Effectiveness of wearable activity trackers to increase physical activity and improve health: a systematic review of systematic reviews and meta-analyses

Ty Ferguson, Timothy Olds, Rachel Curtis, Henry Blake, Alyson J Crozier, Kylie Dankiw, Dorothea Dumuid, Daiki Kasai, Edward O'Connor, Rosa Virgara, Carol Maher

Wearable activity trackers offer an appealing, low-cost tool to address physical inactivity. This systematic review of systematic reviews and meta-analyses (umbrella review) aimed to examine the effectiveness of activity trackers for improving physical activity and related physiological and psychosocial outcomes in clinical and non-clinical populations. Seven databases (Embase, MEDLINE, Ovid Emcare, Scopus, SPORTDiscus, the Cochrane Library, and Web of Science) were searched from database inception to April 8, 2021. Systematic reviews of primary studies using activity trackers as interventions and reporting physical activity, physiological, or psychosocial outcomes were eligible for inclusion. In total, 39 systematic reviews and meta-analyses were identified, reporting results from 163 992 participants spanning all age groups, from both healthy and clinical populations. Taken together, the meta-analyses suggested activity trackers improved physical activity (standardised mean difference [SMD] 0·3–0·6), body composition (SMD 0·7–2·0), and fitness (SMD 0·3), equating to approximately 1800 extra steps per day, 40 min per day more walking, and reductions of approximately 1 kg in bodyweight. Effects for other physiological (blood pressure, cholesterol, and glycosylated haemoglobin) and psychosocial (quality of life and pain) outcomes were typically small and often non-significant. Activity trackers appear to be effective at increasing physical activity in a variety of age groups and clinical and non-clinical populations. The benefit is clinically important and is sustained over time. Based on the studies evaluated, there is sufficient evidence to recommend the use of activity trackers.

Lancet Digit Health 2022; 4: e615–26

Alliance for Research in Exercise, Nutrition, and Activity, University of South Australia, Adelaide, SA, Australia
(T Ferguson BPhysio Hons, Prof T Olds PhD, R Curtis PhD, H Blake BHlthSc Hons, A J Crozier PhD, K Dankiw MRes, D Dumuid PhD, D Kasai MRsch, E O'Connor BPsych Hons, R Virgara PhD, Prof C Maher PhD)

Correspondence to: Prof Carol Maher, Alliance for Research in Exercise, Nutrition, and Activity, University of South Australia, Adelaide 5000, SA, Australia
carol.maher@unisa.edu.au

Introduction

Insufficient physical activity is a major threat to population health. Worldwide, physically inactive lifestyles are common in modern life, partly caused by labour-saving devices and transportation advances, combined with the busy pace of life. Physical inactivity is a key contributor to premature mortality and morbidity, increasing the risk of major non-communicable diseases, including coronary heart disease, stroke, type 2 diabetes, cancers, and mental illness.[1] The estimated cost to health-care systems worldwide as a result of these morbidities was US\$53·8 billion in 2013.[2] Low-cost interventions to address physical inactivity in clinical and non-clinical populations that are suited to the demands of modern lifestyles are needed.

Wearable activity trackers (consumer devices that provide feedback to the wearer, such as fitness trackers, activity-tracking smartwatches, and pedometers) could meet this need. These devices are affordable (costing from \$30),[3] visually appealing and user-friendly,[4] and encourage lifestyle physical activity (ie, accrual of physical activity through everyday activities), reducing the barriers associated with more structured forms of physical activity.[5] Wearable activity trackers promote behaviour-change techniques, such as self-monitoring and goal setting,[6] and their use has been associated with increased physical activity.[7–9] Additionally, using wearable activity trackers has been shown to be associated with improved physiological outcomes, such as reduced BMI,[10,11] reduced blood pressure,[10] and increased aerobic capacity,[9] which might occur via increases in physical activity. Wearable activity trackers also have the potential to improve psychosocial outcomes, such as depression and anxiety, through increases in physical activity, as physical activity is shown to have antidepressant[12] and anxiolytic[13] effects.

The global wearable activity-tracker market has grown tremendously over the past decade. Between 2014 and 2020, the number of wearable activity trackers shipped worldwide increased by an estimated 1444%.[14] In 2020, approximately \$2·8 billion was spent on wearable activity trackers globally.[15] Accordingly, the body of research examining the use of wearable activity trackers for measuring and intervening on physical activity has expanded rapidly.

Despite their prima-facie promise, there is widespread scepticism about the effectiveness of wearable activity trackers within the scientific, medical, and general community.[16–19] Such articles cite scientific studies and experts to highlight concerns about the accuracy of wearable activity trackers,[20] fuelling of obsessive behaviours and eating disorders,[21] exacerbation of health inequities,[22] and being simply ineffective for changing physical activity.[23] Such studies, however, represent the tip of a now voluminous body of evidence regarding wearable activity trackers and physical activity.

Hierarchies of evidence rank the strength of health evidence.[24] At the pinnacle of contemporary hierarchies are systematic reviews of systematic reviews and the gold standard, systematic reviews of meta-analyses.[25] As of 2021, it appears that there are hundreds of randomised controlled trials of wearable activity trackers, and numerous systematic reviews and meta-analyses focused on activity trackers in various clinical and non-clinical populations. Now is a good time to amass the highest-quality evidence and consolidate understanding of the effectiveness of using activity trackers to intervene on physical activity. Specifically, we aimed to do a review of systematic reviews and meta-analyses (an umbrella review) examining the effectiveness of wearable activity trackers for increasing physical activity, and for affecting physiological and psychosocial outcomes that might arise from a change in physical activity.

See Online for appendix

Methods

Protocol and registration

This systematic review of systematic reviews and meta-analyses (hereafter referred to as an umbrella review) was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[26] and was prospectively registered on PROSPERO (CRD42021246494).[27]

Eligibility criteria

The eligibility criteria were structured using the PICOS (ie, population, intervention, comparison, outcome, and study type) framework[28] as follows: any human population of any age as the population; use of a wearable activity tracker as the intervention (ie, pedometer, accelerometer, activity monitor, or a step-counting smartphone application); no intervention, usual care, waiting list control, or similar intervention without an activity-tracker component as the comparator; physical activity (eg, step count or minutes of activity), or down-stream physiological or psychosocial outcomes (ie, outcomes that might derive from physical activity, such as bodyweight and wellbeing) as the outcome; and systematic reviews of experimental studies as the study design. The systematic reviews that used meta-analytic analyses, narrative analyses, or both, were eligible. Systematic reviews were excluded if they had the following characteristics: the wrong population (ie, non-human population); the wrong intervention, because either activity trackers were used only to monitor outcomes rather than being used as an intervention, or activity trackers were not the key intervention of interest (eg, a review exploring physical activity interventions with a subgroup of studies that used activity trackers); the wrong comparator—eg, systematic reviews were excluded if more than a third of the original studies in the review used a design that was inappropriate to address the research question of our meta-analysis (ie, they included activity trackers in both the intervention and control conditions, or activity trackers were used for the control condition in a study of an alternate, more intensive intervention); the wrong outcome, meaning there were no physical activity or downstream outcomes; the wrong study design, such as scoping reviews, literature reviews, or primary research (eg, randomised controlled trials); not published in a peer-reviewed journal or database; or they were a conference abstract (however, full-length conference articles were included).

Search strategy

Search strings related to activity trackers, interventions, and systematic reviews were developed. Search terms are listed in the appendix (pp 1–2). Seven databases were searched from database inception on April 8, 2021: Embase, MEDLINE, Ovid Emcare, Scopus, SPORTDiscus, the Cochrane Library, and Web of Science. When possible, searches were limited to English language and human studies (appendix pp 1–2).

Selection process

Database search results were imported into EndNote (EndNote x9; Clarivate, Philadelphia, PA, USA) and duplicates were removed. Results were then exported to Covidence (Veritas Health Innovation, Melbourne, VIC, Australia) for title, abstract, and full-text screening; each paper was screened by two of eight independent reviewers (TF, HB, AJC, KD, DD, DK, EO'C, RV), with disagreements resolved by discussion. Inter-rater reliability was calculated as the mean proportionate agreement across all pairs of reviewers.

Data extraction, data items and risk of bias

Data were extracted by pairs of eight independent reviewers using a data-extraction form (appendix pp 3–16) developed for the Review based on a template used in previous reviews of systematic reviews.[29,30] Extracted data items included the number of original studies in the review, sample size and characteristics (eg, age and population type), setting (eg, clinical setting or workplace), nature of the included interventions, outcomes measured (physical activity, physiological outcomes such as bodyweight and blood pressure, and psychosocial outcomes such as quality of life and social support), and results.

Risk of bias for each included systematic review was assessed by two of eight independent reviewers using the 16-item AMSTAR 2 tool,[31] with each item scored as no, partial yes, or yes. Seven items that can affect the validity of a review are deemed to be critical and nine non-critical, although these ratings can be altered depending on the context of the review.[31] For the current Review, we deemed one of the critical items to be non-critical (item 7—justification for excluding individual studies—the team agreed that since this was not required in PRISMA reporting guidelines, it should not be considered a critical weakness). The remaining six critical domains were therefore protocol registration, adequacy of search strategy, risk-of-bias assessment, appropriateness of meta-analytical methods, use of risk of bias during interpretation, and assessment of publication bias. Reviews were rated as high confidence (0 critical weaknesses and <3 non-critical weaknesses), moderate confidence (1 critical weakness and

<3 non-critical weaknesses), low confidence (>1 critical weakness and <3 non-critical weaknesses), or critically low confidence (>1 critical weakness and ≥3 non-critical weaknesses).

Deviations from the registered protocol

An exclusion criterion was added after study registration, because we found some systematic reviews that otherwise met the inclusion criteria but became ineligible because they had a single or small number of studies that did not meet the comparator inclusion criterion. Following team discussion, we decided that systematic reviews that met all other inclusion criteria should be included in the Review, provided no more than a third of studies used an inappropriate design for our research question.

After study registration, we opted to use the AMSTAR 2 risk-of-bias tool instead of the Joanna Briggs tool, because it was more detailed and therefore appeared to be more sensitive in differentiating the methodological quality of systematic reviews.

Following feedback from our peers, we also performed an additional analysis presenting the findings disaggregated by device type.

Umbrella review synthesis methods

We wanted to understand the degree of overlap in component studies (eg, the original randomised controlled trials) captured in the various systematic reviews and meta-analyses that were included in this umbrella review. The degree of overlap was quantified using the corrected covered area (CCA) method,[32] whereby a CCA of 100% indicates that every systematic review and meta-analysis identified in this umbrella review comprised exactly the same component studies, whereas a CCA of 0% indicates that every systematic review and meta-analysis in this umbrella review included entirely unique studies. A CCA of 0–5% is said to indicate slight overlap, 6–10% moderate overlap, 11–15% high overlap, and more than 15% very high overlap.[32]

Because of the overlap in some of the component studies included in the various meta-analyses captured in this umbrella review, a meta-analysis of the results of the meta-analyses was not appropriate, given that it would count results from some component studies several times, violating meta-analysis principles.[33] Rather, a narrative synthesis was used,[34] in which meta-analyses were grouped by outcome measure and presented using forest plots that summarised the effect size, population, sample size, and heterogeneity (eg, I^2) of each meta-analysis. Effect sizes are presented as originally reported by each systematic review. For each outcome, results from meta-analyses reporting standardised effect sizes (ie, standardised mean difference [SMD] and Hedge's *g*) were presented separately from results from meta-analyses reporting unstandardised effect sizes (ie, mean difference, weighted mean difference, and ratio of means).

For systematic reviews using narrative synthesis, we calculated the number of outcome-specific studies that found significantly favourable differences (ie, differences that the systematic reviews reported as being statistically significant—eg, with CIs that did not overlap with 0—and which were favourable for health and wellbeing). These totals were expressed as a percentage of the total number of studies. Findings were summarised by plotting the relationship between the percentage of studies showing significant favourable associations and the total number of studies reporting the outcome.

Sensitivity analyses were undertaken descriptively to examine whether findings were consistent when low-confidence and critically low-confidence systematic reviews were excluded. In addition, narrative subgroup analyses were done to examine whether results varied on the basis of age group (child *vs* adults *vs* older adults), clinical characteristics (ie, healthy and various clinical populations), and device type (pedometers *vs* consumer activity trackers such as smart watches and fitness trackers). We considered results to be consistent if the effect sizes were of the same category according to Cohen's *d* (d=0·2 was considered to be a small effect size, d=0·5 a medium effect size, and d≥0·8 a large effect size). A funnel plot was created as a visual aid to detect reporting bias (ie, missing results). Finally, the certainty (or confidence) in the body of evidence was compiled in a figure summarising the findings, with a red, orange, and green colour-coding system used to aid visual interpretation.[34]

Results

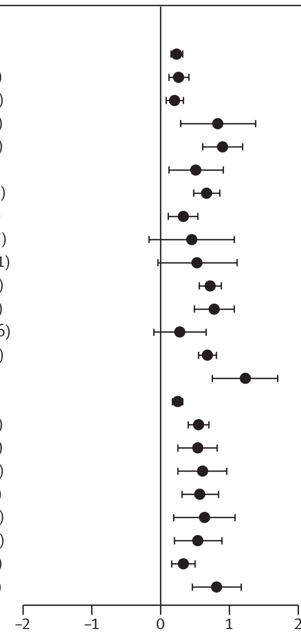
Study selection

The search identified 2382 records. Following removal of duplicates, title and abstract screening, and full-text reviews, 39 systematic reviews were eligible[7–11,35–68] (see appendix p 17 for PRISMA flow chart, and appendix pp 18–21 for the full list of articles excluded at full-text review with reasons).

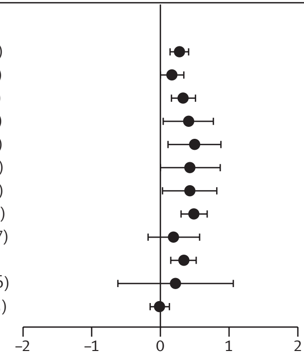
Inter-rater reliability (mean proportionate agreement across all pairs of reviewers) was 0·85 (95% CI 0·75–0·95) for title and abstract screening, and 0·76 (0·61–0·90) for full-text screening.

The characteristics for the systematic reviews included in the umbrella review are provided in the appendix (pp 22–27). These reviews were published between 2007 and 2021, with the majority published since 2018 (n=26). The number of studies in each systematic review ranged from four[54,58] to 70[37] (median 14, IQR 11–25) and the number of participants in each systematic review ranged from 167[62] to 73 440[36] (median 2003, 1262–2839). Most systematic reviews included studies of adults (n=31), whereas four systematic reviews included all ages,[40,49,63,67] three included only older adults,[39,51,59] and one included only children.[62] Half of the reviews focused on clinical populations (n=20; eg, type 2 diabetes,

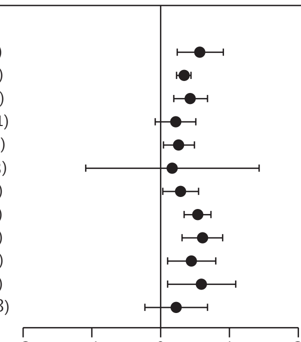
	Population	Analysis grouping	Number of studies	n	Model	Value (95% CI)	p value	I^2 (%)
Step count								
Brickwood et al, 2019[8]	Mixed*, adults	Overall	11	2144	R, SMD	0·23 (0·15 to 0·32)	<0·0001	3
		SG: MF	7	1210	R, SMD	0·26 (0·12 to 0·41)	0·0004	25
		SG: AT-based	5	934	R, SMD	0·20 (0·08 to 0·33)	0·002	0
Davergne et al, 2019[40]	RMD†, all ages	Overall	7	463	R, SMD	0·83 (0·29 to 1·38)	0·003	85
de Vries et al, 2016[41]	OW OB, adults	Overall	5	417	R, SMD	0·90 (0·61 to 1·19)	<0·001	49
Gal et al, 2018[45]	Mixed*, adults	Overall	7	1392	R, SMD	0·51 (0·12 to 0·91)	0·01	90
		SG: low RoB	3	469	R, SMD	0·67 (0·48 to 0·86)	<0·0001	0
		SG: post-intervention	6	573	R, SMD	0·33 (0·11 to 0·54)	0·003	29
Hannan et al, 2019[47]	Cardiac rehab‡, adults	Overall	4	341	R, SMD	0·45 (–0·17 to 1·07)	0·15	81
Kang et al, 2009[49]	Mixed*, all ages	Overall	32	2570	R, G	0·53 (–0·04 to 1·11)	··	··
		SG: combined ages	5	307	R, G	0·72 (0·56 to 0·88)	··	··
		SG: children	10	296	R, G	0·78 (0·49 to 1·07)	··	··
		SG: adults	32	1924	R, G	0·28 (–0·10 to 0·66)	··	··
		SG: older adults	3	43	R, G	0·68 (0·55 to 0·81)	<0·001	77
Liu et al, 2020[53]	General, older adults	Overall	2	83	R, SMD	1·23 (0·75 to 1·70)	<0·0001	0
Lynch et al, 2020[55]	General, adults	Overall	2	904	F, SMD	0·25 (0·17 to 0·32)	<0·01	57
Oliveira et al, 2020[59]	Mixed*, older adults	Overall	23	2766	R, SMD	0·55 (0·40 to 0·70)	0	85·8
		SG: clinical conditions	14	··	R, SMD	0·54 (0·25 to 0·82)	0·001	··
		SG: general population	9	··	R, SMD	0·61 (0·25 to 0·96)	0·002	··
Qiu et al, 2019[9]	COPD, adults	Overall	15	644	R, SMD	0·57 (0·31 to 0·84)	<0·05	75
		SG: AT + PR vs PR	6	146	R, SMD	0·64 (0·19 to 1·08)	<0·05	64
		SG: AT vs usual care	9	498	R, SMD	0·54 (0·20 to 0·89)	<0·05	81
Tang et al, 2020[66]	Healthy, adults	Overall	7	543	F, SMD	0·33 (0·16 to 0·50)	<0·001	2
Vaes et al, 2013[68]	T2DM, adults	Overall	8	633	R, SMD	0·81 (0·46 to 1·17)	<0·0001	73



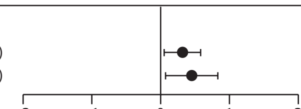
	Population	Analysis grouping	Number of studies	n	Model	Value (95% CI)	p value	I^2 (%)
Moderate-to-vigorous physical activity								
Brickwood et al, 2019[8]	Mixed*, adults	Overall	12	2237	R, SMD	0·28 (0·14 to 0·41)	<0·0001	46
		SG: AT-based	3	522	R, SMD	0·17 (0·00 to 0·34)	0·05	0
		SG: MF	10	1715	R, SMD	0·33 (0·16 to 0·51)	0·0002	58
Davergne et al, 2019[40]	RMD†, all ages	Overall	3	117	R, SMD	0·41 (0·04 to 0·77)	0·02	0
de Vries et al, 2016[41]	OW OB, adults	MF	3	651	R, SMD	0·50 (0·11 to 0·88)	0·01	74
		MF with AT vs MF no AT	3	83	R, SMD	0·43 (0·00 to 0·87)	0·05	0
Gal et al, 2018[45]	Mixed*, adults	Objective PA	6	1151	R, SMD	0·43 (0·03 to 0·82)	0·03	85
		SG: objective, low RoB	4	440	R, SMD	0·49 (0·30 to 0·68)	<0·0001	0
		Subjective PA	3	798	R, SMD	0·19 (–0·18 to 0·57)	0·32	62
Larsen et al, 2019[51]	General, older adults	Overall	8	1686	R, SMD	0·30 (0·2 to 0·52)	0·005	66
Liu et al, 2020[53]	General, older adults	Overall	3	201	R, SMD	0·22 (–0·62 to 1·06)	0·61	88
Lynch et al, 2020[55]	General, older adults	Overall	3	512	F, SMD	–0·01 (–0·15 to 0·13)	··	20



	Population	Analysis grouping	Number of studies	n	Model	Value (95% CI)	p value	I^2 (%)
Physical activity								
Baskerville et al, 2017[35]	T2DM, adults	Overall	10	1213	R, SMD	0·57 (0·24 to 0·91)	0	84
Braakhuis et al, 2019[7]	Former patient§, adults	Overall	14	1432	F, SMD	0·34 (0·23 to 0·44)	<0·0001	49
Cooper et al, 2018[39]	Mixed*, adults	Accelerometers	4	400	R, SMD	0·43 (0·19 to 0·68)	0·30	18·6
		Pedometers	4	488	R, SMD	0·22 (–0·08 to 0·51)	0·12	48·2
Goode et al, 2017[46]	Mixed*, adults	Overall	12	1378	R, SMD	0·26 (0·04 to 0·49)	0·001	64·7
		SG: active control	3	109	R, SMD	0·17 (–1·09 to 1·43)	0·12	52·3
		SG: inactive control	9	1269	R, SMD	0·29 (0·03 to 0·55)	<0·001	70·3
Larsen et al, 2019[51]	General, older adults	Overall	21	2704	R, SMD	0·54 (0·34 to 0·73)	<0·001	79·2
Oliveira et al, 2020[59]	Mixed*, older adults	Overall	3	228	R, SMD	0·61 (0·31 to 0·90)	0·35	10
Tang et al, 2020[66]	Healthy, adults	Overall	12	1356	R, SMD	0·45 (0·10 to 0·80)	0·01	88
		SG: all weights	7	727	R, SMD	0·59 (0·10 to 1·09)	0·02	90
		SG: overweight	5	629	R, SMD	0·23 (–0·23 to 0·68)	0·33	76



	Population	Analysis grouping	Number of studies	n	Model	Value (95% CI)	p value	I^2 (%)
Energy expenditure (kcal/week)								
Brickwood et al, 2019[8]	Mixed*, adults	Overall	5	347	R, SMD	0·32 (0·05 to 0·58)	0·02	33
de Vries et al, 2016[41]	OW OB, adults	MF with AT vs MF no AT	3	110	R, SMD	0·45 (0·07 to 0·83)	0·02	0



chronic obstructive pulmonary disease [COPD], overweight, or obesity),[7,9–11,35,38,40–42,44,47,50,54,56–58,60,61,64,68] eight focused on non-clinical populations,[36,43,51,53,55,62,65,66] and 11 included both.[8,37,39,45,46,48,49,52,59,63,67] The systematic reviews focused on physical activity (either alone [n=26][7,8,10,35,37,39–41,43–45,47,49–56,59,60,62,64,66,68] or combined with other outcomes [n=6][9,36,42,46,63,65]), weight loss (n=4),[11,38,57,61] weight control (n=1),[67] chronic disease management (n=1),[48] and pain and disability (n=1).[58]

Risk-of-bias results

Methodological strengths across the included reviews included providing conflict-of-interest statements (97%), performing study selection in duplicate (77%), and reporting study heterogeneity in meta-analyses (74%). Common methodological weaknesses were failure to provide the search strategy for all databases (only 5% successfully met this criterion), failure to provide a full list of excluded studies (8%), and failure to extract data on funding sources (10%). One review was rated as high confidence,[43] two were rated as moderate confidence,[11,51] two were rated as low confidence,[8,35] and the remaining 34 were rated as critically low confidence (appendix pp 28–29).

Study overlap

The 39 included systematic reviews reported a total of 719 component studies including duplicates, of which 390 were unique component studies (ie, duplicates removed). The CCA was 2·3%, indicating slight overlap.

Results of the meta-analyses

Wearable activity trackers and physical activity

In 25 systematic reviews, the authors meta-analysed the effect of wearable activity trackers on physical activity;[7–10,35,37,39–43,45–47,49–51,53–55,59,60,63,66,68] however, the physical activity metric varied among reviews (step counts, moderate-to-vigorous physical activity [MVPA], physical activity energy expenditure, and walking), as did the statistical approach used to analyse the metric. The meta-analysis results for physical activity outcomes from systematic reviews that reported standardised effect sizes are summarised here (figure 1). Wearable activity trackers increased physical activity outcomes with effect sizes in the order of 0·28 to 0·57 (first and third quartiles, respectively). Notably, wearable activity trackers increased daily step count with an average effect size of 0·6 (medium effect), increased physical activity and energy expenditure with an average effect size of 0·4 (medium effect), and increased MVPA with an average effect size of 0·3 (small effect). When the results for physical activity meta-analyses were presented in mean differences and ratio of means (appendix pp 35–36), interventions using wearable activity trackers increased step counts on average by around 1800 steps per day, walking time by approximately 40 min per day, and MVPA by around 6 min per day.

Wearable activity trackers and physiological outcomes

The effect of wearable activity trackers on physiological outcomes were meta-analysed in 17 systematic reviews.[9–11,35,41–43,46,47,51,57,60,61,63,66–68] The meta-analysis results for physiological outcomes from systematic reviews that reported standardised effect sizes are summarised here (figure 2; see appendix pp 31–32 for meta-analyses based on mean differences). The effects were in favourable directions (ie, negative effects suggesting improvement in BMI, blood pressure, cholesterol, glycosylated haemoglobin, and waist circumference and bodyweight, and positive effects suggesting improvement in aerobic capacity), although in most instances the effect sizes were small in magnitude and only occasionally statistically significant. The outcomes with the strongest evidence for improvement were bodyweight (six of nine meta-analyses showed improvement, with effect sizes of approximately –2·0 indicating a large effect, or –0·5 kg to –1·5 kg),[11,46,57,61,63,67] waist circumference (two of three meta-analyses showed a significant improvement, with an effect size of around –0·7, or –1·5 cm),[42,67] BMI (five of ten meta-analyses reported a significant reduction, with an effect size of around –0·5, or –0·5 kg/m²),[10,11,43,67,68] and aerobic capacity (two of three meta-analyses showed a significant improvement, with an effect size of around 0·3, or 1·7 mL/kg per min improvement in peak maximal oxygen consumption during exercise [VO_2]).[9,47] For systolic blood pressure, three out of five meta-analyses reported a significant reduction in effect size of around –0·2, or 2–4 mmHg,[10,42,68] and a sole meta-analysis addressing resting heart rate suggested a borderline significant reduction (p=0·07) of around two beats per min (appendix p 32).[42] There was little evidence that interventions based on activity trackers significantly affected diastolic blood pressure (SMD –0·1, 95% CI –0·28 to 0·10), cholesterol (SMD –0·06, –0·31 to 0·19), triglycerides (two meta-analyses reported a non-significant improvement),[10,42] glycosylated haemoglobin (two of four meta-analyses reported a small significant improvement, SMD around –0·16),[35,68] and fasting glucose (two meta-analyses reported non-significant improvements).[10,42]

Figure 1:* Summary of meta-analysis results for physical activity outcomes reported in standardised mean differences and Hedge's *g

AT=activity tracker. COPD=chronic obstructive pulmonary disease. F=fixed model. G=Hedge's *g*. MF=multifaceted. OW OB=overweight and obese. PA=physical activity. PR=pulmonary rehabilitation. R=random model. rehab=rehabilitation. RMD=rheumatic and musculoskeletal diseases. RoB=risk of bias. SG=subgroup. SMD=standardised mean difference. T2DM=type 2 diabetes. *Healthy participants and those with disease. †Including lower-extremity osteoarthritis, lower back pain, or chronic inflammatory rheumatic diseases (ie, spondyloarthritis, rheumatoid arthritis, psoriatic arthritis, or juvenile arthritis). ‡Myocardial infarction, acute coronary syndrome, percutaneous coronary intervention, coronary artery disease, and history of cardiac surgery. §Stroke, cardiac conditions, geriatric Parkinson's disease, and COPD.

Wearable activity trackers and psychosocial outcomes

In a small number of the included systematic reviews, the authors meta-analysed the effect of interventions using activity trackers on psychosocial outcomes; quality of life was addressed in four reviews[40,43,51,59] and disability and pain each in a single review.[58] The meta-analysis results for psychosocial outcomes from systematic reviews that reported standardised (SMD) or unstandardised (mean differences) effect sizes are summarised here (figure 3). Overall, there was little evidence that interventions based on activity trackers affected quality of life. One review looking at effects on disability and pain reported non-significant medium effect sizes in a favourable direction (ie, a decrease in disability and pain: –0·81 [95% CI –2·34 to 0·73] to –0·50 [–1·91 to 0·91]).[58]

	Population	Analysis grouping	Number of studies	n	Model	Value (95% CI)	p value	I^2 (%)
BMI								
Baskerville et al, 2017[35]	T2DM, adults	Overall	7	343	R, SMD	0·06 (–0·19 to 0·32)	0·23	26
Larsen et al, 2019[51]	Healthy, older adults	Overall	3	570	R, SMD	0·15 (–0·01 to 0·31)	0·75	0
McDonough et al, 2021[57]	Clinical, adults	AT-based	··	··	R, SMD	–1·99 (–4·95 to 0·96)	··	··
		Multifaceted inc AT	··	··	R, SMD	–0·96 (–2·95 to 1·00)	··	··
Vaes et al, 2013[68]	T2DM, adults	Overall	10	399	R, SMD	–0·24 (–0·48 to –0·01)	0·04	25
Yen and Chiu, 2019[67]	Mixed*, all ages	Overall	12	1826	R, G	–0·84 (–1·29 to –0·40)	<0·001	94
Bodyweight								
McDonough et al, 2021[57]	Clinical, adults	AT-based	··	··	R, SMD	–2·76 (–4·80 to –0·81)	··	··
		Multifaceted inc AT	··	··	R, SMD	–2·00 (–4·04 to 0·04)	··	··
Tang et al, 2020[66]	Healthy, adults	Overall	4	589	F, SMD	0·13 (–0·34 to 0·60)	0·58	0
Yen and Chiu, 2019[67]	Mixed*, all ages	Overall	19	2128	R, G	–0·59 (–0·84 to –0·35)	<0·001	86
Waist circumference								
Yen and Chiu, 2019[67]	Mixed*, all ages	Overall	4	331	R, G	–0·67 (–1·31 to –0·03)	0·041	88
Blood pressure								
Baskerville et al, 2017[35]	T2DM, adults	Systolic blood pressure	7	1055	R, SMD	–0·05 (–2·05 to 1·95)	>0·05	0
Vaes et al, 2013[68]	T2DM, adults	Diastolic blood pressure	11	708	R, SMD	–0·09 (–0·28 to 0·10)	0·34	24
Vaes et al, 2013[68]	T2DM, adults	Systolic blood pressure	11	708	R, SMD	–0·18 (–0·34 to –0·01)	0·04	9
Cholesterol								
Baskerville et al, 2017[35]	T2DM, adults	Total or LDL cholesterol	5	256	R, SMD	–0·06 (–0·31 to 0·19)	0·73	0
Glycosylated haemoglobin (HbA_{1c})								
Baskerville et al, 2017[35]	T2DM, adults	Overall	9	1094	R, SMD	–0·10 (–0·13 to –0·06)	0·69	0
Vaes et al, 2013[68]	T2DM, adults	Overall	16	1054	R, SMD	–0·23 (–0·41 to –0·05)	0·01	43
Aerobic capacity (6 min walk test)								
Larsen et al, 2019[51]	Healthy, older adults	Overall	4	754	R, SMD	0·19 (–0·10 to 0·48)	0·12	49
Qiu et al, 2018[9]	COPD, adults	Overall	13	534	R, SMD	0·30 (0·16 to 0·45)	<0·05	15
		SG: AT + PR *vs* PR	6	193	R, SMD	0·32 (0·01 to 0·62)	<0·05	42
		SG: AT *vs* usual care	7	341	R, SMD	0·29 (0·14 to 0·44)	<0·05	1

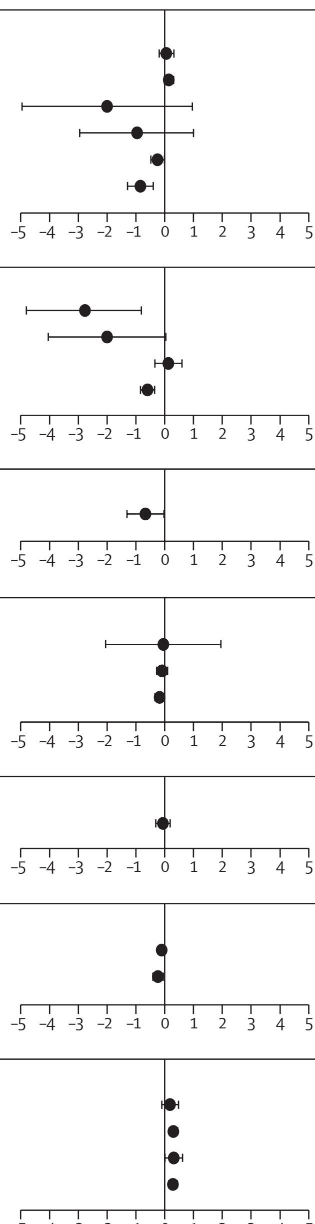
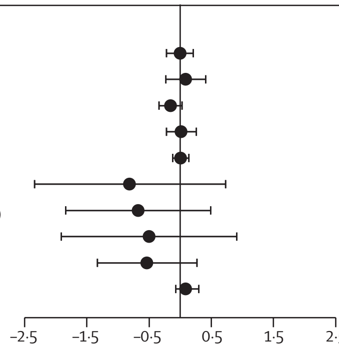


Figure 2:* Summary of meta-analysis results for physiological outcomes reported in standardised mean differences and Hedge's *g

AT=activity tracker. COPD=chronic obstructive pulmonary disease. F=fixed model. G=Hedge's *g*. inc=including. PR=pulmonary rehabilitation. R=random model. SG=subgroup. SMD=standardised mean difference. T2DM=type 2 diabetes. *Healthy participants and those with disease.

	Population	Analysis grouping	Number of studies	n	Model	Value (95% CI)	p value	I^2 (%)
Psychosocial outcomes (SMD)								
Davergne et al, 2019[40]	RMD*, all ages	Disability	5	339	R, SMD	0·00 (–0·22 to 0·21)	0·99	0
		Functional tests	4	321	R, SMD	0·09 (–0·23 to 0·41)	0·56	36
		Pain	7	580	R, SMD	–0·15 (–0·34 to 0·03)	0·11	15
		QoL: health-related	4	281	R, SMD	0·02 (–0·22 to 0·26)	0·88	0
Larsen et al, 2019[51]	General, older adults	QoL: health-related	5	1038	R, SMD	0·01 (–0·12 to 0·14)	0·54	0
Oliveira et al, 2018[58]	Chronic MSKP, adults	Disability: ST FU	2	116	R, SMD	–0·81 (–2·34 to 0·73)	0·30	93
		Disability: IM FU	3	114	R, SMD	–0·67 (–1·84 to 0·49)	0·26	90
		Pain: ST FU	2	116	R, SMD	–0·50 (–1·91 to 0·91)	0·49	92
		Pain: IM FU	3	114	R, SMD	–0·53 (–1·33 to 0·27)	0·19	80
Oliveira et al, 2020[59]	Mixed†, older adults	QoL	9	1433	R, SMD	0·09 (–0·07 to 0·30)	0·051	45
Quality of life (MD)								
Freak-Poli et al, 2020[43]	Employees, adults	QoL: mental, complete	2	168	R, MD	3·18 (–1·32 to 7·68)	0·17	70
		QoL: physical, complete	2	310	R, MD	0·21 (–0·29 to 0·70)	0·42	68
		QoL: physical, FU	2	310	R, MD	0·08 (–0·14 to 0·30)	0·47	0



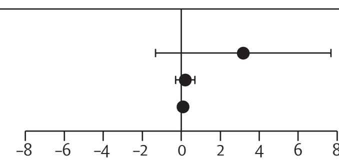


Figure 3: **Summary of meta-analysis results for psychosocial outcomes reported in standardised mean differences and mean differences**
Complete=at study completion. FU=follow-up. IM=intermediate. MD=mean difference. MSKP=musculoskeletal pain. QoL=quality of life. R=random model. RMD=rheumatic and musculoskeletal diseases. SMD=standardised mean difference. ST=short-term. *Including lower-extremity osteoarthritis, lower back pain, or chronic inflammatory rheumatic diseases (ie, spondyloarthritis, rheumatoid arthritis, psoriatic arthritis, or juvenile arthritis). †Healthy participants and those with disease.

Results of the narrative synthesis

Physical activity outcomes (energy expenditure, light physical activity, moderate physical activity, MVPA, vigorous physical activity, overall physical activity, steps, and walking), physiological outcomes (eg, body composition, fitness, and cardiometabolic markers such as cholesterol and diabetes risk), and psychosocial outcomes (eg, quality of life related to depression, anxiety, and stress, and wellbeing) were narratively synthesised in 15,[7,36,38,41,43,44,47,52,54–56,62,64,65,68] 11,[36,38,42,43,44,47,48,52,56,64,65] and six[36,38,47,54,58,59] systematic reviews, respectively. The percentage of original studies reporting significantly favourable differences and the total number of studies reporting each outcome are summarised (figure 4). As the number of studies increased, the percentage of significant favourable studies was about 40–50%. Considering all studies, 41% showed significant favourable associations, 57% non-significant associations, and 3% significant unfavourable associations.

Results according to risk of bias and population subgroups

Only three of the 39 included systematic reviews were graded as being of high or moderate confidence in the risk-of-bias appraisal (appendix pp 28–29).[11,43,51] Focusing on their results alone, effects were similar in magnitude and significance to the values reported for all included studies. For example, for MVPA, the meta-analysis by Larsen and colleagues[51] found an effect size of 0·34 (95% CI 0·20–0·52; p=0·005; figure 1), and the effect sizes for all other systematic reviews also centred around approximately 0·3 (figure 1). For physical activity, Larsen and colleagues[51] reported an effect size of 0·54 (95% CI 0·34–0·73; p<0·0001) and Freak-Poli and colleagues[43] reported a ratio of means of 1·37 (95% CI 1·11–1·71; p=0·004; appendix p 30), which is similar to, or marginally larger, than the effect sizes reported in other reviews (centred around an SMD of 0·4; figure 1). Results for physiological outcomes were also broadly consistent, suggesting a significant improvement in adiposity-related outcomes (BMI and weight), but not cardiometabolic markers.

Physical activity outcomes were consistently shown to improve in children, younger adults, adults, and older adults, with similar effect sizes. Further, physical activity outcomes improved in both general (ie, healthy) and all clinical populations captured in this Review. The beneficial effects for body composition were apparent across a range of populations, including healthy adults and those with type 2 diabetes, COPD, cardiovascular disease, overweight, and obesity. Beneficial effects on cardiometabolic markers were reported for clinical populations (adults) only. Specifically, benefits for blood pressure were reported in people with type 2 diabetes[68] (systolic SMD –0·18, 95% CI –0·34 to –0·01; figure 2) and populations with mixed disease (which included diabetes, COPD, cardiovascular disease, overweight or obesity, and Alzheimer's disease; systolic mean difference –3·79 mmHg, 95% CI –4·54 to –3·04; appendix p 31),[42] benefits for LDL cholesterol in populations of mixed disease (mean difference –0·32 mmol/L, 95% CI –0·51 to –0·12; appendix p 32),[42] and benefits for fitness in adults with COPD (SMD 0·30, 95% CI 0·16 to 0·45; figure 2) or those undergoing cardiac

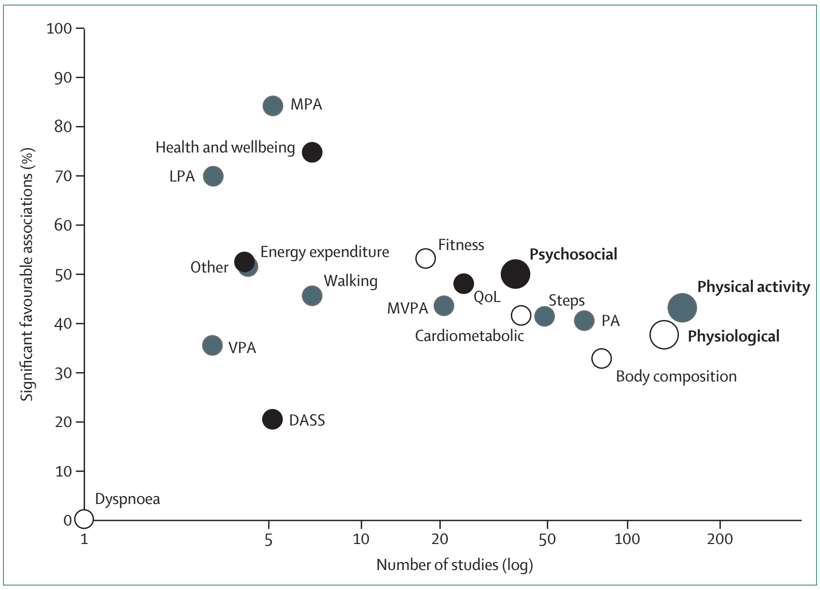


Figure 4: **Relationship between the percentage of studies showing significant favourable associations and the total number of included studies, for reviews that did not use meta-analysis**

The larger dots indicate lumped values for the overall categories of physical activity outcomes, physiological outcomes, and psychosocial outcomes. DASS=Depression, Anxiety and Stress Scale. LPA=light physical activity. MPA=moderate physical activity. MVPA=moderate-to-vigorous physical activity. PA=physical activity. QoL=quality of life. VPA=vigorous physical activity.

rehabilitation (VO_2 peak mean difference 1·65 L/min, 95% CI 0·64 to 2·66; appendix p 32).[9,47]

Results appeared broadly consistent when they were compared by device type (ie, pedometer-based interventions and consumer activity-tracker-based interventions). Direct comparison between device types was only possible for step count and bodyweight. For step count, three meta-analyses of consumer activity trackers reported effect sizes[8,55,66] ranging from 0·23 (95% CI 0·15 to 0·32) to 0·33 (0·16 to 0·50; ie, small effects), all statistically significant (appendix p 35). The single meta-analysis of pedometer-based interventions reported an overall effect size of 0·53 (95% CI –0·04 to 1·11; ie, moderate effect); however, this result was statistically non-significant (appendix p 35).[49] Three other meta-analyses provided mean-difference data for step count; the meta-analysis by Ringeval and colleagues[63] of consumer activity-tracker-based interventions reported an increase of 951 daily steps (95% CI 476 to 1425), whereas two meta-analyses of pedometer-based interventions reported increases of 1822 (95% CI 751 to 2894)[60] to 2491 (1098 to 3885)[10] daily steps (appendix p 35). For bodyweight, two meta-analyses[46,63] of consumer activity-tracker interventions reported an overall weight loss of –1·5 kg (95% CI –2·81 to –0·14) to –1·7 kg (–3·03 to –0·28), whereas two meta-analyses[11,61] of pedometer interventions reported weight loss of –0·7 kg (95% CI –1·12 to –0·17) to –1·3 kg (–1·85 to –0·70; appendix p 37). The meta-analysis results from all reviews that reported results separately for consumer activity trackers[8,46,55,63,66] or pedometers[10,11,43,49,60,61] are provided in the appendix (pp 35–39).

The duration of the effects of wearable activity-tracker interventions was examined in three systematic reviews (appendix p 30).[11,37,43] Results suggested strong effects on step counts at 4–6 months (an increase of 1127 [95% CI 710–1543] steps per day), and smaller but statistically significant effects on step counts up to 4 years (an increase of 494 [251–738] steps per day).[37] Two reviews examined whether effects on body composition were sustained long term following the completion of wearable activity-tracker-based interventions; one review found no sustained effects,[43] whereas another found small but non-significant beneficial effects up to 1–2 years (BMI –0·21 kg/m², 95% CI –1·06 to 0·65; appendix p 31).[11]

Reporting biases: funnel plot

A funnel plot of the relationship between systematic review-level SMDs and the number of studies included in each meta-analysis is presented in the appendix (p 34). SMDs converge towards a value of 0·5–0·6. The plot is reasonably symmetrical with no evidence of publication bias.

Summary of findings

A summary of the certainty (or confidence) in the body of evidence regarding the effectiveness of interventions on the basis of wearable activity trackers on a variety of physical activity, physiological, and psychosocial outcomes is provided (figure 5). This summary shows that interventions using activity trackers have positive effects on physical activity metrics, a mixture of positive and non-significant effects for physiological metrics, and mostly suggestive positive effects for psychosocial outcomes.

Discussion

To the best of our knowledge, this is the first umbrella review regarding the effectiveness of wearable activity-tracker-based interventions. This study analysed 39 systematic reviews of experimental studies on the effectiveness of wearable activity trackers. A large number of meta-analyses and narrative syntheses addressed physical activity outcomes, and found that wearable trackers consistently outperformed controls for physical activity outcomes with moderate effect sizes. Fewer systematic reviews addressed physiological outcomes, and the effect sizes were smaller, and often non-significant. Only seven meta-analyses addressed psychosocial outcomes, finding small effect sizes for quality of life, whereas those for pain and disability were moderate. These patterns were consistent across all ages and disease states, and the gains appeared to be retained to at least 6 months' follow-up.

Together, the results from this umbrella review suggest there is consistent evidence that wearable activity trackers affect physical activity across a wide range of metrics. Our results suggest that interventions using wearables resulted on average in an extra 1800 steps per day, 40 min per day more walking, and a 6 min per day increase in MVPA. These improvements are relatively strong when considered in the broader field of physical activity interventions.[69,70] In public health terms, an increase of 5–10 min per day of MVPA is generally considered meaningful,[71] suggesting the degree of change seen in this umbrella review is clinically significant.

Domain	Metric	Explanation
Physical activity	Steps	19 meta-analyses and ten narrative analyses. 23 (79%) favoured intervention, none favoured control, and six (21%) were non-significant (although two of these were meta-analyses that approached significance, favouring intervention).
	Moderate-to-vigorous physical activity	Nine meta-analyses and six narrative analyses. 12 (80%) favoured intervention, 0 (0%) favoured control, and three (20%) were non-significant.
	Vigorous physical activity	No meta-analyses and three narrative analyses. Two (66%) favoured intervention, none favoured control, and one (33%) was non-significant. Only addressed by a small number of original studies.
	Moderate physical activity	No meta-analyses and five narrative analyses. Five (100%) favoured intervention. Only addressed by a small number of original studies.
	Light physical activity	No meta-analyses and two narrative analyses. Two (100%) favoured intervention. Only addressed by a small number of original studies.
	Physical activity	Eight meta-analyses and 11 narrative reviews. 17 (89%) favoured intervention, none favoured control, and two (11%) were non-significant.
	Walking	One meta-analysis and six narrative reviews. Four (57%) favoured intervention, one (14%) favoured control, and two (29%) were non-significant. Only addressed by a small number of original studies.
	Energy expenditure	Two meta-analyses and three narrative reviews. Four (80%) favoured intervention, none favoured control, and one (20%) was non-significant. Only addressed by a small number of original studies.
Physiological	Body composition	22 meta-analyses and nine narrative reviews. 20 (65%) favoured intervention, none favoured control, and 11 (35%) were non-significant.
	Blood pressure	Five meta-analyses and four narrative reviews. Seven (78%) favoured intervention, none favoured control, and two (22%) were non-significant.
	Fitness	Three meta-analyses and seven narrative reviews. Seven (70%) favoured intervention, none favoured control, and three (30%) were non-significant (although the meta-analysis approached significance, favouring intervention, and the two non-significant narrative analyses were each based on a single study).
	Cholesterol	Three meta-analyses and three narrative reviews. Four (66%) favoured intervention, none favoured control, and two (33%) were non-significant. Only addressed by a small number of original studies.
	Glycosylated haemoglobin	Four meta-analyses and two narrative reviews. Two (33%) favoured intervention, none favoured control, and four (66%) were non-significant. Only addressed by a small number of original studies.
	Blood glucose	Two meta-analyses and two narrative reviews. None favoured intervention, none favoured control, and four (100%) were non-significant. Only addressed by a small number of original studies.
	Triglycerides	Two meta-analyses and one narrative review. None favoured intervention, none favoured control, and three (100%) were non-significant. Only addressed by a small number of original studies.
	Resting heart rate	One meta-analysis and three narrative reviews. Two (50%) favoured intervention, none favoured control, and two (50%) were non-significant (although the meta-analysis approached significance, favouring intervention). Only addressed by a small number of original studies.
	Other, physiological	No meta-analyses and two narrative reviews. None favoured intervention, none favoured control, and two (100%) were non-significant.
Psychosocial	Quality of life	Four meta-analyses and six narrative reviews. Four (40%) favoured intervention, none favoured control, and six (60%) were non-significant.
	Disability	Three meta-analyses and no narrative reviews. None favoured intervention, none favoured control, and three (100%) were non-significant (although moderate-to-large effect size and small sample size).
	Pain	Two meta-analyses and no narrative reviews. None favoured intervention, none favoured control, and two (100%) were non-significant (although one meta-analysis approached significance, favouring intervention, and the other had a moderate effect size, and small sample size, favouring intervention).
	Anxiety, depression, and stress	No meta-analyses and three narrative reviews. One (33%) favoured intervention, none favoured control, and two (66%) were non-significant.
	Emotional wellbeing	No meta-analyses and two narrative reviews. Two (100%) favoured intervention. Only addressed by a small number of original studies.
	Other, psychosocial	No meta-analyses and three narrative reviews. Two (66%) favoured intervention, none favoured control, and one (33%) was non-significant.

Strong negative · Weak negative · Suggestive negative · Non-significant · Suggestive positive · Weak positive · Strong positive

Figure 5: **Summary of evidence for the effectiveness of wearable activity trackers on physical activity, physiological, and psychosocial outcomes**

Given the magnitude of change in physical activity found in this umbrella review, it might seem surprising that more compelling evidence for the effect of wearable activity-tracker-based interventions on physiological outcomes was not found. For many of these outcomes, the effect sizes were small to very small, and often statistically non-significant (except for weight-loss outcomes, which showed moderate effects, equating to a reduction in weight of about 1 kg or in BMI of 0·5 kg/m², and a reduction of 1·5 cm in waist circumference). One possible explanation is that many studies of physical activity interventions are reasonably short in duration (eg, up to 3 months), in which case measurable physiological benefits of sustained physical activity behaviour change might not have had time to fully manifest. Secondly, we applied the conventional definitions for interpreting effect sizes, by which an SMD of 0·2 is considered a small effect. However, effect sizes typically seen in different fields of research vary; in medical research, effect sizes in the range of 0·05 to 0·2 are common and sufficient to meaningfully affect health.[72] In this umbrella review, results for physiological outcomes were generally in a favourable direction, with effect sizes suggestive of small improvements (eg, an SMD of approximately –0·1 for blood pressure, –0·06 for cholesterol, and –0·16 for glycosylated haemoglobin), with a mix of significantly favourable and non-significant results (and an absence of significant, unfavourable results). Given these findings, some readers might consider our interpretation of findings for physiological outcomes as being overly conservative.

This umbrella review focused on intervention studies to capture the highest level of evidence. It embraced a wide range of physical activity, physiological, and psychosocial outcomes. A comprehensive search strategy was used, with study selection and data extraction processes completed in duplicate. There was only 2% overlap (a slight overlap) in studies between reviews. Additionally, we could locate no previously published umbrella reviews on this topic to which this umbrella review can be compared.

Any umbrella review is limited by constraints of the systematic reviews and meta-analyses it captures. Although the systematic reviews captured here together covered a wide range of clinical populations, these populations were not exhaustive. Most reviews were graded as low or critically low in confidence, using the AMSTAR 2 instrument, although this finding is similar to reports from other domains of medicine.[73–75] Results were similar across reviews with different risk-of-bias ratings, suggesting that the findings are robust. Although we subjectively grouped outcomes into broad domains and subcategories, a wide range of specific metrics were used in the reviews. For example, physical activity was variously operationalised as steps per day, meeting step goals, walking time or metabolic-equivalent minute, energy expenditure, accelerometer counts, time spent in moderate physical activity or MVPA, or exercise frequency. Nonetheless, results were quite robust across different metrics.

The reviews included a wide range of study designs. Although we excluded reviews with more than a third of studies with inappropriate designs (eg, studies including a wearable activity tracker in both the intervention and control conditions), some systematic reviews included some studies of these designs. The probable net effect of these inclusions would be to underestimate effect sizes.

Although this umbrella review provides strong assurance of the efficacy of wearable activity trackers in increasing physical activity, few of the systematic reviews and meta-analyses examined the effectiveness of wearable activity trackers alone. Brickwood and colleagues did in 2019,[8] separately meta-analysing 16 studies that used multifaceted interventions and seven that used wearable activity trackers alone. They showed that the multifaceted interventions yielded effects around 50% larger than those using wearable activity trackers alone.

Fewer reviews addressed physiological outcomes than those assessing physical activity, and fewer still a small range of psychosocial outcomes (mainly quality of life). These reviews included fewer studies and yielded mainly non-significant findings. These outcomes are probably downstream of changes in physical activity and might require longer follow-up to manifest.

The volume of included studies investigating the efficacy of wearable trackers is large—39 systematic reviews and 390 component experimental studies covering 163 841 participants (not accounting for overlap). The studies spanned a wide variety of subpopulations (populations with overweight or obesity, cardiovascular disease, diabetes, COPD, stroke, Alzheimer's disease, musculoskeletal pain, and Parkinson's disease) and a wide range of age groups from across the lifespan. Some limitations of the current body of evidence emerged. For example, the majority of component randomised controlled trials were done in high-income countries, and no systematic reviews addressing the activity-tracker-based interventions in populations with mental health illnesses were included in this umbrella review. Further research addressing these research gaps would be beneficial. Longer-term studies examining physiological, and a wider range of psychosocial, outcomes would add to the scientific record. Studies focusing on the efficacy of activity trackers used alone (or if being used as part of a multifaceted intervention, which features of multifaceted interventions work most effectively in combination with activity trackers), would also be valuable.

Wearable activity-tracker interventions are effective in increasing physical activity and supporting modest weight loss in a wide variety of clinical and non-clinical populations and age groups. The magnitude of benefit is of clinical importance, and gains appear to be durable for at least 6 months. Effects on other physiological and psychosocial outcomes are small and often non-significant. There is sufficient evidence to recommend the use of wearable activity trackers at least as an adjunct to programmes aiming to increase physical activity.

Contributors
TF, CM, TO, and RC were responsible for conceptualisation, study design, and search strategy for this Review. TF did all database searching and collating of results. TF, HB, AJC, KD, DD, DK, EO'C, and RV did the article screening, data extraction, and critical appraisal. All authors contributed to conflict resolution during screening. TF was responsible for data curation. TF, CM, TO, and RC contributed to data analysis and data interpretation, and wrote the original draft. All authors contributed to reviewing and editing of the final manuscript.

Declaration of interests
We declare no competing interests.

Data sharing
All data for this Review were obtained from published systematic reviews. Data extracted for this Review and database search strategies will be made available on reasonable request. For access, please email the corresponding author.

References

1 Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012; **380:** 219–29.
2 Ding D, Lawson KD, Kolbe-Alexander TL, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. *Lancet* 2016; **388:** 1311–24.
3 Degroote L, Hamerlinck G, Poels K, et al. Low-cost consumer-based trackers to measure physical activity and sleep duration among adults in free-living conditions: validation study. *JMIR Mhealth Uhealth* 2020; **8:** e16674.
4 Maher C, Ryan J, Ambrosi C, Edney S. Users' experiences of wearable activity trackers: a cross-sectional study. *BMC Public Health* 2017; **17:** 880.
5 Chaddha A, Jackson EA, Richardson CR, Franklin BA. Technology to help promote physical activity. *Am J Cardiol* 2017; **119:** 149–52.
6 Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *J Med Internet Res* 2014; **16:** e192.
7 Braakhuis HEM, Berger MAM, Bussmann JBJ. Effectiveness of healthcare interventions using objective feedback on physical activity: a systematic review and meta-analysis. *J Rehabil Med* 2019; **51:** 151–59.
8 Brickwood KJ, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2019; **7:** e11819.
9 Qiu S, Cai X, Wang X, et al. Using step counters to promote physical activity and exercise capacity in patients with chronic obstructive pulmonary disease: a meta-analysis. *Ther Adv Respir Dis* 2018; **12:** 1753466618787386.
10 Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 2007; **298:** 2296–304.
11 Cai X, Qiu SH, Yin H, et al. Pedometer intervention and weight loss in overweight and obese adults with type 2 diabetes: a meta-analysis. *Diabet Med* 2016; **33:** 1035–44.
12 Kandola A, Ashdown-Franks G, Hendrikse J, Sabiston CM, Stubbs B. Physical activity and depression: towards understanding the antidepressant mechanisms of physical activity. *Neurosci Biobehav Rev* 2019; **107:** 525–39.
13 Stubbs B, Vancampfort D, Rosenbaum S, et al. An examination of the anxiolytic effects of exercise for people with anxiety and stress-related disorders: a meta-analysis. *Psychiatry Res* 2017; **249:** 102–08.
14 International Data Corporation. Forecast wearables unit shipments worldwide from 2014 to 2024. Statista. https://www.statista.com/statistics/437871/wearables-worldwide-shipments/ (accessed May 10, 2021).
15 International Data Corporation. Fitness tracker device sales revenue worldwide from 2016 to 2022. Statista. https://www.statista.com/statistics/610433/wearable-healthcare-device-revenue-worldwide/ (accessed May 10, 2021).
16 Gonzalez R. Science says fitness trackers don't work. https://www.wired.com/story/science-says-fitness-trackers-dont-work-wear-one-anyway/ (accessed Nov 11, 2021).
17 Lieberman B. The 4 biggest problems with your fitness tracker, according to scientists. https://www.self.com/story/fitness-tracker-accuracy-wearables (accessed Nov 11, 2021).
18 Duus R, Cooray M. How we discovered the dark side of wearable fitness trackers. https://.com/how-we-discovered-the-dark-side-of-wearable-fitness-trackers-43363 (accessed Nov 11, 2021).
19 Heitz J. The trouble with fitness trackers. https://parentology.com/the-trouble-with-fitness-trackers/ (accessed Nov 11, 2021).
20 Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. *J Pers Med* 2017; **7:** 3.
21 Simpson CC, Mazzeo SE. Calorie counting and fitness tracking technology: associations with eating disorder symptomatology. *Eat Behav* 2017; **26:** 89–92.
22 Arcaya MC, Figueroa JFJHA. Emerging trends could exacerbate health inequities in the United States. *Health Affairs*; **36:** 992–98.
23 Jakicic JM, Davis KK, Rogers RJ, et al. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. *JAMA* 2016; **316:** 1161–71.
24 Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. *JAMA* 1995; **274:** 1800–04.
25 Moore A, Jull G. The systematic review of systematic reviews has arrived! *Man Ther* 2006; **11:** 91–92.
26 Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339:** b2700.
27 Maher C, Olds T, Curtis R, et al. The effectiveness of fitness tracker interventions: an umbrella review. PROSPERO: international prospective register of systematic reviews. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=246494 (accessed Nov 12, 2021).
28 Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995; **123:** A12–13.
29 Hutchesson MJ, Gough C, Müller AM, et al. eHealth interventions targeting nutrition, physical activity, sedentary behavior, or obesity in adults: a scoping review of systematic reviews. *Obes Rev* 2021; **22:** e13295.
30 Kracht CL, Hutchesson M, Ahmed M, et al. E-&mHealth interventions targeting nutrition, physical activity, sedentary behavior, and/or obesity among children: a scoping review of systematic reviews and meta-analyses. *Obes Rev* 2021; **22:** e13331.
31 Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017; **358:** j4008.
32 Pieper D, Antoine SL, Mathes T, Neugebauer EA, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. *J Clin Epidemiol* 2014; **67:** 368–75.
33 Deeks JJ, Higgins JP, Altman DG, Group CSM. Analysing data and undertaking meta-analyses. In: Cochrane Training, ed. Cochrane handbook for systematic reviews of interventions. London: Cochrane Training, 2019: 241–84.
34 Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid-Based Healthc* 2015; **13:** 132–40.
35 Baskerville R, Ricci-Cabello I, Roberts N, Farmer A. Impact of accelerometer and pedometer use on physical activity and glycaemic control in people with type 2 diabetes: a systematic review and meta-analysis. *Diabet Med* 2017; **34:** 612–20.
36 Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. *Digit Health* 2019; **5:** 2055207619839883.
37 Chaudhry UAR, Wahlich C, Fortescue R, Cook DG, Knightly R, Harris T. The effects of step-count monitoring interventions on physical activity: systematic review and meta-analysis of community-based randomised controlled trials in adults. *Int J Behav Nutr Phys Act* 2020; **17:** 129.

38 Cheatham SW, Stull KR, Fantigrassi M, Motel I. The efficacy of wearable activity tracking technology as part of a weight loss program: a systematic review. *J Sports Med Phys Fitness* 2018; **58:** 534–48.

39 Cooper C, Gross A, Brinkman C, et al. The impact of wearable motion sensing technology on physical activity in older adults. *Exp Gerontol* 2018; **112:** 9–19.

40 Davergne T, Pallot A, Dechartres A, Fautrel B, Gossec L. Use of wearable activity trackers to improve physical activity behavior in patients with rheumatic and musculoskeletal diseases: a systematic review and meta-analysis. *Arthritis Care Res* 2019; **71:** 758–67.

41 de Vries HJ, Kooiman TJ, van Ittersum MW, van Brussel M, de Groot M. Do activity monitors increase physical activity in adults with overweight or obesity? A systematic review and meta-analysis. *Obesity* 2016; **24:** 2078–91.

42 Franssen WMA, Franssen GHLM, Spaas J, Solmi F, Eijnde BO. Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. *Int J Behav Nutr Phys Act* 2020; **17:** 57.

43 Freak-Poli R, Cumpston M, Albarqouni L, Clemes SA, Peeters A. Workplace pedometer interventions for increasing physical activity. *Cochrane Database Syst Rev* 2020; **7:** CD009209.

44 Funk M, Taylor EL. Pedometer-based walking interventions for free-living adults with type 2 diabetes: a systematic review. *Curr Diabetes Rev* 2013; **9:** 462–71.

45 Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. *Sports Med Open* 2018; **4:** 42.

46 Goode AP, Hall KS, Batch BC, et al. The impact of interventions that integrate accelerometers on physical activity and weight loss: a systematic review. *Ann Behav Med* 2017; **51:** 79–93.

47 Hannan AL, Harders MP, Hing W, Climstein M, Coombes JS, Furness J. Impact of wearable physical activity monitoring devices with exercise prescription or advice in the maintenance phase of cardiac rehabilitation: systematic review and meta-analysis. *BMC Sports Sci Med Rehabil* 2019; **11:** 14.

48 Jo A, Coronel BD, Coakes CE, Mainous AG 3rd. Is there a benefit to patients using wearable devices such as Fitbit or health apps on mobiles? A systematic review. *Am J Med* 2019; **132:** 1394–400.e1.

49 Kang M, Marshall SJ, Barreira TV, Lee JO. Effect of pedometer-based physical activity interventions: a meta-analysis. *Res Q Exerc Sport* 2009; **80:** 648–55.

50 Kirk MA, Amiri M, Pirbaglou M, Ritvo P. Wearable technology and physical activity behavior change in adults with chronic cardiometabolic disease: a systematic review and meta-analysis. *Am J Health Promot* 2019; **33:** 778–91.

51 Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance amount of physical activity in older adults: a systematic review and meta-analysis. *Eur Rev Aging Phys Act* 2019; **16:** 7.

52 Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J. Using an electronic activity monitor system as an intervention modality: a systematic review. *BMC Public Health* 2015; **15:** 585.

53 Liu JY, Kor PP, Chan CP, Kwan RY, Sze-Ki D. The effectiveness of a wearable activity tracker (WAT)-based intervention to improve physical activity levels in sedentary older adults: a systematic review and meta-analysis. *Arch Gerontol Geriatr* 2020; **91:** 104211.

54 Lynch EA, Jones TM, Simpson DB, et al. Activity monitors for increasing physical activity in adult stroke survivors. *Cochrane Database Syst Rev* 2018; **7:** CD012543.

55 Lynch C, Bird S, Lythgo N, Selva-Raj I. Changing the physical activity behavior of adults with fitness trackers: a systematic review and meta-analysis. *Am J Health Promot* 2020; **34:** 418–30.

56 Martínez-García MDM, Ruiz-Cárdenas JD, Rabinovich RA. Effectiveness of smartphone devices in promoting physical activity and exercise in patients with chronic obstructive pulmonary disease: a systematic review. *COPD* 2017; **14:** 543–51.

57 McDonough DJ, Su X, Gao Z. Health wearable devices for weight and BMI reduction in individuals with overweight/obesity and chronic comorbidities: systematic review and network meta-analysis. *Br J Sports Med* 2021; **55:** 917–25.

58 Oliveira CB, Franco MR, Maher CG, et al. Physical activity-based interventions using electronic feedback may be ineffective in reducing pain and disability in patients with chronic musculoskeletal pain: a systematic review with meta-analysis. *Arch Phys Med Rehabil* 2018; **99:** 1900–12.

59 Oliveira JS, Sherrington C, Zheng ERY, Franco MR, Tiedemann A. Effect of interventions using physical activity trackers on physical activity in people aged 60 years and over: a systematic review and meta-analysis. *Br J Sports Med* 2020; **54:** 1188–94.

60 Qiu S, Cai X, Chen X, Yang B, Sun Z. Step counter use in type 2 diabetes: a meta-analysis of randomized controlled trials. *BMC Med* 2014; **12:** 36.

61 Richardson CR, Newton TL, Abraham JJ, Sen A, Jimbo M, Swartz AM. A meta-analysis of pedometer-based walking interventions and weight loss. *Ann Fam Med* 2008; **6:** 69–77.

62 Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. *JMIR Mhealth Uhealth* 2016; **4:** e129.

63 Ringeval M, Wagner G, Denford J, Paré G, Kitsiou S. Fitbit-based interventions for healthy lifestyle outcomes: systematic review and meta-analysis. *J Med Internet Res* 2020; **22:** e23954.

64 Schaffer K, Panneerselvam N, Loh KP, et al. Systematic review of randomized controlled trials of exercise interventions using digital activity trackers in patients with cancer. *J Natl Compr Canc Netw* 2019; **17:** 57–63.

65 Sypes EE, Newton G, Lewis ZH. Investigating the use of an electronic activity monitor system as a component of physical activity and weight-loss interventions in nonclinical populations: a systematic review. *J Phys Act Health* 2019; **16:** 294–302.

66 Tang MSS, Moore K, McGavigan A, Clark RA, Ganesan AN. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2020; **8:** e15576.

67 Yen HY, Chiu HL. The effectiveness of wearable technologies as physical activity interventions in weight control: a systematic review and meta-analysis of randomized controlled trials. *Obes Rev* 2019; **20:** 1485–93.

68 Vaes AW, Cheung A, Atakhorrami M, et al. Effect of ‘activity monitor-based’ counseling on physical activity and health-related outcomes in patients with chronic diseases: a systematic review and meta-analysis. *Ann Med* 2013; **45:** 397–412.

69 Watson A, Timperio A, Brown H, Best K, Hesketh KD. Effect of classroom-based physical activity interventions on academic and physical activity outcomes: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2017; **14:** 114.

70 Conn VS, Hafdahl AR, Cooper PS, Brown LM, Lusk SL. Meta-analysis of workplace physical activity interventions. *Am J Prev Med* 2009; **37:** 330–39.

71 Kraus WE, Powell KE, Haskell WL, et al. Physical activity, all-cause and cardiovascular mortality, and cardiovascular disease. *Med Sci Sports Exerc* 2019; **51:** 1270–81.

72 Cumming G, Calin-Jageman R. Introduction to the new statistics: estimation, open science, and beyond. New York: Routledge, 2016.

73 Storman M, Storman D, Jasinska KW, Swierz MJ, Bala MMJOR. The quality of systematic reviews/meta-analyses published in the field of bariatrics: a cross-sectional systematic survey using AMSTAR 2 and ROBIS. *Obes Rev* 2020; **21:** e12994.

74 Almeida MO, Yamato TP, Parreira PdCS, Costa LOP, Kamper S, Saragiotto BTJBjopt. Overall confidence in the results of systematic reviews on exercise therapy for chronic low back pain: a cross-sectional analysis using the Assessing the Methodological Quality of Systematic Reviews (AMSTAR) 2 tool. *Braz J Phys Ther* 2020; **24:** 103–117.

75 Kolaski K, Romeiser Logan L, Goss KD, Butler C. Quality appraisal of systematic reviews of interventions for children with cerebral palsy reveals critically low confidence. *Dev Med Child Neurol* 2021; **63:** 1316–26.