Effectiveness of Wearable Trackers on Physical Activity in Healthy Adults: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Matilda Swee Sun Tang¹, BA, MD; Katherine Moore¹, BSc, PhD, MD; Andrew McGavigan¹,², MBChB, MD; Robyn A Clark³,⁴, RN, ICUcert, DipAppliSci, BN, MEd, PhD, ACCCN, FCNA, FAHA, FCSANZ; Anand N Ganesan¹,²,⁴, MBBS, PhD, FRACP, FCSANZ

¹College of Medicine and Public Health, Flinders University, Adelaide, Australia
²Department of Cardiovascular Medicine, Flinders Medical Centre, Adelaide, Australia
³College of Nursing and Health Sciences, Flinders University, Adelaide, Australia
⁴South Australian Health and Medical Research Institute, Adelaide, Australia

Corresponding Author:
Anand N Ganesan, MBBS, PhD, FRACP, FCSANZ
College of Medicine and Public Health
Flinders University
Level 5, Room 5E209 Flinders Medical Centre
Adelaide
Australia
Phone: 61 (08) 7221 8200
Email: anand.ganesan@flinders.edu.au

Abstract

Background: Wearable trackers are an increasingly popular tool among healthy adults and are used to facilitate self-monitoring of physical activity.

Objective: We aimed to systematically review the effectiveness of wearable trackers for improving physical activity and weight reduction among healthy adults.

Methods: This review used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology and reporting criteria. English-language randomized controlled trials with more than 20 participants from MEDLINE, CINAHL, Cochrane Library, Web of Science, PubMed, and Scopus (2000-2017) were identified. Studies were eligible for inclusion if they reported an intervention group using wearable trackers, reporting steps per day, total moderate-to-vigorous physical activity, physical activity, energy expenditure, and weight reduction.

Results: Twelve eligible studies with a total of 1693 participants met the inclusion criteria. The weighted average age was 40.7 years (95% CI 31.1-50.3), with 64.4% women. The mean intervention duration was 21.4 weeks (95% CI 6.1-36.7). The usage of wearable trackers was associated with increased physical activity (standardized mean difference 0.449, 95% CI 0.10-0.80; P=.01).

In the subgroup analyses, however, wearable trackers demonstrated no clear benefit for physical activity or weight reduction.

Conclusions: These data suggest that the use of wearable trackers in healthy adults may be associated with modest short-term increases in physical activity. Further data are required to determine if a sustained benefit is associated with wearable tracker usage.

(JMIR Mhealth Uhealth 2020;8(7):e15576) doi: 10.2196/15576

KEYWORDS
wearable activity tracker; physical activity; healthy adults; randomized controlled trials

Introduction

Wearable activity trackers have rapidly emerged in the past decade as consumer devices to support self-monitoring of physical activity [1,2]. The use of these devices has increased exponentially, and the global sales of wearables in health care are expected to reach US $4.4 billion in 2019 and US $4.5 billion by 2020 [3].
In the past, structured lifestyle interventions have utilized education with behavior change techniques, provision of written information materials, and telephone counseling in a series of combination and permutation [4,5]. These interventions are successful in the short term but not in the long term, and they tend to be labor-intensive and costly [2,6,7]. Today, the availability and accessibility of wearable trackers equip consumers with the ability to monitor their physical activity along with online applications with motivational and tracking tools. Several systematic reviews have shown that wearable trackers are effective [4,6,8].

Contemporary wearable trackers differ from conventional pedometers as they are sophisticated devices providing real-time multidimensional feedback on physiological and health parameters including steps, calories burned, distance covered, active time, sleep assessment, and heart rate, and may include mobile connectivity or an internet application to provide personalized feedback reports [9].

To date, most of the literature on wearable trackers has focused on their feasibility, validity, and reliability [10] with limited data on the impact using these devices has on improving physical activity [6,11]. The primary purpose of this study is to evaluate the effectiveness of wearable trackers and their impact on physical activity levels in healthy adult populations with secondary outcomes of weight change in overweight populations.

**Methods**

**Protocol and Registration**

The protocol for this study is registered under PROSPERO with registration number CRD42019131868.

**Eligibility Criteria, Information Sources, Search and Study Selection**

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [12] methods and reporting were used to perform a systematic literature search with a professional medical librarian (Multimedia Appendix 1) on English-language randomized controlled trials published between January 1, 2000, and August 1, 2017.

We considered English-language studies eligible for inclusion if they reported an intervention with at least one of the groups using wearable trackers to provide objective feedback on physical activity to the wearer, alone or in combination with other interventions to enhance physical activity. Only randomized controlled trials with more than 20 participants in the adult outpatient and community setting that reported a change in physical activity behavior (total steps, total activity, the proportion of participants at activity goal) were included. We are looking at the healthy population and therefore excluded studies that required participants to be hospitalized or confined to a research center, studies in disease populations, and obese populations. We excluded studies that were predominantly pedometer-based interventions since we were only considering the effect of wearable trackers.

**Data Collection Process and Data Items**

Two authors independently abstracted three categories of variables from each of the included studies, with differences resolved by consensus: intervention variables (intervention duration); participant variables; quality variables (method of blinding control participants to step counts, the use of validity- and reliability-tested wearable trackers, the extent of affordability of wearable trackers and the extent to which co-interventions may have affected physical activity). If the study reported results from a different period, we used the final immediate post-intervention data in our primary analysis. For studies that reported different intensities of physical activity instead of steps per day, we chose the walking intensity results for the primary analysis.

**Risk of Bias and Quality Assessment**

The risk of bias was assessed using the Cochrane risk of bias tool [13] across seven domains. Each domain was scored with low (L), unclear (U), or high (H) risk of bias. The domains assessed were as follows:

- **Random sequence generation**: Was there selective bias (biased allocation to interventions) due to inadequate generation of a randomized sequence?
- **Allocation concealment**: Was there selective bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment?
- **Selective outcome reporting**: Was there reporting bias due to knowledge of the allocated interventions by outcome assessors?
- **Blinding outcome assessment**: Was there detection bias due to knowledge of the allocated interventions by outcome assessors?
- **Blinding participants and personnel**: Was there performance bias due to knowledge of the allocated interventions by participants and personnel during the study?
- **Incomplete outcome data**: Was there knowledge of the allocated interventions by outcome assessors?
- **Other sources of bias**: Was there bias due to problems not covered elsewhere?

The GRADE (Grades of Recommendation, Assessment, Development, and Evaluation) system was used to rank the quality of evidence for each study [14]. The GRADE approach uses five considerations (study limitations, consistency of effect, imprecision, indirectness, and publication bias). The criteria for the grade of evidence are as follows:

- **High**: We are very confident that the true effect lies close to that of the estimated effect.
- **Moderate**: We are moderately confident in the effect estimate; the true effect is likely close to the estimated effect, but there is a possibility that it is substantially different
- **Low**: Our confidence in the effect estimate is limited; the true effect may be substantially different from the estimated effect.
- **Very low quality**: We have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect.
Synthesis of Results

Statistical analysis was performed with Comprehensive Meta-Analysis Version 3 [15]. For each of the included studies, we calculated the effect sizes of physical activity for our primary interest outcome using steps/day, total moderate-vigorous physical activity (MVPA), activity units or physical activity energy expenditure, depending on the available results from the studies with steps/day taking priority. We calculated the summary outcomes or the standardized mean difference (SMD) (95% CI) using random-effects calculations. SMD is used since the included studies all assess the same outcome, physical activity, but they measure it in a variety of ways (steps/day, MVPA, energy expenditure). Hence, differences in means that are the same proportion of the standard deviation will have the same SMD, regardless of the actual scales used to make the measurements [16].

The I² statistic was used as a measure of variability in observed effects estimates attributable to between-study heterogeneity [17]. For variables exhibiting mild heterogeneity (I²≤25%), pooled estimates were derived with fixed effects models. For variables exhibiting more than moderate heterogeneity (I²>25%), pooled estimates were derived with random-effects models. Sub-group analysis was used to assess secondary outcomes.

Results

Study Selection

The primary search identified a total of 9591 studies from MEDLINE (n=1910), CINAHL (n=62), Cochrane Library (n=2871), Web of Science (n=1633), PubMed (n=1280), and Scopus (n=1835), as detailed in Multimedia Appendix 1. There were 4620 unique citations in which 103 full texts were retrieved for further review after applying the Population, Intervention, Comparison, Outcome, and Study type [18] criteria of inclusion and exclusion studies (Multimedia Appendix 2) at the title-and-abstract screening level. A total of 12 unique papers were retained for data abstraction (Figure 1) [19].

Figure 1. PRISMA Flow Chart.

![PRISMA Flow Chart](https://example.com/prisma_flow_chart.png)
Study Characteristics

Characteristics of the included studies are presented in Table 1, and the intervention and comparators are presented in Table 2. A total of 1693 participants were included from 12 randomized trials. Included studies were published from 2007 to 2017.

Table 1. Characteristics of participants in the included studies.

<table>
<thead>
<tr>
<th>Study name, publication year</th>
<th>Sample size</th>
<th>Participant characteristics</th>
<th>Health status</th>
<th>Mean age (years)</th>
<th>Proportion</th>
<th>Study duration (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashe, 2016 [20]</td>
<td>25</td>
<td>Healthy participants</td>
<td></td>
<td>40.1</td>
<td>Women: 100%</td>
<td>26</td>
</tr>
<tr>
<td>Buis, 2017 [21]</td>
<td>40</td>
<td>Healthy overweight adults</td>
<td></td>
<td>61.5</td>
<td>Women: 66%; white ethnicity: 66%</td>
<td>12</td>
</tr>
<tr>
<td>Cadmus-Bertram, 2015 [22]</td>
<td>51</td>
<td>Healthy and overweight participants</td>
<td></td>
<td>60.0</td>
<td>White ethnicity: 92%</td>
<td>16</td>
</tr>
<tr>
<td>Godino, 2013 [23]</td>
<td>466</td>
<td>Healthy participants</td>
<td></td>
<td>47.7</td>
<td>Women: 46%; white ethnicity: 96%</td>
<td>8</td>
</tr>
<tr>
<td>Jakicic, 2016 [25]</td>
<td>471</td>
<td>Healthy overweight adults</td>
<td></td>
<td>30.9</td>
<td>Women: 71%; white ethnicity: 77%</td>
<td>104</td>
</tr>
<tr>
<td>Martin, 2015 [26]</td>
<td>49</td>
<td>Healthy overweight adults</td>
<td></td>
<td>58.0</td>
<td>Women: 46%; white ethnicity: 79%</td>
<td>5</td>
</tr>
<tr>
<td>Melton, 2016 [27]</td>
<td>69</td>
<td>Healthy participants</td>
<td></td>
<td>19.7</td>
<td>Women: 100%; black ethnicity: 100%</td>
<td>6</td>
</tr>
<tr>
<td>Poirier, 2007 [28]</td>
<td>264</td>
<td>Healthy participants</td>
<td></td>
<td>39.9</td>
<td>Women: 66%; white ethnicity: 77%</td>
<td>7</td>
</tr>
<tr>
<td>Shrestha, 2013 [29]</td>
<td>28</td>
<td>Healthy overweight adults</td>
<td></td>
<td>32.1</td>
<td>Women: 54%</td>
<td>26</td>
</tr>
<tr>
<td>Thompson, 2014 [30]</td>
<td>49</td>
<td>Healthy participants</td>
<td></td>
<td>79.1</td>
<td>Women: 91%; white ethnicity: 66%</td>
<td>26</td>
</tr>
<tr>
<td>Thorndike, 2014 [31]</td>
<td>104</td>
<td>Healthy participants</td>
<td></td>
<td>29.0</td>
<td>Women: 54%</td>
<td>12</td>
</tr>
</tbody>
</table>

Study sample sizes varied from 21 to 471 participants. The participants’ weighted average age was 40.7 years (95% CI 31.1-50.3), and 64.4% of the participants were women. The duration of interventions ranged from 6 to 104 weeks, with a mean intervention duration of 21.4 weeks (95% CI 6.1-36.7).
<table>
<thead>
<tr>
<th>Study Name, Year</th>
<th>Intervention Device</th>
<th>Intervention</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashe, 2016 [20]</td>
<td>Fitbit</td>
<td>26 weeks of group-based education, social support, individualized physical activity prescription, given Fitbit</td>
<td>26 weeks: only received health-related information</td>
</tr>
<tr>
<td>Buis, 2017 [21]</td>
<td>Jawbone Up24</td>
<td>Received a Jawbone Up24 monitor, a tablet with Jawbone Up app installed, and brief weekly telephone counseling</td>
<td>Waitlist control (did not receive any intervention until after their final assessment where they were provided the intervention in full)</td>
</tr>
<tr>
<td>Cadmus-Bertram, 2015 [22]</td>
<td>Fitbit</td>
<td>16 weeks of Web-Based Tracking Group: Fitbit, instructional session, follow-up call at the fourth week</td>
<td>16 weeks of standard pedometer</td>
</tr>
<tr>
<td>Godino, 2013 [23]</td>
<td>Combined HR monitor and accelerometer (Actiheart)</td>
<td>8 weeks of wearing of Actiheart with one of three different types of feedback (simple, visual, contextualized)</td>
<td>8 weeks of wearing of Actiheart but with no feedback until the end of the trial</td>
</tr>
<tr>
<td>Hurling, 2007 [24]</td>
<td>Wrist-worn accelerometer</td>
<td>9 weeks of wristworn accelerometer, weekly exercise schedule, email reminders, real-time feedback via the internet</td>
<td>9 weeks of wrist-worn accelerometer with no feedback</td>
</tr>
</tbody>
</table>
| Martin, 2015 [26] | Fitbug Orb | 3-arm study
- Phase 1 (1 week): blinded run-in
- Phase 2 (2 weeks): unblinded versus blinded tracking
- Phase 3 (2 weeks): smart texts versus no texts | Blinded participants with no feedback |
| Melton, 2016 [27] | Jawbone UP | 6 weeks of wearing Jawbone UP band and engaging with the application with weekly reminders | 6 weeks of using MyFitnessPal application |
| Poirier, 2007 [28] | Variety of activity trackers | 2-arm study
- 6 weeks of walking program, Walkadoo, and wireless activity tracker
- 1 week of follow-up with wearing of activity tracker for at least 10 hours a day | 2-arm study
- 6 weeks of not wearing activity trackers and maintaining daily activity routine
- 1 week of follow-up wearing of activity tracker for 10 hours a day |
| Shrestha, 2013 [29] | Polar FA20 accelerometer | 1 time 1.5-hour lifestyle instruction and 26 weeks of continuous accelerometer use and feedback | 26 weeks of self-directed exercise and/or US Army mandated physical training |
| Thompson, 2014 [30] | Fitbit | 26 weeks of accelerometer use and feedback, weekly brief telephone counseling sessions focused on accelerometer feedback, 6 in-person brief counseling sessions | 26 weeks of accelerometer without feedback |
| Thorndike, 2014 [31] | Fitbit e3 | 2-arm study
- Phase 1: 6 weeks RCT\textsuperscript{a} comparing daily steps displaying feedback about steps and energy consumed
- Phase 2: 6 weeks non-RCT team steps competition where all participants wore monitors with feedback | 2-arm study
- Phase 1: 6 weeks blinded monitor
- Phase 2: 6 weeks non-RCT team steps competition where all participants wore monitor with feedback |

\textsuperscript{a}RCT: randomized controlled trial.

**Impact of Wearable Tracker Use and Physical Activity**

The primary outcome for this review was the impact of Wearable Tracker Use and Physical Activity. The overall summary estimate from 12 studies [20-31] showed a modest increase in physical activity with the usage of wearable trackers (SMD 0.449, 95% CI 0.10-0.80; \( P = .01 \)). There was significant heterogeneity (\( I^2 = 88\% \)) (Figure 2). Subgroup analyses were performed for studies using steps/day or weight as reported endpoints, and in healthy versus overweight populations to explore mechanisms of heterogeneity.

We performed subgroup analyses and assessed heterogeneity to evaluate the robustness of our results [32].
Impact of Wearable Tracker Use in Studies With Steps/Day as Primary Outcome Variable

A total of 7 [20-22,26-28,31] of 12 studies utilized steps/day as the primary endpoint. In these studies, steps/day were significantly increased by the end intervention (SMD 0.332, 95% CI 0.16-0.50; \( P < 0.001 \)). Heterogeneity in this analysis was low (\( I^2 = 2\% \)) (Figure 3).

<table>
<thead>
<tr>
<th>Study name</th>
<th>Statistics for each study</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashe, 2016</td>
<td>0.959 (0.049, 1.869)</td>
<td>7, 12</td>
</tr>
<tr>
<td>Bius, 2017</td>
<td>0.563 (0.049, 1.175)</td>
<td>20, 20</td>
</tr>
<tr>
<td>Cadmus-Bertram, 2015</td>
<td>0.197 (-0.339, 0.733)</td>
<td>25, 25</td>
</tr>
<tr>
<td>Godine, 2013</td>
<td>-0.068 (-0.345, 0.209)</td>
<td>60, 99</td>
</tr>
<tr>
<td>Hulting, 2007</td>
<td>2.272 (1.699, 2.845)</td>
<td>20, 47</td>
</tr>
<tr>
<td>Jakelios, 2016</td>
<td>-0.318 (-0.499, -0.136)</td>
<td>233, 201</td>
</tr>
<tr>
<td>Martin, 2015</td>
<td>0.394 (-0.196, 0.983)</td>
<td>19, 32</td>
</tr>
<tr>
<td>Melton, 2016</td>
<td>0.584 (0.103, 1.066)</td>
<td>89, 99</td>
</tr>
<tr>
<td>Poirier, 2007</td>
<td>0.320 (0.054, 0.586)</td>
<td>110, 107</td>
</tr>
<tr>
<td>Shresha, 2013</td>
<td>0.628 (-0.213, 1.469)</td>
<td>11, 9</td>
</tr>
<tr>
<td>Thompson, 2014</td>
<td>0.370 (-0.185, 0.925)</td>
<td>24, 24</td>
</tr>
<tr>
<td>Thordike, 2014</td>
<td>0.024 (-0.365, 0.413)</td>
<td>91, 49</td>
</tr>
<tr>
<td></td>
<td>0.449 (0.102, 0.796)</td>
<td>50, 50</td>
</tr>
</tbody>
</table>

Heterogeneity: \( \tau^2 = 0.30; \) df(\( Q \)) = 11 (\( P = 0.00 \)); \( I^2 = 88\% \)
Test for overall effect: \( Z = 2.53 (P = 0.01) \)

Impact of Wearable Tracker Use in Studies With Weight Loss as an Outcome

A total of 4 [22,25,29,30] of 12 studies reported weight change. However, no significant effect on weight change was observed (SMD 0.133, 95% CI –0.34 to 0.60, \( P = .58 \)). Heterogeneity in this analysis was low (\( I^2 = 0\% \)) (Figure 4).

<table>
<thead>
<tr>
<th>Study name</th>
<th>Statistics for each study</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashe, 2016</td>
<td>0.959 (0.049, 1.869)</td>
<td>7, 12</td>
</tr>
<tr>
<td>Bius, 2017</td>
<td>0.563 (0.049, 1.175)</td>
<td>20, 20</td>
</tr>
<tr>
<td>Cadmus-Bertram, 2015</td>
<td>0.197 (-0.339, 0.733)</td>
<td>25, 25</td>
</tr>
<tr>
<td>Godine, 2013</td>
<td>-0.068 (-0.345, 0.209)</td>
<td>60, 99</td>
</tr>
<tr>
<td>Hulting, 2007</td>
<td>2.272 (1.699, 2.845)</td>
<td>20, 47</td>
</tr>
<tr>
<td>Jakelios, 2016</td>
<td>-0.318 (-0.499, -0.136)</td>
<td>233, 201</td>
</tr>
<tr>
<td>Martin, 2015</td>
<td>0.394 (-0.196, 0.983)</td>
<td>19, 32</td>
</tr>
<tr>
<td>Melton, 2016</td>
<td>0.584 (0.103, 1.066)</td>
<td>89, 99</td>
</tr>
<tr>
<td>Poirier, 2007</td>
<td>0.320 (0.054, 0.586)</td>
<td>110, 107</td>
</tr>
<tr>
<td>Shresha, 2013</td>
<td>0.628 (-0.213, 1.469)</td>
<td>11, 9</td>
</tr>
<tr>
<td>Thompson, 2014</td>
<td>0.370 (-0.185, 0.925)</td>
<td>24, 24</td>
</tr>
<tr>
<td>Thordike, 2014</td>
<td>0.024 (-0.365, 0.413)</td>
<td>91, 49</td>
</tr>
<tr>
<td></td>
<td>0.449 (0.102, 0.796)</td>
<td>50, 50</td>
</tr>
</tbody>
</table>

Heterogeneity: \( \tau^2 = 0.001; \) df(\( Q \)) = 6 (\( P = 0.41 \)); \( I^2 = 2\% \)
Test for overall effect: \( Z = 3.80 (P < 0.001) \)
A total of 5 [21,22,25,26,29] of 12 studies reported physical activity outcomes in overweight adults. In these studies, no significant increase in physical activity occurred (SMD 0.225, 95% CI –0.23 to 0.68, \( P = .33 \)). Heterogeneity in this analysis was high (\( I^2 = 76\% \)) (Figure 5). Seven [20,23,24,27,28,30,31] out of 12 studies reported physical activity outcomes in healthy adults unselected by weight. In these studies, a significant increase in physical activity was observed (SMD 0.594, 95% CI 0.10-1.09; \( P = .02 \)) Heterogeneity in this analysis was high (\( I^2 = 90\% \); Figure 6).

### Impact of Wearable Tracker Use in Studies on Overweight Reduction

A total of 5 [21,22,25,26,29] of 12 studies reported physical activity outcomes in overweight adults. In these studies, no significant increase in physical activity occurred (SMD 0.225, 95% CI –0.23 to 0.68, \( P = .33 \)). Heterogeneity in this analysis was high (\( I^2 = 76\% \)) (Figure 5). Seven [20,23,24,27,28,30,31] out of 12 studies reported physical activity outcomes in healthy adults unselected by weight. In these studies, a significant increase in physical activity was observed (SMD 0.594, 95% CI 0.10-1.09; \( P = .02 \)) Heterogeneity in this analysis was high (\( I^2 = 90\% \); Figure 6).

### Figures

**Figure 4.** Forest plot of standardized mean difference (95% CI) in the effect of wearable trackers on weight loss for intervention and control group.

**Figure 5.** Forest plot of standardized mean difference (95% CI) in the effect of wearable trackers on physical activity in overweight adults.

**Figure 6.** Forest plot of standardized mean difference (95% CI) in the effect of wearable trackers on healthy adults.
Synthesis of Results

A GRADE summary of results is presented in Figure 7. We assessed outcomes using the Cochrane GRADE approach. Results were downgraded when there was serious risk of bias, inconsistency, indirectness, imprecision, upgrading of a large effect size, or a dose-response gradient, all of which are possible confounding effects. Such confounding effects may create the appearance of an effect when there is none or reduce the appearance of an effect [33]. A GRADE Summary of Evidence table is provided in Multimedia Appendix 4.

![Figure 7. GRADE Working Group grades of evidence summary.](image)

### Risk of Bias Across Results

Risk of bias (ROB) is measured using the Cochrane risk of bias tool [13] for randomized controlled trials, and the summary of the evaluation is shown in Table 3. Selection bias was low as randomization was considered high in all of the studies. In 2/12 studies [20,28], methods for allocation concealment were described in insufficient details resulting in high ROB, and one [21] of the studies showed an unclear ROB. For the outcome of physical activity, the blinding of participants and personnel was highly variable as there are challenges to blinding physical activity interventions.

Most of the trials (9/12, 75%) [21-26,29-31] reported details of the measured outcomes that were sufficient. Reporting bias was judged to be at low ROB because most of the trials (9/12, 75%) [21-26,29-31] reported sufficient details and judged to be low ROB [26,29,31].

Detection bias was judged to be low since most of the trials (10/12, 83%) [20-30] provided sufficient information regarding outcome blinding assessment, and the remaining trials (16%) provided insufficient information.

Attrition bias was assessed to be low since all the trials reported the numbers reported to each group. The majority of trials (11/12, 92%) [20,21,23-31] included information on attrition and exclusions from the analysis. One trial [22] did not disclose the reason for attrition/exclusion in sufficient detail, resulting in a judgment of high ROB.

Of the 12 studies measuring physical activity as an outcome, only 2 were judged to be of high ROB in terms of performance bias due to lack of blinding, and 1 study showed inadequate information regarding blinding.
Overall ROB was assessed, and the majority of studies (9/12, 75%) \cite{21,23-30} were judged to be at low ROB and the remaining studies \cite{20,22,31} were judged at unclear ROB.

Publication bias was determined by visual inspection of funnel plots comparing physical activity against effect size. There was visual evidence of publication bias with at least two studies falling outside the range of expected precision for their effect size (Multimedia Appendix 3).

**Table 3.** Risk of bias (Cochrane Critical Appraisal Skills Program Tool\textsuperscript{a}).

<table>
<thead>
<tr>
<th>Study</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Total score/7</th>
<th>Quality Score, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashe, 2016 \cite{20}</td>
<td>L\textsuperscript{b}</td>
<td>L</td>
<td>H\textsuperscript{c}</td>
<td>U\textsuperscript{d}</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>4.5</td>
<td>64</td>
</tr>
<tr>
<td>Buis, 2017 \cite{21}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>5.5</td>
<td>79</td>
</tr>
<tr>
<td>Cadmus-Bertram, 2015 \cite{22}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>4.5</td>
<td>64</td>
</tr>
<tr>
<td>Godino, 2013 \cite{23}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>6.5</td>
<td>93</td>
</tr>
<tr>
<td>Hurling, 2007 \cite{24}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>5.5</td>
<td>79</td>
</tr>
<tr>
<td>Jakicic, 2016 \cite{25}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>6.5</td>
<td>93</td>
</tr>
<tr>
<td>Martin, 2015 \cite{26}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>5.5</td>
<td>79</td>
</tr>
<tr>
<td>Melton, 2016 \cite{27}</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>6.5</td>
<td>93</td>
</tr>
<tr>
<td>Poirier, 2007 \cite{28}</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>5.5</td>
<td>79</td>
</tr>
<tr>
<td>Shrestha, 2013 \cite{29}</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>6</td>
<td>86</td>
</tr>
<tr>
<td>Thomdike, 2014 \cite{31}</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>5</td>
<td>71</td>
</tr>
<tr>
<td>Thompson, 2014 \cite{30}</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>6</td>
<td>86</td>
</tr>
<tr>
<td><strong>Category Score (%)</strong></td>
<td>100</td>
<td>88</td>
<td>79</td>
<td>45</td>
<td>54</td>
<td>88</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}Cochrane risk of bias tool. Q1: Were there selection bias (biased allocation to interventions) due to inadequate generation of a randomized sequence? Q2: Were there selective bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment? Q3: Were there reporting bias due to selective outcome report? Q4: Were there bias due to problems not covered elsewhere in the table? Q5: Were there performance bias due to knowledge of the allocated interventions by participants and personnel during the study? Q6: Were there detection bias due to knowledge of the allocated interventions by outcome assessors? Q7: Were there attribution bias due to amount, nature, or handling or incomplete outcome data?

\textsuperscript{b}L: low risk.

\textsuperscript{c}H: high risk.

\textsuperscript{d}U: unclear risk.

**Discussion**

This meta-analysis examined the effects of wearable trackers on physical activity and is based on 12 randomized controlled trials involving 1693 participants. Overall, wearable tracker usage was associated with improvements in physical activity (SMD 0.594, 95% CI 0.10-1.09; \(P=0.018\)). Interventions that included a consumer-based wearable tracker demonstrated an improvement in physical activity as compared to control groups, especially with daily steps. No clear evidence of benefit was seen from wearable tracker use in the endpoints of weight reduction or physical activity of overweight populations. Indeed, a recent review showed a potential increase in physical activity but no evidence for its effectiveness in weight loss \cite{34}.

**Public Health Implications**

This data may have both individual and public health implications. Wearable trackers can support continuous health monitoring at both the individual and the population level. Wearable trackers are activity monitoring tools that help to engage patients as advocates in their personalized care and have been proposed to encourage healthy behavior. Benefits are thought to include prevention or reduction of health problems, support of chronic disease self-management, enhanced provider knowledge, reduced number of healthcare visits, and personalized, localized, and on-demand interventions in ways not previously possible \cite{35}. The low cost of delivery and the feasibility of wearable trackers makes them an attractive potential tool to facilitate self-monitoring of physical activity. The data presented in this meta-analysis demonstrates that wearable tracker usage was associated with short-term gains in physical activity \cite{36}.

A wide range of wearable trackers was used. Four \cite{20,22,30,31} of 12 studies used Fitbit, two \cite{21,27} used Jawbone Up, and the remaining studies used a variety of other wearable trackers. The relative proportion of commercial wearable trackers in the included studies is similar to global market shares, with Fitbit having the largest market share (20%) and hence applicable to the real world \cite{37}. The Fitbit and Jawbone Up used in this meta-analysis have the same selected measures of “steps, distance, calories, and sleep” and are worn on the wrist \cite{21,22,27,30,31}. A systematic review assessing the validity and reliability of Fitbit and Jawbone found that the validity and inter-device reliability of steps counts was generally high \cite{10}.
Study Limitations
This study has several limitations. First, most studies included in this meta-analysis included small study sizes with short intervention durations and limited follow-up, highlighting the necessity for longer-term studies. The variety of study designs may have been reflected by the statistical heterogeneity of the outcomes. Hence, it might be challenging to generalize the results due to heterogeneity. Future research should explore the long-term effectiveness of wearable trackers in increasing physical activity. Second, statistical estimates of publication bias identified individual, small studies with relatively large effect sizes, which may be a reflection of a file drawer effect. Third, it is difficult to establish the independent contribution of conjunctive interventions (eg, behavioral counseling, interactive health coach, weekly reminders, and text messages) that were often offered alongside wearable tracker usage. The last study was obtained in 2017, since which time there have been changes in market share due to the volatile nature of the industry. The entrance of new brands like Apple, Xiaomi, and Samsung means that an updated review is needed. Next, the skewed demographics to white females would mean that the translation to other demographics might be limited. Lastly, we restricted our search to only full-text published articles in the English language and, thus, may have excluded relevant studies outside our current scope.

Recommendations for Further Research
The current study examined the utility of wearable activity trackers in healthy adults and found a modest but measurable benefit in the population studied. Important questions for future research could include the identification of appropriate deployment strategies for these novel technologies in cardiac rehabilitation, aged care, and youth. A related question that we could not address was the role of social engagement in modulating the response of participants to wearable tracking devices.

Recommendations for Clinical Practice
The current study provides qualified support for the use of wearable activity trackers in healthy populations, showing evidence of short-term gains in physical activity, but not weight. As technology advances and these devices improve over time, future studies will be necessary to delineate the optimal use case for these devices.

Conclusion
In conclusion, this meta-analysis demonstrates the efficacy of wearable trackers in facilitating short-term increases in consumer physical activity. Future studies will be required to determine the durability of the influence of wearable tracker use on consumer physical activity behavior.

Acknowledgments
AG is supported by a Future Leader Fellowship of the National Heart Foundation of Australia (111086) and Prof Robyn A Clark Heart Foundation Future Leader Fellowship (APP ID. 100847).

Conflicts of Interest
None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File , 17 KB-Multimedia Appendix 1]

Multimedia Appendix 2
PICOS criteria for inclusion and exclusion of studies.
[DOCX File , 13 KB-Multimedia Appendix 2]

Multimedia Appendix 3
Funnel plot showing publication bias.
[DOCX File , 35 KB-Multimedia Appendix 3]

Multimedia Appendix 4
GRADE Summary of Evidence Table.
[DOCX File , 16 KB-Multimedia Appendix 4]

References


22. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and Acceptability of a Wearable Technology Physical Activity Intervention With Telephone Counseling for Mid-Aged and Older Adults: A Randomized Controlled Pilot Trial. JMRI Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]


33. Ryan R, Hill S. How to GRADE. How to GRADE the quality of the evidence Cochrane Consumers and Communication 2018 Jul 25 [FREE Full text] [doi: 10.26181/5b5795632a2c]


Abbreviations
- GRADE: Grades of Recommendation, Assessment, Development, and Evaluation
- MVPA: moderate-vigorous physical activity
- RCT: randomized controlled trial
- ROB: risk of bias
- SMD: standardized mean difference

Edited by G Eysenbach; submitted 22.07.19; peer-reviewed by A Henriksen, A Rickman, L Toth, S Haghayegh; comments to author 30.09.19; revised version received 18.11.19; accepted 14.05.20; published 22.07.20

Please cite as:
Tang et al. Effectiveness of Wearable Trackers on Physical Activity in Healthy Adults: Systematic Review and Meta-Analysis of Randomized Controlled Trials
JMIR Mhealth Uhealth 2020;8(7):e15576
URL: http://mhealth.jmir.org/2020/7/e15576/
doi: 10.2196/15576
PMID: 32706685