sizes were too small to provide good precision of the point estimates in most other cohorts. The median (IQR) number of people for

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

Table 1 Depression symptoms associated with physical activity: crude OR and heterogeneity

	Contrast	K (N)	OR	95% CI	Q	I^2 (95% CI)
Region						
North America		41 (20)	0.72****	0.64 to 0.82	469.1****	91.7 (90.7 to 92.6)
Europe		30 (18)	0.64****	0.52 to 0.79	301.5****	90.7 (89.4 to 91.9)
Asia		22 (9)	0.71****	0.63 to 0.78	446.4****	75.4 (70.5 to 79.5)
Australia		11 (4)	0.64****	0.57 to 0.71	13.5	63.7 (50.5 to 73.4)
Gender						
Females only	1	29 (8)	0.73***	0.62 to 0.86	385.8****	93.0 (92.1 to 93.8)
Males only	−1	13 (6)	0.72**	0.58 to 0.91	23.4*	53.0 (34.8 to 66.1)
Mixed (55% female, 8%–94%, IQR=49%–59%)	0	62 (37)	0.67****	0.60 to 0.76	931.0****	93.6 (93.0 to 94.1)
Age (years)						
18–44	1	38 (15)	0.74****	0.67 to 0.82	162.3****	77.8 (74.1 to 81.0)
45–64	2	38 (18)	0.66***	0.53 to 0.81	791.8****	95.5 (95.0 to 95.9)
65+	3	28 (18)	0.68****	0.60 to 0.77	282.1****	90.8 (89.4 to 92.0)
Study quality						
Low	1	26 (17)	0.71****	0.58 to 0.88	432.4****	94.5 (93.7 to 95.1)
Acceptable	2	41 (22)	0.64****	0.56 to 0.73	996.1****	96.1 (95.7 to 96.4)
Good	3	37 (12)	0.73****	0.67 to 0.80	125.4****	72.1 (67.1.76.4)
Study design						
Baseline exposure and incident cases	−0.5	73 (31)	0.70****	0.64 to 0.74	1170.9****	93.4 (93.5 to 94.4)
Baseline exposure and change in symptoms	−0.5	9 (8)	0.84**	0.76 to 0.92	27.5***	74.6 (63.8 to 82.1)
Change in exposure and incident cases	0.5	10 (6)	0.73****	067 to 0.80	4.0	0.0 (0.0 to 0.0)
Change in exposure and change in symptoms	0.5	12 (9)	0.56***	0.37 to 0.84	381.8****	87.1 (85.0 to 88.8)
PA measure						
Frequency	−0.5	42 (21)	0.65****	0.57 to 0.75	623.5****	93.6 (92.9 to 94.2)
Time or distance	−0.5	10 (8)	0.73**	0.59 to 0.90	41.0****	76.1 (70.4 to 80.8)
Meeting guidelines (sufficient)	0.5	18 (5)	0.80****	0.73 to 0.87	106.4****	85.0 (81.5 to 87.8)
Metabolic equivalent	0.5	34 (19)	0.70***	0.59 to 0.83	473.2****	93.2 (92.4 to 94.0)
PA domain						
Leisure time	−0.5	89 (40)	0.70****	0.63 to 0.78	1531.2****	94.3 (93.9 to 94.7)
Total	0.5	15 (11)	0.65**	0.59 to 0.73	24.2****	46.3 (26.7 to 60.7)
Depression measure						
Symptoms	−1	21 (14)	0.65***	0.51 to 0.84	333.9****	94.3 (93.4 to 95.1)
Screening cut-point	0.33	53 (25)	0.70****	0.63 to 0.77	495.9****	89.7 (88.6 to 90.7)
Self-reported diagnosis	0.33	8 (4)	0.81****	0.77 to 0.86	58.5****	89.7 (86.2 to 92.4)
Clinical diagnosis	0.33	22 (8)	0.72***	0.62 to 0.84	394.1****	94.9 (94.2 to 95.6)
Exposure dose						
Binary or correlation	0	48 (35)	0.68****	0.60 to 0.77	710.0****	93.5 (92.8 to 94.1)
Low	−1	20 (16)	0.78****	0.70 to 0.87	61.3****	71.6 (64.2 to 77.4)
Moderate	0.5	17 (15)	0.69****	0.60 to 0.79	56.6****	73.5 (66.1 to 79.3)
High	0.5	19 (7)	0.69***	0.59 to 0.81	404.7****	95.8 (95.2 to 96.3)
Recommended exposure						
Not meeting guidelines	−1	15 (9)	0.68**	0.52 to 0.90	100.0***	87.0 (83.8 to 89.6)
Not reported	0	52 (34)	0.67****	0.59 to 0.76	1028.9****	95.1 (94.7 to 95.5)
Meeting guidelines	1	37 (18)	0.76****	0.69 to 0.83	464.1****	92.5 (91.6 to 93.3)

K is number of effects and N is number of studies for each category or level of each modifier.
Significant at *p<0.05, **p<0.01, ****p<0.001, ****p<0.0001.
IQR, interquartile range; PA, physical activity.

each comparison was 2247 (810–7558) people for crude odds and 1404 (597–3308) people for adjusted odds. Figures 1 and 2 illustrate forest plots and the distribution of crude and adjusted odds is annotated.

Crude odds

Physical activity was associated with 31% lower crude odds of study participants having incident depression or an increase in subclinical symptoms (OR=0.69 (95% CI: 0.63 to 0.75); Q(103)=1605.7, p<0.0001; I^2=93.7 (95% CI: 93.2 to 94.0); k=104, n=51). ORs were correlated within studies (ICC=0.23 (95% CI: −0.025 to 0.49)) and cohorts (ICC=0.25 (95% CI: −0.015 to 0.52)) but not regions (ICC=0.00). Fail-safe N_1 was 135, estimating that a null outcome from a cohort of 7.55 million people would be required to overturn the mean odds reduction. Odds reduction was similar when the outcome was incident depression (k=83) (OR=0.69 (95% CI: 0.66 to 0.77)) or an increase in subclinical depressive symptoms (k=21) (OR=0.65 (95% CI: 0.51 to 0.84)) (F(1,102)=0.73; p=0.3943).

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

Table 2 Depression symptoms associated with physical activity: adjusted OR and heterogeneity

	Contrast	K (N)	OR	95% CI	Q	I^2 (95% CI)
Region						
North America		64 (32)	0.81****	0.75 to 0.86	910.6****	93.2 (92.6 to 93.7)
Europe		73 (37)	0.78****	0.73 to 0.83	306.6****	88.6 (87.0 to 90.0)
Asia		29 (16)	0.73****	0.66 to 0.81	109.9****	75.4 (70.5 to 79.5)
Australia		13 (6)	0.83*	0.72 to 0.96	30.3**	63.7 (50.5 to 73.4)
Gender						
Females only	1	50 (23)	0.78****	0.71 to 0.85	309.5****	84.5 (82.5 to 86.3)
Males only	−1	23 (13)	0.80****	0.70 to 0.92	352.0****	94.0 (93.2 to 94.8)
Mixed (54% female, 8%–94%, IQR=46%–57%)	0	106 (65)	0.79****	0.75 to 0.82	429.1****	75.8 (73.4 to 77.9)
Age (years)						
18–44	1	58 (20)	0.81****	0.76 to 0.87	241.8****	76.8 (73.8 to 79.6)
45–64	2	55 (35)	0.77****	0.72 to 0.82	624.7****	91.5 (90.7 to 92.3)
65+	3	66 (37)	0.81**	0.70 to 0.93	328.7****	80.5 (78.2.82.6)
Adjustments						
Age, sex, demographics	1	20 (13)	0.81****	0.75 to 0.88	672.6****	97.3 (97.0 to 97.6)
Plus modifiable risks	2	42 (22)	0.80****	0.73 to 0.86	128.7****	68.9 (63.5 to 73.5)
Plus chronic disease or disability	3	117 (56)	0.77****	0.73 to 0.82	568.6****	79.8 (78.0 to 81.4)
Study quality						
Low	1	28 (16)	0.80**	0.71 to 0.91	49.6**	47.6 (34.4 to 58.1)
Acceptable	2	93 (50)	0.79****	0.75 to 0.83	1121.0****	91.9 (91.3 to 92.5)
Good	3	58 (26)	0.77****	0.72 to 0.82	216.5****	74.1 (70.6.77.3)
Study design						
Baseline exposure and incident cases	−0.5	74 (36)	0.80****	0.76 to 0.85	938.7****	92.3 (91.7 to 92.9)
Baseline exposure and change in symptoms	−0.5	55 (34)	0.81****	0.76 to 0.86	201.7****	73.7 (70.0 to 77.0)
Change in exposure and incident cases	0.5	18 (8)	0.64***	0.52 to 0.79	23.1	30.7 (7.2 to 48.3)
Change in exposure and change in symptoms	0.5	32 (18)	0.71**	0.55 to 0.91	231.6****	87.1 (85.0 to 88.8)
PA measure						
Frequency	−0.5	66 (35)	0.79****	0.73 to 0.84	239.3****	73.3 (69.8 to 76.3)
Time or distance	−0.5	21 (14)	0.78***	0.69 to 0.87	79.6****	76.1 (70.4 to 80.8)
Meeting guidelines (sufficient)	0.5	20 (7)	0.88****	0.86 to 0.91	82.4****	78.2 (72.9 to 82.4)
Metabolic equivalent	0.5	72 (36)	0.78****	0.73 to 0.83	1013.1****	93.1 (92.5 to 93.6)
PA domain						
Leisure time	−0.5	145 (71)	0.80****	0.77 to 0.83	687.8****	79.2 (77.6 to 80.7)
Total	0.5	34 (20)	0.72**	0.65 to 0.80	192.9****	30.9 (14.3 to 44.2)
Depression measure						
Symptoms	−1	51 (27)	0.78****	0.72 to 0.86	260.5****	81.2 (78.7 to 83.4)
Screening cut-point	0.33	111 (53)	0.79****	0.75 to 0.83	455.4****	76.1 (73.8 to 78.1)
Self-reported diagnosis	0.33	6 (4)	0.77****	0.69 to 0.85	76.2****	94.8 (92.9 to 96.1)
Clinical diagnosis	0.33	11 (7)	0.79****	0.73 to 0.86	9.1	1.1 (00.0 to 8.3)
Exposure dose						
Binary or correlation	0	93 (64)	0.80****	0.76 to 0.84	459.1****	80.2 (78.2 to 82.0)
Low	−1	33 (26)	0.84****	0.78 to 0.90	170.1****	81.8 (78.7 to 84.4)
Moderate	0.5	31 (24)	0.72****	0.67 to 0.78	142.7****	79.7 (76.0 to 82.8)
High	0.5	22 (11)	0.76****	0.69 to 0.83	67.9****	70.5 (63.2 to 76.4)
Recommended exposure						
Not meeting guidelines	−1	35 (21)	0.82****	0.73 to 0.92	129.0****	74.4 (69.8 to 78.4)
Not reported	0	105 (59)	0.78****	0.73 to 0.82	499.6****	79.4 (77.5 to 81.2)
Meeting guidelines	1	39 (23)	0.77****	0.73 to 0.81	179.0****	79.3 (76.0 to 82.2)

K is number of effects and N is number of studies for each category or level of each modifier.
Significant at *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.
IQR, interquartile range; PA, physical activity.

The meta-regression model ($F(9,94)=10.8$; $p<0.0001$, $R^2=0.11$ (95% CI: 0.09 to 0.13); $Q_E(94)=1010.6$, $p<0.0001$) indicated that crude odds of depressive symptoms were negatively related to study design (ie, physical activity exposure only at baseline vs change in physical activity) (beta=−0.17; p=0.0053), exposure dose (beta=−0.019; p<0.0001) and physical activity measure (ie, frequency or time vs volume) (beta=−0.18; p<0.0001) (see table 3). Specifically, incident depression was negatively related to exposure dose (beta=−0.018; p<0.0001) and physical activity measure (beta=−0.17; p=0.0003), and an increase in subclinical symptoms was negatively related to study design (beta=−0.22; p=0.0051). The model was materially

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

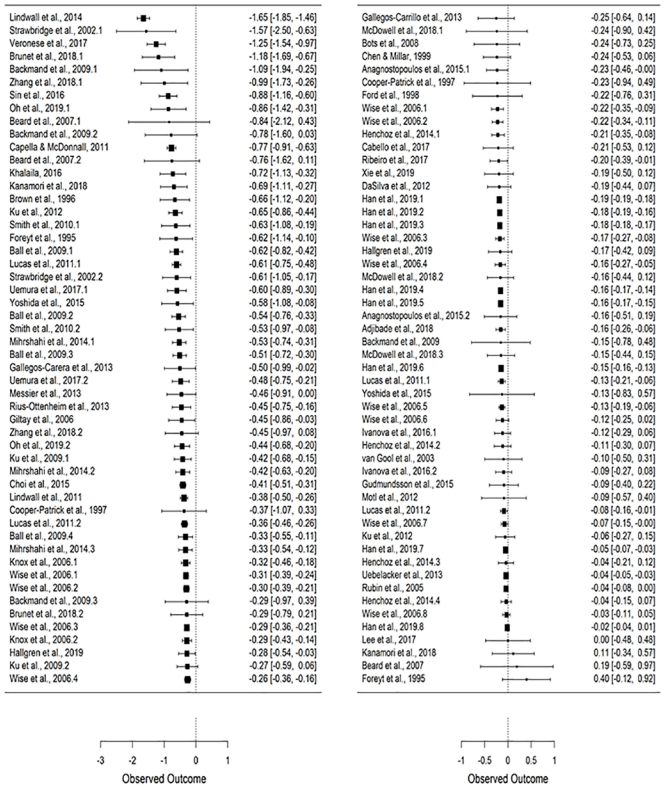


Figure 1 Forest plot of crude effects.

the same when exposure was defined as meeting recommended guidelines of moderate or vigorous physical activity and when using Bayes estimation.

The odds of depression were lower when: (1) exposure dose was moderate or high (OR=0.68 (95% CI: 0.60 to 0.76), k=36) compared with low (OR=0.76 (95% CI: 0.70 to 0.87), k=20) or when participants met recommended physical activity exposure (OR=0.76 (95% CI: 0.69 to 0.83), k=37) compared with not having met the recommendations (OR=0.68 (95% CI: 0.52 to 0.84), k=15); (2) participants had increased their physical activity exposure during a study (OR=0.63 (95% CI: 0.48 to 0.82), k=20) compared with when participants reported a single exposure at baseline (OR=0.72 (95% CI: 0.66 to 0.78), k=84).

The funnel plot (online supplemental figure 2) did not indicate marked asymmetry. Egger's test was significant (p<0.001). The rank correlation test of bias was not significant (p=0.751).

Adjusted odds

Physical activity was associated with 22% lower adjusted odds of study participants having incident depression or an increase in subclinical symptoms (OR=0.79 (95% CI: 0.75 to 0.82); Q(178)=1424.4, p<0.0001; I^2=87.6 (95% CI=86.8 to 88.3); k=179, n=91). ORs were correlated within studies (ICC=0.27 (95% CI: 0.12 to 0.41)) and cohorts (ICC=0.14 (95% CI: −0.020 to 0.30)) but not regions (ICC=0.00). Fail-safe N_1 was 848, estimating that a null outcome from a cohort of 3.99 million people would be required to overturn the mean odds reduction. Odds reduction was the same when the outcome was incident depression (k=128) (OR=0.79 (95% CI: 0.76 to 0.83)) or an increase in subclinical symptoms (k=51) (OR=0.78 (95% CI: 0.72 to 0.86)) (F(1,177)=0.015; p=0.9021).

The meta-regression model (F(10,168)=14.5; p<0.0001, R^2=0.09 (95% CI: 0.07 to 0.10), Q_E(168)=599.9, p<0.0001) indicated that adjusted odds of depression were negatively related to physical activity dose (beta=−0.075; p<0.0001) and study design (beta=−0.16; p=0.0002) (see table 3). Specifically, incident depression was negatively related to physical activity dose (beta=−0.072; p<0.0001), and an increase in subclinical symptoms was negatively related to physical activity dose (beta=−0.18; p=0.0001) and study design (beta=−0.37; p<0.0001). The model was materially the same when exposure was defined as meeting recommended guidelines of moderate or vigorous physical activity or when Bayes estimation was used.

The odds of depression were lower when: (1) exposure dose was moderate or high (OR=0.73 (95% CI: 0.68 to 0.78)) compared with low (k=32, OR=0.84 (95% CI: 0.78 to 0.90), k=54) or when participants met recommended exposure (OR=0.77 (95% CI: 0.73 to 0.81), k=39) compared with not having met the recommendation (OR=0.80 (95% CI: 0.74 to 0.87), k=35); and (2) participants had increased their physical activity exposure in a study (OR=0.69 (95% CI: 0.61 to 0.79), k=50) compared with when participants reported a single exposure at baseline (OR=0.81 (95% CI: 0.78 to 0.84), k=129).

Egger's test was significant (p<0.001), but the funnel plot (online supplemental figure 3) did not indicate asymmetry, excepting a few studies that had the smallest sample sizes (N<1000) and reported mainly reduced odds. The rank correlation test of bias was non-significant (p=0.310).

Sensitivity analysis

Dose–response by measures of depression and physical activity

We applied the meta-regression model (including other confounders) to each depression measure to clarify that physical activity exposure was inversely related to adjusted odds of depression, whether measured by incident depression or by an increase in subclinical symptoms, and according to physical activity measure or change. Physical activity was inversely related to odds of incident depression measured by diagnosed cases (k=17; beta=−0.081; p<0.0001) or by screening cut-points (k=111; beta=−0.062; p<0.0001) and to odds of an increase in subclinical symptoms (k=51; beta=−0.183, p=0.0001). All results were independent of type of physical activity measure (p≥0.3561). Physical activity measured by frequency or time was inversely related to odds of incident depression measured by screening cut-points (k=53; beta=−0.07; p<0.0001) or an increase in subclinical symptoms (k=25; beta=−0.16, p=0.0282) but not to diagnosed cases (k=9; beta=0.02, p=0.8645). Physical activity measured by volume was inversely related to odds of incident depression measured by diagnosed cases (k=8; beta=−0.08; p=0.0742), screening cut-points (k=58; beta=−0.06; p=0.0002) and to odds of an increase in subclinical symptoms (k=26; beta=−0.18, p=0.0072). Change in physical activity was inversely related to odds of incident depression measured be screening cut-points (k=30; beta=−0.07; p<0.0001) and an increase in subclinical symptoms (k=20; beta=−0.16; p<0.0721). No studies of diagnosed depression reported change in physical activity.

Comparison with prior review

We restricted analysis to 35 studies of adults and odds of incident depression after physical activity exposure that were also included in the meta-analysis by Schuch *et al*.[20] We found similar average adjusted (OR=0.78 (95% CI: 0.73 to 0.83); k=69, n=31; Q(68)=200.8, p<0.0001, I^2=66.6 (95% CI: 62.1 to 70.6)) and crude odds (OR=0.65 (95% CI: 0.57 to 0.74); k=39, n=18; Q(38)=172.5, p<0.0001, I^2=78.6 (95% CI: 75.1 to 81.6)), but our results properly weighted

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

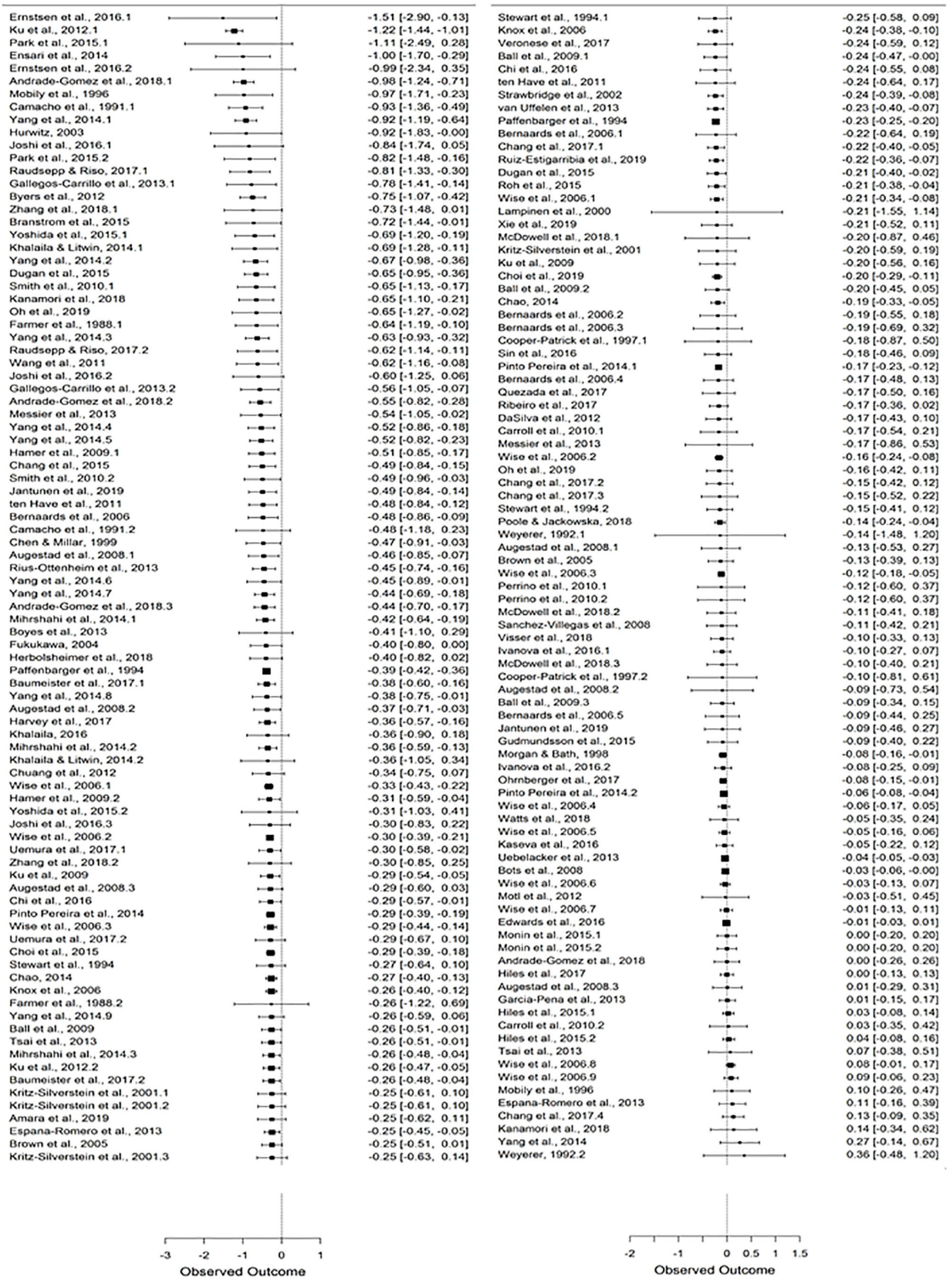


Figure 2 Forest plot of adjusted effects.

here by the inverse variance were heterogeneous not homogeneous. Results of the meta-regression model were similar to the analysis of all studies reported here but not tested in that review.[20] The meta-regression model ($F(9,59)=2.53$; $p=0.0158$, $R^2=0.08$ (95% CI: 0.05 to 0.12), $Q_E(59)=159.1$, $p<0.0001$) indicated that adjusted odds of depressive symptoms were negatively related to physical activity dose ($p=0.0002$), whether exposure met guidelines for sufficient

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

activity (p<0.0001) and study design (p=0.0393). Results were materially the same using Bayes estimation.

Cohort bias

Fifty-nine adjusted ORs were retrieved from 27 studies reported from 12 cohorts. Cohort accounted for 5% of variation in adjusted effects (ICC=0.05 (95% CI: −0.09 to 0.19)) while study accounted for 33% of variation in adjusted effects (ICC=0.33 (95% CI: 0.12 to 0.53)). Results were similar to those obtained for all studies (OR=0.79 (95% CI: 0.76 to 0.82)). Odds were equally low when the outcome was incident depression (k=40) (OR=0.79 (95% CI: 0.76 to 0.83)) or an increase in subclinical depressive symptoms (k=19) (OR=0.82 (95% CI: 0.71 to 0.95) (F(1,57)=1.3; p=0.2526). The meta-regression model indicated that adjusted odds of depression were negatively related to physical activity dose (p=0.0009), whether exposure met guidelines for sufficient activity (p=0.0350), and increased physical activity exposure (p=0.0017). Results were similar to those for 151 effects from 74 studies representing a single cohort (OR=0.78 (95% CI: 0.74 to 0.81)). Similarly, the meta-regression model indicated that odds of depression were negatively related to physical activity dose, whether exposure met guidelines for sufficient activity, and increased physical activity exposure (p≤0.0003).

DISCUSSION

This systematic review of 111 studies of more than three million people found that lower odds of depression outcomes were associated with: (1) moderate-to-high exposure doses or exposure that met public health guidelines for physical activity; and (2) increases in physical activity exposure across time. These findings are novel and strengthen the case that there is a temporal association of physical activity with primary prevention of depression.

The association between physical activity and lower odds of depression was seen for incident cases and also for lowered subclinical symptoms of depression—the latter is a new finding because that outcome was not considered in prior reviews.[19 20] Because depression is a continuum from normal to pathological function,[66] a favourable association of physical activity with change in subclinical depressive symptoms might provide clinical protection in people at risk for incident depression.

Exposure dose

Our data suggest that any level of physical activity mitigates depression risk, but relatively moderate and higher physical activity levels are more strongly associated with lower risk. Only 10 studies included three physical activity levels needed to directly test whether depression odds differed between high and moderate physical activity. We highlight that various measures and criteria were used to classify people into activity exposure groups—these were not identical across studies. It was not possible to convert the findings to a standard estimate of physical activity volume at each level (eg, MET-hours), as has been done in hazard studies not included here.[20 67 68] Among the nearly 40% of studies that provided enough information to determine whether active people were meeting existing WHO and US Federal public health recommendations for sufficient physical activity,[20 28] odds of depression were lower when people met or exceeded the recommended exposure. Randomised trials have provided limited evidence of dose-dependent effects of exercise among depressed patients,[69 70] and trials showing no effect of exercise training may have had an inadequate exposure dose.[71] Taken together, we conclude that there is a dose response, but studies comparing three or more standardised amounts of physical activity are needed to clarify a more precise dose–response association.

Table 3 Meta-regression of effect modifiers

	Estimate (LOR)	SE	t value	P value	95% CI
Crude model 51 studies (k=104)					
Main effects					
Intercept	−0.35	0.05	−6.8	<0.0001	−0.46 to −0.25
Female (%)	0.06	0.18	0.34	0.7335	−0.30 to 0.43
Age	−0.06	0.07	−0.89	0.3733	−0.19 to 0.07
Design	−0.17	0.06	−2.82	0.0059	−0.28 to −0.05
Follow-up	0.004	0.008	0.53	0.5965	−0.012 to 0.02
Exposure dose	−0.019	0.002	−8.6	<0.0001	−0.02 to −0.14
Depression measure	0.07	0.12	0.57	0.5707	−0.17 to 0.31
Physical activity measure	−0.18	0.04	−4.2	<0.0001	−0.27 to −0.10
Physical activity domain	−0.09	0.13	−0.70	0.4876	−0.34 to 0.16
Study quality	0.006	0.07	0.82	0.9352	−0.14 to 0.15
Adjusted model 91 studies (k=179)					
Intercept	−0.26	0.022	−11.9	<0.0001	−0.31 to −0.22
Female (%)	−0.011	0.04	−0.26	0.7986	−0.10 to 0.08
Age	0.009	0.026	0.36	0.7212	−0.04 to 0.06
Design	−0.17	0.05	−3.8	0.0002	−0.26 to −0.08
Follow-up	−0.002	0.004	−0.60	0.5552	−0.01 to 0.005
Exposure dose	−0.08	0.007	−11.2	<0.0001	−0.09 to −0.06
Depression measure	−0.05	0.05	−0.99	0.3250	−0.14 to 0.05
Physical activity measure	0.017	0.05	0.35	0.7291	−0.08 to 0.11
Physical activity domain	−0.09	0.06	−1.57	0.1192	−0.20 to 0.024
Study quality	−0.032	0.036	−0.88	0.3805	−0.10 to 0.040
Adjustments	−0.012	0.032	−0.38	0.7062	−0.07 to 0.05

K, number of effects; LOR, log odds ratio.

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

Outcome measures

Our study conclusions—that physical activity/exercise is beneficial—apply whether the depression outcome was incident depression or an increase in subclinical symptoms. About half the studies (n=58) defined depression outcomes based on scores above varying predictive cut-points for probable depression on validated screening tools such as the Centre for Epidemiological Studies Depression scale and the Hospital Anxiety and Depression scale. In another 35 studies, the primary outcomes were an increase in depressive symptoms of unspecified clinical meaning that cannot be categorised precisely. Eleven cohorts used a standardised diagnostic interview and another seven used self-reported physician diagnosis to measure incident depression. Despite high-quality outcome measures,[72] nine of those studies used weak measures of physical activity exposure (eg, single items or estimates other than volume). Importantly, the association of physical activity with lower odds of depression was not limited to self-rated symptoms.

Temporal sequence

Could our results be affected by reverse causality? The prospective cohorts measured physical activity exposure prior to depression onset. However, we also scrutinised studies that assessed change in physical activity exposure over time and prior to incident cases of depression or a change in subclinical depressive symptoms. On average, those studies of change in physical activity reported lower odds of depression than studies of a single physical activity exposure at baseline. While not fully ruling out reverse causation (ie, elevated depression leads to lower physical activity), this consolidation of studies that measured increased physical activity prior to depression outcomes is an advancement towards characterising the potential bias of reverse causation. In addition, we excluded from our review results of studies that defined incident cases by binary screening test cut-points if baseline depressive symptoms were not controlled or adjusted; outcome estimates would be biased if physically active people had lower symptoms than inactive people at baseline.

A 2018 meta-analysis of 35 cohort studies of adults[20] reported significant reductions in crude and adjusted rate ratios for incident depression. Schuch and colleagues concluded that the results were mainly homogeneous across studies. Hence, the authors' attempts to examine modification of the exposure–outcome association by cohort size or gender composition, length of follow-up and study quality were unsuccessful. Homogeneity in that analysis likely resulted from weighting of each OR by cohort size, which is not the sampling error of an OR.[22] We also note that Schuch and colleagues limited their analysis to the highest physical activity exposure group, which restricts conclusions about whether outcomes depended on physical activity dose. Their review also was limited to studies of a single, baseline exposure to physical activity. An earlier systematic review of physical activity and incident depression risk by Mammen and Faulkner[18] located 10 adult studies of change in physical activity and 13 of dose response, but meta-analysis was not used to aggregate and compare the mixed findings of those studies.

Exposure measures

Studies were mainly limited to self-reported physical activity. Only four studies[32 73–75] used a device (ie, a pedometer) to measure physical activity. In addition to devices, cardiorespiratory fitness provides an objective, surrogate measure of change in physical activity exposure that has been underused in studies of depression risk.[76–80] In cohorts of men and women assessed repeatedly at four medical clinic visits across 10–12 years, maintenance of fitness compared with lessened fitness reduced odds of incident depression, even after time-varying adjustments for relevant covariates.[81]

Change in exposure

There is risk of misclassification bias in cohort studies that only assessed a single physical activity exposure at baseline, and there are very few population-based studies that track trajectories of change and periodicity of physical activity across multiple timepoints.[82] A quarter of the studies reviewed here (12/51 studies that reported crude odds; 22/91 studies that reported adjusted odds) assessed physical activity more than once. They allowed us to estimate change in exposure across follow-up, which likely provides a truer estimate of customary physical activity than a single exposure measured at baseline.

Risk of bias

There currently is no consensus on the best procedure/tool to assess risk of bias in observational designs.[83 84] The method we used[42] focuses on common domains of bias: selection and attrition, confounding, and exposure and outcome measures, consistent with criteria endorsed by the GRADE working group.[43] Here, it yielded high rater agreement. Although few studies fully accounted for participants lost to follow-up, we judged the quality of the studies that most fully adjusted for putative confounders as moderate-to-high in quality when they used a valid physical activity measure. Acceptable measures of physical activity were those based on validated questions that permit an estimate of the quantity of physical activity expressed as volume (ie, duration, frequency and intensity), metabolic equivalence of volume or as meeting a recommended criterion of volume. Weak measures of physical activity were limited to single-item indicators of only frequency, distance or time that are not equivalent to volume of exposure. We have confidence in the summary evidence from about 40 studies that adjusted for confounders and used a validated measure of exposure, although derived from self-reported rather than device-measured physical activity. Only 10 studies assessed change in physical activity using a measure of volume and also adjusted for confounders.[35 39 40 85–91]

Half the included studies reported odds that were fully adjusted for many potential confounders common to risks of physical inactivity and chronic diseases. There was possible publication bias of the fully adjusted odds. However, funnel plot asymmetry and Egger's test can yield false-positive indication of bias when the true effect is heterogeneous,[92] as observed here. Statistical tests of funnel asymmetry have uncertain validity as an indicator of publication bias when effects are heterogeneous and are not recommended as a sole indicator.[93] Although the outcome estimate in each analysis had sufficient precision and consistency and low publication bias, heterogeneity of the cumulative evidence, even after adjusting for variation within studies, suggests residual confounding in the studies or effect modification not accounted for by the analysis here. This might be because many cohort studies of physical activity and depression were not specifically designed to examine depression; rather a measure of depressive symptoms had been included in studies of broader health. Hence, there likely was residual confounding and selection bias even in the most fully adjusted studies reviewed here. For example, no reviewed study used propensity matching of exposure groups on risk factors, as is common in retrospective or

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

case–control designs[94] and randomised trials when risks of a medical or health outcome is well established. Although modifiable risk factors of depression are incompletely understood, future cohort studies should include widowhood, physical abuse during childhood, obesity, metabolic risk factors, sexual dysfunction and job strain as confounders or effect modifiers.[95]

CONCLUSIONS

The likelihood of residual confounding and selection bias notwithstanding, studies reporting adjusted odds collectively were of moderate (ie, acceptable to good) quality. The cumulative evidence supports that moderate-to-vigorous physical activity is inversely associated with odds both of incident depression and of increased subclinical depressive symptoms among adults, regardless of global region, gender, age, follow-up period, and timing or measures of physical activity and depression. These observational findings are sufficiently positive to encourage randomised trials that experimentally test the efficacy and effectiveness of physical activity intervention in the primary prevention of depression.

Key messages

What is already known

- Exercise therapy in the management of depression can improve symptoms.
- Whether exercise confers protection against the onset of depression is uncertain.
- Whether odds of depression vary according to physical activity dose or change in physical activity is not fully understood.

What are the new findings

- Physical activity is inversely associated with odds of incident depression.
- Physical activity is associated with lower odds of having more subclinical depressive symptoms.
- Odds reduction depended on amount of physical activity. Moderate-to-vigorous physical activity was associated with lower odds more than light physical activity.
- Odds of depression were lower in those studies where researchers reported an increase in physical activity than in those studies where physical activity was measured only at baseline.

Correction notice This article has been corrected since it published Online First. The first affiliation and tables 1 and 2 have been updated.

Twitter Matthew Payton Herring @mph8

Contributors All authors had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis. All authors contributed to study concept and design; acquisition, analysis or interpretation of data; critical revision of the manuscript and administrative, technical or material support. RKD drafted the manuscript and contributed to statistical analysis. RKD and MPH supervised the study.

Funding CPMcD is funded by the Irish Research Council under the Government of Ireland Postdoctoral Programme.

Disclaimer No funding was used in the design, collection, management, analysis, interpretation of the data, preparation, review, approval of the manuscript, and decision to submit the manuscript for publication.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Cillian P McDowell http://orcid.org/0000-0001-7825-1123
Matthew Payton Herring http://orcid.org/0000-0002-6835-5321

REFERENCES

1 World Health Organization. *Global health estimates*. Geneva: World Health Organization, 2017.
2 Chisholm D, Sweeny K, Sheehan P, *et al*. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry* 2016;3:415–24.
3 Correll CU, Solmi M, Veronese N, *et al*. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry* 2017;16:163–80.
4 Ferrari AJ, Somerville AJ, Baxter AJ, *et al*. Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature. *Psychol Med* 2013;43:471–81.
5 Freeman MP, Fava M, Lake J. Complementary and alternative medicine in major depressive disorder: the American psychiatric association Task force report. *J Clin Psychiatry* 2010;71:669–81.
6 Trivedi MH, Greer TL, Church TS, *et al*. Exercise as an augmentation treatment for nonremitted major depressive disorder: a randomized, parallel dose comparison. *J Clin Psychiatry* 2011;72:677–84.
7 Trivedi MH, Greer TL, Grannemann BD, *et al*. Exercise as an augmentation strategy for treatment of major depression. *J Psychiatr Pract* 2006;12:205–13.
8 Gartlehner G, Gaynes BN, Amick HR. *Nonpharmacological Versus Pharmacological Treatments for Adult Patients With Major Depressive Disorder. Comparative Effectiveness Review No. 161. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No.290-2012-00008-I.) AHRQ Publication No. 15(16)-EHC031-EF*. Rockville, MD: Agency for Healthcare Research and Quality, 2015. http://www.effectivehealthcare.ahrq.gov/reports/final.cfm
9 Burton R. *The anatomy of melancholy*. Printed by Ion Lichfield for Henry Cripps: Oxford, 1632.
10 Cooney G, Dwan K, Mead G. Exercise for depression. *JAMA* 2014;311:2432–3.
11 Krogh J, Nordentoft M, Sterne JAC, *et al*. The effect of exercise in clinically depressed adults: systematic review and meta-analysis of randomized controlled trials. *J Clin Psychiatry* 2011;72:529–38.
12 Mead GE, Morley W, Campbell P. Exercise for depression. *Cochrane Database Syst Rev* 2009;3:CD004366.
13 Herring MP, Puetz TW, O'Connor PJ, *et al*. Effect of exercise training on depressive symptoms among patients with a chronic illness: a systematic review and meta-analysis of randomized controlled trials. *Arch Intern Med* 2012;172:101–11.
14 Krogh J, Hjorthøj C, Speyer H, *et al*. Exercise for patients with major depression: a systematic review with meta-analysis and trial sequential analysis. *BMJ Open* 2017;7:e014820.
15 Schuch FB, Vancampfort D, Richards J, *et al*. Exercise as a treatment for depression: a meta-analysis adjusting for publication bias. *J Psychiatr Res* 2016;77:42–51.
16 Physical Activity Guidelines Advisory Committee. *2018 physical activity guidelines Advisory Committee scientific report*. Washington, DC: U.S. Department of Health and Human Services, 2018.
17 Dunn AL, Trivedi MH, O'Neal HA. Physical activity dose-response effects on outcomes of depression and anxiety. *Med Sci Sports Exerc* 2001;33:S587–97.
18 Mammen G, Faulkner G. Physical activity and the prevention of depression: a systematic review of prospective studies. *Am J Prev Med* 2013;45:649–57.
19 Physical Activity Guidelines Advisory Committee. *Physical activity guidelines Advisory Committee report, 2008*. Washington, DC: US Department of Health and Human Services, 2008: A1–14.
20 Schuch FB, Vancampfort D, Firth J, *et al*. Physical activity and incident depression: a meta-analysis of prospective cohort studies. *Am J Psychiatry* 2018;175:631–48.
21 Dishman RK, Heath GW, Schmidt MD. *Physical activity epidemiology*. 3rd edn. Champagne, IL: Human Kinetics, 2021.
22 Lipsey MW, Wilson DB. *Practical meta-analysis*. Thousand Oaks, CA: Sage, 2001.
23 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9.
24 Stroup DF, Berlin JA, Morton SC, *et al*. Meta-Analysis of observational studies in epidemiology: a proposal for reporting. meta-analysis of observational studies in epidemiology (moose) group. *JAMA* 2000;283:2008–12.

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

25 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
26 Borenstein M, Hedges LV, Higgins JPT, *et al*. *Introduction to meta-analysis (PP. 45-49)*. Chichester, UK: John Wiley & Sons, 2011.
27 Wang F, DesMeules M, Luo W, *et al*. Leisure-time physical activity and marital status in relation to depression between men and women: a prospective study. *Health Psychol* 2011;30:204–11.
28 Bull FC, Al-Ansari SS, Biddle S, *et al*. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med* 2020;54:1451–62.
29 Augestad LB, Slettemoen RP, Flanders WD. Physical activity and depressive symptoms among Norwegian adults aged 20-50. *Public Health Nurs* 2008;25:536–45.
30 Carroll DD, Blanck HM, Serdula MK, *et al*. Obesity, physical activity, and depressive symptoms in a cohort of adults aged 51 to 61. *J Aging Health* 2010;22:384–98.
31 Farmer ME, Locke BZ, Mościcki EK, *et al*. Physical activity and depressive symptoms: the NHANES I epidemiologic follow-up study. *Am J Epidemiol* 1988;128:1340–51.
32 Hiles SA, Baker AL, de Malmanche T, *et al*. Unhealthy lifestyle may increase later depression via inflammation in older women but not men. *J Psychiatr Res* 2015;63:65–74.
33 Kritz-Silverstein D, Barrett-Connor E, Corbeau C. Cross-sectional and Prospective Study of Exercise and Depressed Mood in the Elderly : The Rancho Bernardo Study. *Am J Epidemiol* 2001;153:596–603.
34 McDowell CP, Dishman RK, Hallgren M, *et al*. Associations of physical activity and depression: results from the Irish longitudinal study on ageing. *Exp Gerontol* 2018;112:68–75.
35 Monin JK, Levy B, Chen B, *et al*. Husbands' and wives' physical activity and depressive symptoms: longitudinal findings from the cardiovascular health study. *Ann Behav Med* 2015;49:704–14.
36 Weyerer S. Physical inactivity and depression in the community. *Int J Sports Med* 1992;13:492–6.
37 Yang X, Hirvensalo M, Hintsanen M, *et al*. Longitudinal associations between changes in physical activity and depressive symptoms in adulthood: the young Finns study. *Int J Behav Med* 2014;21:908–17.
38 Andrade-Gómez E, Martínez-Gómez D, Rodríguez-Artalejo F, *et al*. Sedentary behaviors, physical activity, and changes in depression and psychological distress symptoms in older adults. *Depress Anxiety* 2018;35:884–97.
39 Jantunen H, Wasenius N, Salonen MK, *et al*. Change in physical activity and health-related quality of life in old age-A 10-year follow-up study. *Scand J Med Sci Sports* 2019;29:1797–804.
40 Knox S, Barnes A, Kiefe C, *et al*. History of depression, race, and cardiovascular risk in cardia. *Int J Behav Med* 2006;13:44–50.
41 Center for Behavioral Health Statistics and Quality. *Key substance use and mental health indicators in the United States: results from the 2015 national survey on drug use and health. HHS publication No. SMA 16-4984, NSDUH series H-51*. Rockville, MD: Substance Abuse and Mental Health Services, 2016. https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR1-2015/NSDUH-FFR1-2015/NSDUH-FFR1-2015.pdf
42 Jarde A, Losilla J-M, Vives J, *et al*. Q-Coh: a tool to screen the methodological quality of cohort studies in systematic reviews and meta-analyses. *Int J Clin Heal Psychol* 2013;13:138–46.
43 Guyatt G, Oxman AD, Akl EA, *et al*. Grade guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383394.
44 Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psychol Bull* 1995;117:167–78.
45 Peterson RA, Brown SP. On the use of beta coefficients in meta-analysis. *J Appl Psychol* 2005;90:175–81.
46 Carvalho AF, Maes M, Solmi M, *et al*. Is dynapenia associated with the onset and persistence of depressive and anxiety symptoms among older adults? Findings from the Irish longitudinal study on ageing. *Aging Ment Health* 2019;63:1–8.
47 Frank P, Kaushal A, Poole L, *et al*. Systemic low-grade inflammation and subsequent depressive symptoms: is there a mediating role of physical activity? *Brain Behav Immun* 2019;80:688–96.
48 Cole TJ. Setting number of decimal places for reporting risk ratios: rule of four. *BMJ* 2015;350:h1845.
49 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
50 Konstantopoulos S. Fixed effects and variance components estimation in three-level meta-analysis. *Res Synth Methods* 2011;2:61–76.
51 Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J Educ Behav Stat* 2005;30:261–93.
52 Viechtbauer W. Conducting Meta-Analyses in *R* with the **metafo***r* Package. *J Stat Softw* 2010;36:1–48.
53 Viechtbauer W. Meta-Analysis package for R, 2018. Available: https://cran.rproject.org/web/packages/metafor/metafor.pdf [Accessed 27 Jun 2020].
54 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018. http://www.R-project.org/
55 Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med* 2003;22:2693–710.
56 Viechtbauer W, López-López JA, Sánchez-Meca J, *et al*. A comparison of procedures to test for moderators in mixed-effects meta-regression models. *Psychol Methods* 2015;20:360–74.
57 Jackson D, Law M, Stijnen T, *et al*. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio. *Stat Med* 2018;37:1059–85.
58 Muthén LK, Muthén BO. *Mplus User's Guide*. Eighth Edition. Los Angeles: CA, 1998-2017.
59 Cheung MW-L. A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychol Methods* 2008;13:182–202.
60 Hox JJ. *Multilevel analysis: techniques and applications*. New York, NY: Routledge, 2010.
61 Maas CJM, Hox JJ. Sufficient sample sizes for multilevel modeling. *Methodology* 2005;1:86–92.
62 Higgins JPT, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
63 Egger M, Davey Smith G, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
64 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.
65 Rosenberg MS. The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* 2005;59:464–8.
66 Regier DA, Kuhl EA, Kupfer DJ. The DSM-5: classification and criteria changes. *World Psychiatry* 2013;12:92–8.
67 Kim S-Y, Park J-H, Lee MY, *et al*. Physical activity and the prevention of depression: a cohort study. *Gen Hosp Psychiatry* 2019;60:90–7.
68 Kuwahara K, Honda T, Nakagawa T, *et al*. Intensity of leisure-time exercise and risk of depressive symptoms among Japanese workers: a cohort study. *J Epidemiol* 2018;28:94–8.
69 Dunn AL, Trivedi MH, Kampert JB, *et al*. Exercise treatment for depression: efficacy and dose response. *Am J Prev Med* 2005;28:1–8.
70 Singh NA, Stavrinos TM, Scarbek Y, *et al*. A randomized controlled trial of high versus low intensity weight training versus general practitioner care for clinical depression in older adults. *J Gerontol A Biol Sci Med Sci* 2005;60:768–76.
71 Krogh J, Videbech P, Thomsen C, *et al*. DEMO-II trial. aerobic exercise versus stretching exercise in patients with major depression-a randomised clinical trial. *PLoS One* 2012;7:e48316.
72 Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009;374:609–19.
73 Brunet J, O'Loughlin JL, Gunnell KE, *et al*. Physical activity and depressive symptoms after breast cancer: cross-sectional and longitudinal relationships. *Health Psychol* 2018;37:14–23.
74 Fukukawa Y, Nakashima C, Tsuboi S, *et al*. Age differences in the effect of physical activity on depressive symptoms. *Psychol Aging* 2004;19:346–51.
75 Raudsepp L, Riso E-M. Longitudinal association between objectively measured walking and depressive symptoms among Estonian older adults. *J Aging Phys Act* 2017;25:639–45.
76 Åberg MAI, Waern M, Nyberg J, *et al*. Cardiovascular fitness in males at age 18 and risk of serious depression in adulthood: Swedish prospective population-based study. *Br J Psychiatry* 2012;201:352–9.
77 Baumeister SE, Leitzmann MF, Bahls M, *et al*. Associations of leisure-time and occupational physical activity and cardiorespiratory fitness with incident and recurrent major depressive disorder, depressive symptoms, and incident anxiety in a general population. *J Clin Psychiatry* 2017;78:e41–7.
78 Kandola AA, Osborn DPJ, Stubbs B, *et al*. Individual and combined associations between cardiorespiratory fitness and grip strength with common mental disorders: a prospective cohort study in the UK Biobank. *BMC Med* 2020;18:1.
79 Sui X, Laditka JN, Church TS, *et al*. Prospective study of cardiorespiratory fitness and depressive symptoms in women and men. *J Psychiatr Res* 2009;43:546–52.
80 Willis BL, Leonard D, Barlow CE, *et al*. Association of midlife cardiorespiratory fitness with incident depression and cardiovascular death after depression in later life. *JAMA Psychiatry* 2018;75:911–7.
81 Dishman RK, Sui X, Church TS, *et al*. Decline in cardiorespiratory fitness and odds of incident depression. *Am J Prev Med* 2012;43:361–8.
82 Dishman RK, Vandenberg RJ, Motl RW, *et al*. Using constructs of the transtheoretical model to predict classes of change in regular physical activity: a multi-ethnic longitudinal cohort study. *Ann Behav Med* 2010;40:150–63.
83 Hartling L, Milne A, Hamm MP, *et al*. Testing the Newcastle Ottawa scale showed low reliability between individual reviewers. *J Clin Epidemiol* 2013;66:982–93.
84 Losilla J-M, Oliveras I, Marin-Garcia JA, *et al*. Three risk of bias tools lead to opposite conclusions in observational research synthesis. *J Clin Epidemiol* 2018;101:61–72.
85 Bäckmand H, Kaprio J, Kujala U, *et al*. Influence of physical activity on depression and anxiety of former elite athletes. *Int J Sports Med* 2003;24:609–19.
86 Brown WJ, Ford JH, Burton NW, *et al*. Prospective study of physical activity and depressive symptoms in middle-aged women. *Am J Prev Med* 2005;29:265–72.
87 Lindwall M, Larsman P, Hagger MS. The reciprocal relationship between physical activity and depression in older European adults: a prospective cross-lagged panel design using share data. *Health Psychol* 2011;30:453–62.

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.

88 Gallegos-Carrillo K, Flores YN, Denova-Gutiérrez E, *et al*. Physical activity and reduced risk of depression: results of a longitudinal study of Mexican adults. *Health Psychol* 2013;32:609–15.
89 Rius-Ottenheim N, Geleijnse JM, Kromhout D, *et al*. Physical activity after myocardial infarction: is it related to mental health? *Eur J Prev Cardiol* 2013;20:399–408.
90 Visser M, Wijnhoven HAH, Comijs HC, *et al*. A healthy lifestyle in old age and prospective change in four domains of functioning. *J Aging Health* 2019;31:1297–314.
91 Watts AS, Mortby ME, Burns JM. Depressive symptoms as a barrier to engagement in physical activity in older adults with and without Alzheimer's disease. *PLoS One* 2018;13:e0208581.
92 Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;53:1119–29.
93 Sterne JAC, Sutton AJ, Ioannidis JPA, *et al*. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
94 Hong I, Aaron SE, Li C-Y, *et al*. Physical activity and the risk of depression in community-dwelling Korean adults with a history of stroke. *Phys Ther* 2017;97:105–13.
95 Köhler CA, Evangelou E, Stubbs B, *et al*. Mapping risk factors for depression across the lifespan: an umbrella review of evidence from meta-analyses and Mendelian randomization studies. *J Psychiatr Res* 2018;103:189–207.

Br J Sports Med: first published as 10.1136/bjsports-2020-103140 on 5 January 2021. Downloaded from http://bjsm.bmj.com/ on January 14, 2021 by guest. Protected by copyright.