This is an Accepted Manuscript of an article published by Taylor & Francis in Health Psychology Review on 28th April 2025 available at: https://doi.org/10.1080/17437199.2025.2492042

Promoting Physical Activity During Retirement Age with Psychological Components: Multilevel Meta-analysis

Paula Collazo-Castiñeira[1,2*] (pcollazo@comillas.edu; 0000-0003-0543-1822), Ignacio Echegoyen[1] (iechegoyen@comillas.edu; 0000-0001-8016-7836), Josje Schoufour[3](j.d.schoufour@hva.nl), Noelia Álvarez-Díaz[4] (nalvarezd@salud.madrid.org; 0000-0002-5904-3609), Juan Botella[5] (juan.botella@uam.es; 0000-0001-8633-4981), Katarzyna janiszewska[6] (katarzyna@janiszewska.com; 0000-0002-4471-7543), Cristina Jiménez-Domínguez[2](cjdominguez@salud.madrid.org; 0009-0009-1165-7407), Vicente Lavilla[2] (vicentelavillag@gmail.com; 0000-0002-4303-8039), Ana López-Iglesias[2](alopeziglesias@salud.madrid.org; 0009-0003-07173625), Manuel V. Mejía Ramírez-Arellano[2] (manuelvicente.mejia@salud.madrid.org; 0000-0002-4647-0177), Amalia Tsagari [7,8] (tsagaridiet@yahoo.gr), Marije H. Verwijs[3](verwijs@voedingscentrum.nl; 0000-0002-3137-8757), Alfonso J. Cruz-Jentoft[2] (alfonsojose.cruz@salud.madrid.org; 0000-0001-7628-4861), Macarena Sánchez-Izquierdo[1](msizquierdo@comillas.edu; 0000-0002-6444-0545)

[1] Department of Psychology, Universidad Pontificia Comillas, Spain

[2] Servicio de Geriatría, Hospital Universitario Ramón y Cajal (IRYCIS), Spain

[3] Department of Nutrition and Dietetics, Faculty of Health, Sport and Physical Activity, Amsterdam University of Applied Sciences, Amsterdam, Netherlands

[4] Medical Library, Hospital Universitario Ramón y Cajal (IRYCIS), Spain

[5]Department of Methodology, Universidad Autónoma de Madrid, Spain

[6]European Specialist Dietetic Network for Sports and Physical Activity, The European Federation of the Associations of Dietitians, Netherlands

[7]European Federation of the Associations of Dietitian's European Specialist Dietetic Network for Older Adults

[8]Department of Clinical Nutrition, KAT General Hospital, Greece

*Corresponding author: Paula Collazo-Castiñeira, Department of Psychology, Comillas Pontifical University, C. Universidad Comillas, 3-5, 28049 Madrid, Spain.
Email: pcollazo@comillas.edu

This is an Accepted Manuscript of an article published by Taylor & Francis in Health Psychology Review on 28th April 2025 available at: https://doi.org/10.1080/17437199.2025.2492042

Promoting Physical Activity During Retirement Age with Psychological Components: Multilevel Meta-analysis

Abstract

Background. Physical activity plays a critical role in preventing and treating age-related chronic diseases and mortality. During retirement, individuals restructure their routines and behaviors, becoming a perfect period to promote and establish health behaviors. Identifying effective intervention components to promote health behaviors during this period is important for future intervention designs. **Purpose.** We aimed to identify which intervention components and behavior change techniques predicted effectiveness in interventions promoting physical activity. **Methods.** We conducted a meta-analysis using four databases (PsycInfo, Pubmed, Web of Science, and Scopus). We included studies reporting randomized controlled trials that (p)targeted retirement-age adults (50-70 years old), (i)intervened with behavior change techniques, (c)with any comparator, and (o)promoted physical activity. Screening, full-text, and extraction were done independently by at least two reviewers. We fitted a multilevel random effects model, reporting three effect sizes, and a meta-regression, testing several moderators. **Results.** 67 studies were included for meta-analysis (N=12,147). However, a high risk of bias predicted larger effect sizes. Thus, these studies were excluded from the main analysis. While individual studies often showed non-significant results, the pooled overall effect was small yet statistically significant across all effect sizes. Significant predictors varied across effect sizes, including action planning, motivational interviewing, and prompts/cues. Furthermore, the use of emails and a website to deliver the intervention was related to lower effect sizes. **Conclusions.** The effectiveness of lifestyle interventions is heterogeneous and presented small effects; implementing action planning, motivational interviewing, and prompts could improve the effectiveness. However, many BCTs that are not frequently used remain unexplored.

Keywords: Retirement, exercise, physical activity, lifestyle, meta analysis, behavior change

This is an Accepted Manuscript of an article published by Taylor & Francis in Health Psychology Review on 28th April 2025 available at: https://doi.org/10.1080/17437199.2025.2492042

Promoting Physical Activity During Retirement Age with Psychological Components: Multilevel Meta-analysis

Introduction

The global population is experiencing a notable demographic shift towards aging, accompanied by an increase in age-related diseases that significantly impact autonomy, quality of life, and overall longevity [1]. In this regard, adopting preventive behaviors, particularly engaging in regular physical activity, emerges as an essential factor for promoting healthy aging and longevity [2–4]. The importance of being physically active intensifies with age, acting as a countermeasure against the usual decline associated with aging, including sensory, motor, and specific cognitive functions, as well as overall physical performance and intrinsic capacity [5,6]. However, older adults often fail to engage sufficiently in physical activity [2–4].

Retirement, a significant life transition, presents a unique window of opportunity for individuals to establish new health behaviors, as one's routine is about to change anyway. However, retirees' new routines can follow different trajectories. On one hand, research indicates that individuals going through this phase often exhibit openness and motivation to embrace healthier behaviors [7]. On the other hand, despite this potential, the inherent impact of retirement on health behaviors remains under debate, as some studies found detrimental changes in physical activity and diet, whereas others supported positive changes in health behaviors [8,9]. Therefore, this transition offers a strategic moment to promote health behaviors and prevent otherwise unhealthy trajectories. Thus, identifying effective intervention components for this aging population becomes essential.

Significant research efforts have been dedicated to developing and testing lifestyle interventions tailored to promote and sustain health behaviors in retirement age, such as physical activity [10]. However, these interventions encounter challenges that limit their effectiveness, including issues with adherence and long-term maintenance [11]. Notably, interventions explicitly grounded in psychological theoretical frameworks (i.e., where intervention techniques are selected based on theoretical constructs and mechanisms to target specific predictors of behavior) appear to be more successful in facilitating and sustaining behavior change [12]. Yet, attempts to identify effective psychological techniques (i.e. behavior change techniques; BCTs) and models, often through meta-analysis, have yielded conflicting results in the existing literature. In this regard, there is significant diversity in the BCTs that meta-analyses found to be significant moderators [13,14]. Potential reasons for this inconsistency include poorly described interventions failing to identify active ingredients [15], heterogeneity in the description of BCTs [16], and even its inadequate utilization [17,18]. Several initiatives have sought to address these challenges, such as the TIDieR checklist, designed to ensure an adequate description of the interventions [15], and taxonomies like Michie et al.'s [16], which encompasses 93 BCTs. Even though the use of these tools is growing in reviews and meta-analyses, interventions often neglect these guidelines. In this regard, it is important not only to test which BCTs and models elicit effective results but also to determine if quality measures (such as TIDieR or risk of bias) are influencing these results and possibly explaining the discrepancies found. Thus, these measures should also be considered as potential moderators in the intervention's success and considered when identifying active ingredients.

Ultimately, significant gaps remain in our understanding of the optimal strategies for promoting and maintaining health behaviors, especially during the retirement transition. Various reviews have explored physical activity interventions in adults of retirement age (e.g., Baxter et al. [10]). Yet, to our knowledge, none of them have undertaken quantitative analyses to discern the effective components. Among other reasons, this might be due to the inherent heterogeneity among the studies, for example, reporting on diverse outcome measures, which also leads to an issue of dependence that is usually overlooked when conducting a meta-analysis [10,19].

This meta-analysis sought to identify intervention components that predicted effectiveness in randomized controlled trials (RCTs) promoting physical activity in retirement-age adults. To address the mentioned issue of heterogeneity, we studied the role of several potential moderators (quality measures, type of physical activity, assessment tools, target population, delivery format, BCTs, theoretical framework, and interventionist). Furthermore, to use all the measures reported by primary studies and the different intervention arms, we followed a multi-level approach, accounting for the interdependence among the different observations presented in each study [19].

Method

This meta-analysis focused only on physical activity, following the protocol preregistered in PROSPERO (ID: blinded for review), and it was reported according to the Preferred Reporting Items for Systematic Reviews (PRISMA [20]; Supplementary Material 1, Appendix A).

Study Eligibility

Published studies in English, Spanish, Portuguese, Greek, or Dutch (as per the linguistic capabilities of the research team) that fulfilled the inclusion criteria based on the following PICO question scheme were included:

Population

Community-dwelling older adults around retirement age, defined as participants' mean age between 55 and 65 years old, or both mean age and standard deviations fell between the 50 and 70 years old range. Studies specifically targeting participants with serious mental or physical illnesses (e.g., COPD, active cancer) were excluded, as interventions designed for these populations are often tailored to their unique needs and may not be generalizable to the broader population of retirement-age adults [21]. However, studies targeting participants with common and prevalent chronic conditions among this age group, such as diabetes or cardiac conditions, were included to ensure the findings remained relevant to the general population. This criterion is common in studies assessing this population [22].

Intervention

Eligible interventions targeted the promotion of physical activity and included psychological components as active ingredients of the intervention. Lifestyle interventions frequently use psychological components (e.g., BCTs) unintentionally [23], we defined psychological components as specific techniques and strategies intentionally designed and implemented within the intervention to promote behavior change. Two psychologists independently evaluated each intervention, rating whether the intervention (1) purposefully implemented techniques to promote behavior change (e.g., action planning), or (2) was designed grounded on a psychological theory (e.g., Self-Determination Theory). Originally, interventions that were not sufficiently described were to be discarded; however, we opted later to include these interventions to statistically explore the influence of the level of description on the results.

Comparison

Any comparison group was deemed eligible, ranging from usual care to alternative versions of the intervention. Initially, we planned to exclude studies where all study arms shared the same psychological components, as traditional meta-analyses calculate effect sizes based on differences between intervention and control groups. However, we opted to include these studies by calculating additional effect sizes and using a multilevel model, which allowed us to account for the lack of variability between groups. Thus, studies sharing the same psychological components in their arms were just excluded from two analyses: standardized mean difference (SMD) and standardized mean change difference (SMCD).

Outcomes

Studies that pursued change in physical activity as their primary outcome were included. Any form of measure or type of physical activity was included (e.g., strength exercise, walking, reducing sedentary time, etc.) as long as it was a direct behavioral measure. Thus, indirect measures such as intentions or physiological values were discarded.

Study Designs

We included randomized controlled trials (RCTs) that performed randomization at the individual or stratified level. Cluster-randomized trials were excluded, as combining individual and cluster designs in meta-analyses of continuous outcomes may introduce bias due to differences in design and analysis requirements [24,25]. Cross-over trials, including step-wedge designs, were included only if post-intervention results for intervention and control groups were reported separately, ensuring the intervention effect could be accurately isolated.

Search Strategy

The search strategy was designed with the support of a librarian (XX; Supplementary Material 1, Appendix B) and was searched in four databases (PsycInfo, Pubmed, Scopus, and Web of Science) on July 1st, 2021. No restrictions based on language or dates were applied.

Study Selection

Records were managed with Covidence software (http://www.covidence.org) Once duplicates were removed, at least two reviewers screened titles and abstracts independently (XXX). Conflicts were resolved through discussion or with a third reviewer (XXX). Full-text screening was performed by two reviewers independently (XXX), and reasons for exclusion were registered (Figure 1). Conflicts were solved through discussion or a third reviewer (XXX).

Data Extraction

Data extraction was performed independently by two clinical psychologists (XXX) following a predesigned template and reaching a consensus through discussion. Data were extracted on studies' characteristics (title, authors, year, and country), sample demographics (sample size, mean age, gender, race, education, and diseases when they were inclusion criteria for studies), Intervention characteristics (Interventionist, duration, format, and delivery channel, theoretical framework, and BCTs following a standardized Taxonomy: V1 Michie et al [14]; an additional technique was coded as "motivational interviewing", as it was not sufficiently covered by the taxonomy [19]), outcome characteristics (Type of physical activity, unit of measure, tendency intended, and assessment tool). A dummy-coding strategy was followed for most variables, as 0= absence, 1= presence, as categories usually were not mutually exclusive (e.g., an intervention could use different BCTs and delivery formats). An additional code was used when the presence was common in both the intervention and control groups (Code=2). When there were missing details, the code "non-specified" was used. Furthermore, when there were unclear data (e.g., unclear whether a BCT was used), it was coded as absence, following a conservative approach. Furthermore, quantitative data were recorded for the metanalysis, extracting the sample size, mean, and standard deviation for all included outcomes, arms, and time points; if this was not presented, but other data were displayed (graphs, student's t-test), transformations were conducted when possible. When the primary outcomes could not be analyzed quantitatively, equivalent secondary outcomes were considered. We contacted the corresponding authors whenever data were unavailable (supplementary material, wrong tables, etc.). Only intention-to-treat analyses were included, excluding two additional studies.

Quality assessment

Two reviewers (XXX) independently assessed two quality measures, reaching a consensus through discussion. First, the TIDieR checklist was used to evaluate the description of the intervention and its replicability [15]. This checklist does not offer a decision tree for evaluating the level of description, it rather outlines what information needs to be described in each intervention component (e.g., procedure, interventionist, etc.). Thus, at least two reviewers rated each item and the overall intervention independently using the following categories: (1) not described (not mentioned in the study), (2) somewhat described (it is mentioned, but not sufficiently described), and (3) well described (it is described with enough detail). Second, the risk of bias for RCTs was assessed with RoB 2 [26].

Data Analysis

Analyses were conducted by a methodologist (XXX) using R [27] and the *metafor* package [28]. Three different types of effect sizes (ESs) were calculated when possible, using data from the first follow-up, namely the standardized mean difference (SMD; comparing control and intervention groups in the post-measurement), the standardized mean change (SMC; comparing the pre and post-measurement in the intervention groups), and the standardized mean change difference (SMCD; comparing the pre-post-measurements between intervention and control groups). Follow-up length was tested as a moderator but was not significant (data not shown). Thus, we did not conduct different analyses for short- and long-term results. More details on effect size estimation can be found in Supplementary Material 3, section A. In total, from the 67 studies, we obtained M = 160 realizations of the SMD (from N = 57 studies), M = 195 of the SMC (N = 65), and M = 124 of the SMCD (N = 41). Nonetheless, effect sizes coming from the same study were mutually dependent (procedure, experimenters, and subjects are the same). Hence, to analyze the effect of the intervention and its particular characteristics (covariates and BCTs) on the effect sizes, we fitted a multilevel random-effects model (Restricted Maximum Likelihood; REML), adjusting the level of correlation between ESs with a sensitivity analysis (see figure A-B.2 in the Supplementary Material 3) and a robust variance estimation (RVE) method to impute the studies' covariance matrices [29,30].

The model explicitly accounted for interdependence between ESs, extending the classical two-level model by adding a cluster (study) effect level (more details and the complete equations to replicate the model can be found in Supplementary Material 3, section B).

We then curated the database, removing studies meeting at least one of the following criteria: a) a high risk of bias (RoB); or b) a Cook's distance (CD) above 4/N (with N being the number of studies considered). The former came from the fact that a high RoB was a predictor of the ES when we added it to the model (SMD: $F = 3.393$; p-val = .036; SMC: $F = 4.846$; p-val = .009; SMCD: $F = 6.007$; p-val = .003). In particular, the difference between a high and a low RoB (SMD: β=-0.302; p-val = 0.018; SMC: β=-1.07; p-val = 0.006; SMCD: β=0.044; p-val = 0.914) indicated that studies with a high risk of bias tended to report significantly greater ESs (more details can be found in Supplementary Material 3, section C, and table A-C.1). The latter (a CD > 4/N) filtered for outliers in the sample, which can be visually observed in the forest, funnel, residuals, and Cook's distance plots (Supplementary Material 3, section C, figures A-C.1 to 6). To sum up, we removed 19 studies from the SMD effect size group (resulting in M = 123 estimates from N = 38 studies), 21 from the SMC (M = 158; N = 44), and 14 from the SMCD (M = 100; N = 27), which improved the quality of the data (see figures A-C.7 and 8 in the Supplementary Material 3). The studies included in each analysis can be found in Supplementary Material 3, section C, figures A-C.1 to 6. Furthermore, qualitative results are provided for the 67 studies, and forest plots for each ES for the original 67 studies can be found in Supplementary Material 3 A-C.1, 3, and 5. After curating the database and eliminating outliers, we calculated the Rosenberg fail-safe number, Egger regression, funnel plots, and Trim and Fill for each ES (Supplementary Material 3, section E) to detect a potential threat of publication bias. Lastly, the main results of our investigation were obtained by recomputing the full multilevel model for each ES, which showed a variable level of heterogeneity (more information in Supplementary Material 3, Appendix D), always indicating that studies are indeed estimating a different parametric effect [31]. Thus, moderator analyses were conducted, adding (one by one, as a predictor) the set of covariates of interest in a meta-regression. In this regard, moderators that appeared in less than 5 studies were removed.

Results

Description of the studies

A total of 67 studies met our inclusion criteria and were eligible for meta-analysis (table 1), obtaining a pooled sample of N = 12,147 and 150 outcomes reported. Studies were published between 2002 and 2020, most originally from the United States (53.7%). Usually, studies had a single follow-up visit, with a mean of 4.8 months. Some studies used up to 3 follow-ups, reaching a maximum of 36 months. Regarding population, participants had a pooled mean age of 60.62, and considering the standard deviation, it was always within the range of 50-70. Regarding race/ethnicity, studies conducted in the USA usually report this, with most participants being Caucasian. For studies conducted in other countries, only 8 out of 31 studies reported race or ethnic origin. Not all studies reported educational level, and when they did, there was heterogeneity in the data presented (e.g., mean years of education [32], frequency of high school completion [33]). We considered post-high school studies to be higher education and reported the percentage of participants in this category. As can be seen in Table 1, there is variability in the sample's education level. Most of the studies selected their sample based on them having a disease (62.69%), the most frequent being diabetes (25.4%; Table 1).

Regarding the measures used, the most frequent ones evaluated overall physical activity (25.3%) and moderate-vigorous physical activity (21.3%). Regarding the assessment tools used, 42.7% of measures were objective (e.g., collected through pedometers). For self-report measures (57.3%; n=86), the most used scales were IPAQ (n=26) and CHAMPS (n=12). For objective assessment tools, 54 measures were assessed with accelerometers and 9 with pedometers. Additionally, one measure employed a wristwatch, though it was unclear whether it functioned as a pedometer or accelerometer. Most of the studies' measures showed some concerns about the risk of bias (63.3%), whereas only 18% showed a low risk of bias (considering that each outcome has its own RoB evaluation, N = 150). Information for each study can be found in Supplementary Material 2, Table 1.

Interventions's characteristics

Most studies used a two-arms design, with an intervention and control group. However, some studies used up to three and four arms, evaluating a total of 90 interventions clustered in 67 studies through a multilevel approach (characteristics of interventions are presented in Supplementary Material 2, Table 2). Interventions lasted a mean of 4.7 months. Most interventions were individual (77.8%), some were in group (14.4%), and a few used both (7.8%). Regarding the delivery format, 34.4% were e-health interventions (their main component was delivered through an electronic device). More specifically, the frequency of delivery formats was: face-to-face (n=42), telephone calls (n=37), printed information (n=21), website (n=18), app (n=10), text messages (n=8), and DVDs (n=7). Regarding the person delivering the intervention (excluding interventions without people delivering it, such as e-health and print information; n=37), approximately half the interventions were delivered by professionals who received some kind of behavior change training for delivering the psychological components (52.8%), from which only 4 were delivered by a psychologist. Furthermore, only 43.3% of studies adequately described the interventions according to TIDieR criteria. Focusing on the psychological components, starting with the theoretical framework, 11 interventions did not specify any theory, whereas some studies combined different theories and others used more general frameworks (e.g., based on learning and behavioral principles). The most frequent ones were Social Cognitive Theory (n=28), Transtheoretical Model (n=13), Self-Determination Theory (n=9), and Theory of Planned Behavior (n=6). Furthermore, some studies referred to general behavioral models and other theories (n=30). Regarding BCTs, in Supplementary Material 2, Table 3 and 4, the reader can find the frequency with which each BCT and cluster was used in interventions (we count only the ones that are present in active intervention arms, not in control groups). In summary, the most frequently used BCTs were goal setting (behavior), social support (unspecified), problem-solving, and self-monitoring (behavior). It is worth noting that among the 93 techniques, only 16 appeared in more than 10 studies. Regarding the prevalence of clusters, the most common were "1. Goals and planning" and "2. Feedback and monitoring".

Intervention effects

The three models yield statistically significant pooled estimates (SMD: 0.223; SE = 0.027; CI (95%) [0.168 – 0.278]; $p < .001$; SMC: 0.359; SE = 0.056; CI (95%) [0.245 – 0.472]; $p < .001$; SMCD: 0.358; SE = 0.063; CI (95%) [0.229 – 0.488]; $p < .001$), although the within (τ^2) and between study (ω^2) variances are very small (SMD: $\tau^2 = 0$; $\omega^2 = 0.063$; SMC: $\tau^2 = 0.077$; $\omega^2 = 0.1$; SMCD: $\tau^2 = 0.064$; $\omega^2 = 0.066$). While the pooled estimate is statistically significant (ranging from SMD: 0.22 to SMCD /SMC:0.36), individual studies tend, in general, to a small ES, with confidence intervals that indeed contain the zero value. Furthermore, there were differences among estimators. For example, for SMD, the variances were higher and confidence intervals wider, with all individual studies going through zero. Meanwhile, SMC and SMCD intervals are tighter, with a higher number of studies showing individual significance (see Figure 2 for SMD and Material Supplementary 3 Appendix F for SMC and SMCD).

Moderators

To further investigate moderators' potential influence on the outcome, we added one covariate at a time (as categorical predictors), retaining its coefficient and p-value. For this analysis,

This is an Accepted Manuscript of an article published by Taylor & Francis in Health Psychology Review on 28th April 2025 available at: https://doi.org/10.1080/17437199.2025.2492042

moderators needed to be present in at least 5 studies (Table 2 shows the results). The following moderators showed a small (ranging from β = -0.12 to β = 0.37) but statistically significant coefficient predicting larger effect sizes: action planning (SMC), motivational interviewing (SMD, SMCD), the use of prompts/cues and not using a website nor sending emails (SMD).

Publication bias

Different analyses were conducted to assess publication bias for each ES; complete results can be found in Supplementary Material 3, section E. Starting with SMD, there was no evident asymmetry in the funnel plot, supported by a non-significant result in Egger's regression testing for asymmetry. However, while Rosenberg's fail-safe number indicates that the number of unpublished articles with null results needed for the effect to become non-significant is unrealistic, Orwin's calculation shows that this amount is below the reference number, indicating a potential risk of publication bias. We conducted Trim and Fill analyses, and 6 studies were filled on the left side of the plot, resulting in a pooled estimate of 0.181, SE= 0.041, CI(95%):[0.101 - 0.261], p<0.001. Thus, the effect could be even lower than the one identified in this meta-analysis, but it seems it would remain significant when accounting for publication bias. Regarding SMC and SMCD, Rosenberg and Orwin's fail-safe numbers are above the reference number. In the funnel plot, some outliers can be identified on the right, congruently with a significant result in Egger's regression in both cases. However, no outliers were identified or filled out when Trim and Fill analyses were conducted. Thus, it seems that publication bias could slightly affect these results, but not a significant effect.

Discussion

The effectiveness of interventions with psychological components promoting physical activity in retirement-age adults was studied across 67 studies. Effectiveness was tested through all possible ESs, from the simplest and most frequently used in meta-analysis (SMD) to the most complex, studying the interaction of time and group assignation (SMDC). The overall effect, excluding studies with a high risk of bias, was statistically significant but small in all cases, ranging from 0.22 to 0.36. However, looking into individual interventions in the forest plot, most confidence intervals appear wide, going through zero and indicating apparent non-effectiveness. This discrepancy shows that the effect of most lifestyle interventions is small and can only be found with great statistical power, such as the one in a meta-analysis. In this field, similar meta-analyses yield the same findings, concluding that interventions evoke a small but positive overall effect in physical activity without clear effectiveness at individual interventions [13,34]. Thus, the positive results in meta-analyses studying these types of interventions must be considered cautiously, as it is not clear whether statistical significance translates into tangible benefits for health and/or meaningful real changes in their health behaviors. Although the combined effect was effective and significant for all estimations, it must be noted that both the global effect and the individual estimations with their confidence intervals differed among estimators. For example, for SMD, the variances were higher and confidence intervals wider, with all individual studies going through zero. Meanwhile, SMC and SMCD intervals are tighter, with a higher number of studies showing individual significance. This might be due to two reasons. First, the improvements experienced by the intervention groups might be so little that compared to a control group, this difference is non-significant; however, if compared to oneself, even if improvements are small, these are consistent across participants. Thus, only when we consider the time factor do we find higher improvements. Considering this, primary studies should provide sufficient data to calculate all of the ESs' estimators, and future meta-analyses should consider different estimates, as results can vary from one estimate to another. The second reason might be the evolving standards of usual care over the years, potentially resulting in fewer differences compared to intervention groups [34].

Additionally, although finding overall effectiveness is useful to contrast whether interventions are indeed effective to some extent, this needs to be considered cautiously as lifestyle interventions addressing physical activity in this population appear to differ significantly, yielding diverse parametric effects (i.e., showing great heterogeneity) [31]. This is not surprising, as there are several possible sources of heterogeneity: the type of physical activity promoted (e.g., light, moderate, less sedentarism), the measurement (self-report, objective), the population (e.g., diabetic, at cardiac risk), the delivery of the intervention (in person, e-health, etc.), the BCTs (e.g., goal setting), theoretical framework (e.g., Social Cognitive Theory), and the interventionist (psychologist, trained, etc.). Furthermore, there were additional sources of heterogeneity, such as study-related factors (e.g., risk of bias). In this regard, we tested all of the mentioned variables as possible moderators in a meta-regression analysis. Some predictors appeared significant; however, these results must be considered cautiously as they changed according to the type of ES.

Quality of the studies

One predictor of effectiveness was the study's risk of bias, with higher bias correlating with larger ESs, congruently with prior meta-analyses [35]. Consequently, we excluded these studies from the main analyses. The significant impact of bias on ESs is concerning, particularly given the low proportion of studies with measures showing a low risk of bias (18%). These results call for a standardized approach to mitigate the risk of bias in primary research and to carefully consider the risk of bias in meta-analyses, as it can overestimate the effect of lifestyle interventions. Furthermore, we tested some other quality measures as moderators, such as the level of replicability of an intervention following its description (TIDieR), but it did not appear significant. Nevertheless, it is worth noting the low proportion of well-described interventions (43.28%). The assessment of intervention replicability and the transparency of intervention descriptions are essential for adequately identifying the components used and thus advancing the field of health intervention research. This limits the trust with which we can test the components of an intervention for effectiveness, as some of the techniques used might not be mentioned. This is a common problem in studies reporting lifestyle interventions and limits the conclusions of moderator analyses, such as meta-regressions [15].

Type of physical activity and measurements

We did not find evidence that the effectiveness depended on the type of physical activity promoted or the measurement used to evaluate it. This was an unexpected result, as some meta-analyses have found greater effectiveness in promoting physical activity compared to reducing sedentary behavior [13]. Nevertheless, studies aiming to reduce sedentary behavior were a minority in this study. Thus, we might need a larger sample of studies to find this difference. Furthermore, outcomes measured with objective tools (e.g., pedometer) were expected to show lower ESs, as participants could overestimate the activity performed when using self-reporting measures, as a previous meta-analysis has found [36]. In our analysis, this effect might have diluted as, unlike in previous meta-analyses, we included all possible measures clustered at the study's level.

Population

Population characteristics, such as the percentage of females included, or whether the intervention targeted a population with a particular disease (diabetes, cardiac risk, cancer survivors, or arthritis), did not appear as significant moderators. However, relevant variables such as education level, or race were not included in the analysis as they were reported heterogeneously, making them difficult to categorize accordingly.

Delivery of the intervention

Some delivery formats appeared to be related to lower effectiveness, such as the use of a website and sending emails (SMD). However, we did not find formats that did relate to higher effectiveness. Furthermore, we tested whether using e-health interventions (dummy-coded as whether the main intervention was delivered through electronic devices or not) could be a moderator, but it was not. In this regard, previous meta-analyses found higher effectiveness in face-to-face formats in comparison to digital interventions [14]; however, it seems that it might depend on the type of digital tools used. Further research is needed to identify the appropriate categories. Regarding whether the intervention was conducted individually or in groups, it did not appear to influence effectiveness, congruently with previous meta-analyses [35] and reviews finding effectiveness for both formats [37].

BCTs and theoretical framework

Regarding BCTs, it seems that the following are related to higher effectiveness: action planning (SMC), motivational interviewing (SMD, SMCD), and the use of prompts/cues (SMD). However, the coefficients were small for all cases. Furthermore, the number of BCTs that showed a significant association with effectiveness was limited and did not allow for testing the effectiveness of BCT combinations. When comparing these results with previous meta-analyses, we find high diversity in the BCTs found as effective, but some of them are congruent with the ones identified in our findings (action planning, prompts/cues, and motivational interviewing [13,37]). Furthermore, most meta-analyses identify as significant moderators common techniques such as "goal setting" or "self-monitoring" [14,35]. Although these techniques appear to improve effectiveness, they are the most frequently used [34] and are present in a significant proportion of the interventions (58.5-42.3%). Thus, implementing these strategies might be necessary to improve effectiveness as fundamental components of a lifestyle intervention, but not enough to achieve clinically significant effects on their own. This was supported by these results, as these BCTs did not emerge as significant predictors. Furthermore, among the 93 BCTs, only a small proportion of them were usually implemented (only 16 BCTs were used in more than 10 studies). Thus, there are many BCTs that could enhance effectiveness but cannot yet be tested. Considering these results, the exploration of a broader range of BCTs is imperative, as their potential to significantly enhance the effectiveness of lifestyle interventions remains unexplored.

In addition to identifying effective BCTs, we aimed to find whether the use of specific theoretical frameworks was related to higher or lower effect sizes. However, no theory showed a significant association with intervention effectiveness. Similarly to BCTs, the low number of theories adopted could explain these results. Furthermore, there was a lack of age-tailored theories in the included studies, which could be promising for tailoring interventions for this population. For example, the Socioemotional Selectivity Theory outlines how older adults' time perspective can directly impact the health behaviors they pursue [38,39], prioritizing goals that offer emotional rewards [40,41]. Similarly, the Selective Optimization with Compensation model outlines that older adults prioritize personally meaningful goals when resources are limited [42]. These frameworks suggest that strategies to promote and maintain health behaviors in older adults might significantly differ from those effective in younger populations. For example, planning has been identified as a particularly effective strategy, especially for individuals with a limited time perspective [43]. This aligns with our findings, as action planning was one of the few BCTs identified as a significant moderator of intervention effectiveness (i.e., was related to larger effect sizes). Future studies targeting older adults might benefit from including additional motivation constructs and theoretical frameworks, particularly those developed to address older populations' unique psychological and motivational needs.

Interventionist

Contrary to what we expected, we did not find evidence that being a psychologist or having received psychological training improved the effectiveness of the interventions, congruently with previous meta-analyses [35]. However, only 4 studies implemented interventions delivered by a psychologist, and only 52.8% of other professionals were trained. This limited representation prevents drawing a clear conclusion about the impact of psychological expertise on intervention outcomes. Furthermore, it could explain the variability in the application and reporting of BCTs within interventions, as even with tools like TIDieR and taxonomies, we found that BCTs are sometimes incorrectly implemented or identified [17,18,35]. This indicates a gap between theoretical understanding of psychology and its practical application, underscoring the need for more precise guidelines and training in the application of psychological strategies in health interventions. To address this, a collaborative effort involving psychologists and other behavior change professionals in the intervention design process is recommended to ensure the integration of comprehensive psychological principles. Thus, interventions should be designed not only considering BCTs that have been found effective in the literature -as we mentioned, these results are not robust- but rather considering a solid theoretical framework and BCTs that align with it as well as with the target population, delivery format, etc.

Strengths and limitations

This study follows a solid methodological approach that addresses common limitations faced by similar meta-analyses. For example, the numerous measures for physical activity used in primary studies usually are simplified to one due to the problem of interdependence. Thus, the most effective measures are sometimes selected, or a mean of ESs for the different measures is calculated [44]. The same issue takes place when a study uses more than one active arm, usually selecting the most intensive intervention [44]. By building a multilevel model, this meta-analysis meticulously evaluated the different measures reflecting physical activity and the plurality of intervention arms. Furthermore, different ES estimates were calculated, considering the time effect, group-assignment effect, and the interaction of both, providing a multi-faceted understanding of the effectiveness and moderator effects. Nevertheless, acknowledging the limitations is crucial, particularly the timing of the search strategy, which may not encompass the most recent studies. This time-lapse is frequent in meta-analyses [34], highlighting this methodology's meticulous and time-consuming nature in collecting and analyzing extensive datasets. Furthermore, there are significant limitations in the included evidence, as interventions are not usually well described, and studies usually present some risk of bias. Thus, these results should be considered cautiously.

Additionally, even though we followed a robust and validated taxonomy to identify BCTs, some additional strategies may not have been well-documented. Motivational interviewing was included because it was frequently mentioned as a central intervention component, though often without sufficient detail to identify the implemented BCTs. However, omitting potential newer BCTs represent a limitation. Future studies might consider the inclusion of additional BCTs [45,46] , such as streaking [47] or social prescribing [48].

Finally, the lack of data on participants' retirement status or duration is an additional limitation. Although the inclusion criteria targeted adults around retirement age, few studies focused explicitly on retired adults or provided detailed information on this variable. This limited the possibility of exploring retirement timing as a potential moderator, an area requiring further investigation.

By addressing these limitations, future meta-analyses and primary studies may improve the design and implementation of interventions targeting this demographic and promoting physical activity.

Conclusions

Lifestyle interventions promoting physical activity at the retirement age have small positive effects. Thus, improving existing interventions is needed to ensure more significant and meaningful changes. Furthermore, when trying to identify significant predictors for success, we have found that these differ according to the ES' estimate and that might not be robust. Furthermore, a high risk of bias predicted success in interventions. In conclusion, despite the number of published RCTs promoting physical activity, more primary studies with a low risk of bias and with a comprehensive design of interventions implementing adequate BCTs and theories are needed to yield robust meta-analytic results on effective psychological components for promoting physical activity.

Acknowledgments

We thank Ricardo Olmos and Laura Bermejo for their impeccable advice. We also thank the SO-NUTS consortium for designing and developing this project.

Data availability statement

The data that support the findings of this study are openly available in Figshare https://figshare.com/s/dfc03eabcdda5b1cf05e; https://figshare.com/s/e20fb8c9c0eedae61fc7.

Declaration of interest statement

The authors declare no conflict of interest.

Funding

The SO-NUTS project is funded by JPI HDHL PREPHOBES call, the funding agencies supporting this work are: the Spanish Agencia Estatal de Investigación (AEI: PCI2020-120683-2; AEI/10.13039/501100011033), the Netherlands Organisation for Health Research and Development (ZonMw), French national Research Agency (ANR), Federal Ministry of Education, Science and Research represented by the Austrian Research Promotion Agency (BMBWF represented by FFG), Ministry of Education, Youth and Sports Department of Research and Development (MSMT). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the ERA-NET Cofund action N° 727565.

Author Contribution

P.C.C.: Conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing-original draft, writing-review editing; I.E.: Data curation, formal analysis, methodology, software, visualization, writing-original draft, writing-review editing; J.S.: Investigation, writing-review editing; N.A.: Investigation, writing-review editing; J.B.: methodology, writing-review editing; K.J.: Investigation, writing-review editing; C.J.: Investigation, writing-review editing; V.L.: Investigation, writing-review editing; A.L.: Investigation, writing-review editing; M.M.: Investigation, writing-review editing; A.T.: Investigation, writing-review editing M.V.: Investigation, writing-review editing; A.C.J.: Conceptualization, funding acquisition, investigation, project administration, supervision, writing-review editing; M.S.I.: Conceptualization, data curation, investigation, methodology, project administration, supervision, writing-original draft, writing-review editing

References

1. Permanyer I, Bramajo O: El aumento de la longevidad en Europa: ¿Añadiendo años a la vida o vida a los años? Perspect Demogràfiques. 2022; 1–4.

2. Wullems JA, Verschueren SMP, Degens H, Morse CI, Onambélé GL: A review of the assessment and prevalence of sedentarism in older adults, its physiology/health impact and non-exercise mobility counter-measures. Biogerontology. 2016; 17:547–565.

3. Musich S, Wang SS, Hawkins K, Greame C: The Frequency and Health Benefits of Physical Activity for Older Adults. Popul Health Manag. 2017; 20:199–207.

4. Fernández-Ballesteros R, Valeriano-Lorenzo E, Sánchez-Izquierdo M, Botella J: Behavioral Lifestyles and Survival: A Meta-Analysis. Front Psychol. 2021; 12:786491.

5. Organization WH: World Report on Ageing and Health. World Health Organization,

2015.

6. Murtagh EM, Murphy MH, Murphy NM, Woods C, Nevill AM, Lane A: Prevalence and correlates of physical inactivity in community-dwelling older adults in Ireland. PloS One. 2015; 10:e0118293.

7. Smeaton D, Barnes H, Vegeris S: Does Retirement Offer a "Window of Opportunity" for Lifestyle Change? Views From English Workers on the Cusp of Retirement. J Aging Health. 2017; 29:25–44.

8. Zantinge EM, van den Berg M, Smit HA, Picavet HSJ: Retirement and a healthy lifestyle: opportunity or pitfall? A narrative review of the literature. Eur J Public Health. 2014; 24:433–439.

9. Tam ACT, Steck VA, Janjua S, et al.: A systematic review of evidence on employment transitions and weight change by gender in ageing populations. PLOS ONE. 2022; 17:e0273218.

10. Baxter S, Johnson M, Payne N, et al.: Promoting and maintaining physical activity in the transition to retirement: a systematic review of interventions for adults around retirement age. Int J Behav Nutr Phys Act. 2016; 13:12.

11. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF: Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: A systematic review and meta-analysis. Soc Sci Med 1982. 2017; 192:125–133.

12. Webb T, Joseph J, Yardley L, Michie S: Using the Internet to Promote Health Behavior Change: A Systematic Review and Meta-analysis of the Impact of Theoretical Basis, Use of Behavior Change Techniques, and Mode of Delivery on Efficacy. J Med Internet Res. 2010; 12:e1376.

13. Howlett N, Trivedi D, Troop NA, Chater AM: Are physical activity interventions for

healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med. 2019; 9:147–157.

14. Carraça E, Encantado J, Battista F, et al.: Effective behavior change techniques to promote physical activity in adults with overweight or obesity: A systematic review and meta-analysis. Obes Rev Off J Int Assoc Study Obes. 2021; 22 Suppl 4:e13258.

15. Hoffmann TC, Glasziou PP, Boutron I, et al.: Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014; 348:g1687.

16. Michie S, Richardson M, Johnston M, et al.: The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. Ann Behav Med. 2013; 46:81–95.

17. 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.

18. Rovniak LS, Hovell MF, Wojcik JR, Winett RA, Martinez-Donate AP: Enhancing Theoretical Fidelity: An E-mail—Based Walking Program Demonstration. Am J Health Promot. 2005; 20:85–95.

19. Fernández-Castilla B, Jamshidi L, Declercq L, Beretvas SN, Onghena P, Van den Noortgate W: The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. Behav Res Methods. 2020; 52:2031–2052.

20. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021; 372:n71.

21. Armstrong M, Winnard A, Chynkiamis N, Boyle S, Burtin C, Vogiatzis I: Use of

pedometers as a tool to promote daily physical activity levels in patients with COPD: a systematic review and meta-analysis. Eur Respir Rev. 2019; 28:.

22. Hobbs N, Godfrey A, Lara J, et al.: Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. BMC Med. 2013; 11:75.

23. Collazo-Castiñeira P, Sánchez-Izquierdo M, Reiter LJ, et al.: Analysis of behavioral change techniques used in exercise and nutritional interventions targeting adults around retirement age with sarcopenic obesity in a systematic review. Arch Gerontol Geriatr. 2024; 123:105437.

24. Leyrat C, Caille A, Eldridge S, Kerry S, Dechartres A, Giraudeau B: Intervention effect estimates in cluster randomized versus individually randomized trials: a meta-epidemiological study. Int J Epidemiol. 2018; 48:609–619.

25. Donner A, Klar N: Issues in the meta-analysis of cluster randomized trials. Stat Med. 2002; 21:.

26. Sterne JAC, Savović J, Page MJ, et al.: RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019; 366:l4898.

27. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2023; .

28. Viechtbauer W: Conducting Meta-Analyses in R with the metafor Package. J Stat Softw. 2010; 36:1–48.

29. Hedges LV, Tipton E, Johnson MC: Robust variance estimation in meta-regression with dependent effect size estimates. Res Synth Methods. 2010; 1:39–65.

30. Pustejovsky JE, Tipton E: Meta-analysis with robust variance estimation: Expanding the range of working models. Prev Sci. 2022; 23:425–438.

31. Botella J, Sánchez-Meca J: Meta-análisis en ciencias sociales y de la salud. Síntesis,

2015.

32. Lorig KR, Ritter PL, González VM: Hispanic chronic disease self-management: a randomized community-based outcome trial. Nurs Res. 2003; 52:361–9.

33. Collins TC, Lu L, Valverde MG, Silva MX, Parra-Medina D: Efficacy of a multi-component intervention to promote physical activity among Latino adults: A randomized controlled trial. Prev Med Rep. 2019; 16:.

34. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R: mHealth Technologies to Influence Physical Activity and Sedentary Behaviors: Behavior Change Techniques, Systematic Review and Meta-Analysis of Randomized Controlled Trials. Ann Behav Med Publ Soc Behav Med. 2017; 51:226–239.

35. Samdal GB, Eide GE, Barth T, Williams G, Meland E: Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. Int J Behav Nutr Phys Act. 2017; 14:42.

36. Chase J-AD: Interventions to Increase Physical Activity Among Older Adults: A Meta-Analysis. The Gerontologist. 2015; 55:706–718.

37. Greaves CJ, Sheppard KE, Abraham C, et al.: Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health. 2011; 11:119.

38. Hill KC, Allemand M, Hill P: Daily Limited Future Time Perspective Is Associated With More Health Behavior Within Older Adults. J Gerontol B Psychol Sci Soc Sci. 2024; 79:gbae161.

39. Stahl ST, Patrick JH: Adults' future time perspective predicts engagement in physical activity. J Gerontol B Psychol Sci Soc Sci. 2012; 67:413–416.

40. Carstensen LL, Isaacowitz DM, Charles ST: Taking time seriously. A theory of socioemotional selectivity. Am Psychol. 1999; 54:165–181.

41. Carstensen LL: Social and emotional patterns in adulthood: Support for socioemotional selectivity theory. Psychol Aging. 1992; 7:331–338.

42. Baltes PB, Baltes MM: Psychological perspectives on successful aging: The model of selective optimization with compensation. In: Baltes MM, Baltes PB, editors. Successful Aging: Perspectives from the Behavioral Sciences. Cambridge: Cambridge University Press, 1990. p. 1–34.

43. Gellert P, Ziegelmann JP, Lippke S, Schwarzer R: Future time perspective and health behaviors: temporal framing of self-regulatory processes in physical exercise and dietary behaviors. Ann Behav Med Publ Soc Behav Med. 2012; 43:208–218.

44. Cooper H: Research synthesis and meta-analysis: A step-by-step approach. Sage publications, 2015.

45. Knittle K, Heino M, Marques MM, et al.: The compendium of self-enactable techniques to change and self-manage motivation and behaviour v.1.0. Nat Hum Behav. 2020; 4:215–223.

46. Kok G, Gottlieb NH, Peters G-JY, et al.: A taxonomy of behaviour change methods: an Intervention Mapping approach. Health Psychol Rev. 2016; 10:297–312.

47. Curran M, Larade N, Özakinci G, Tymowski-Gionet G, Dombrowski SU: Look, over there! A streaker! – Qualitative study examining streaking as a behaviour change technique for habit formation in recreational runners. Health Psychol Behav Med. 2024; 12:2416505.

48. Cunningham KB, Rogowsky RH, Carstairs SA, Sullivan F, Ozakinci G: Social prescribing and behaviour change: proposal of a new behaviour change technique concerning the "connection" step. Health Psychol Behav Med. 2022; 10:121–123.

This is an Accepted Manuscript of an article published by Taylor & Francis in Health Psychology Review on 28th April 2025 available at: https://doi.org/10.1080/17437199.2025.2492042

Tables

Table 1. Basic characteristics of the included studies.

Author, year, reference	Country	N	#fu (month)	Population	Race (%)	% females	% higher education
Hartman et al. (2016) [1]	USA	54	1(6)	Women at cancer risk	91.7-100% white	100	39.4-66.7
Lyons et al. (2017) [2]	USA	40	1(3)	General	65% white, 13% African American, 28% Latin 15% other	85	68
Thomsen et al. (2017) [3–5]	Denmark	150	2(4,22)	Arthritis	x	81	66
Patel et al. (2017) [6]	USA	200	2(3,6)	General	100% white	56	93.5-94.8
Mayer et al. (2018) [7]	USA	284	1(6)	Cancer survivors	89% white, 11% other	51,7	79.2
Krebs et al. (2017) [8]	USA	85	1(3)	Cancer survivors	81% white, 6% African American, 6% Latin, 2% Asian, 5% other	96	92
James et al. (2017) [9]	Australia	203	1(12)	General	Origin (81.3% Australia, 18.7% other)	70	59.1
Scott et al. (2019) [10]	UK	35	2(3,6)	Chronic diseases (diabetes, cardiac)	97.1% white british, 2.9% black caribean	43	x
Chapman et al. (2018) [11]	Australia	101	2(1,3)	Women, cancer survivors	Origin (79.2% Australia; 7.9% British, 12.9% other)	100	80.2
Balducci et al. (2019) [12,13]	Italy	300	3(12,24, 36)	Diabetes	x	39	32
Wirth et al. (2019) [14]	USA	36	2(1,4)	Cancer survivors	x	72	x
Robin et al. (2019) [15]	France	85	2(0.5,1)	General	98.8% white	65,9	50-55
Poppe et al. (2019) [16]	Belgium	54	1(1.25)	Diabetes	x	37	54
Poppe et al. (2019) [16]	Belgium	63	1(1.25)	General	x	75	60.32
Schlenk et al. (2020) [17]	USA	182	2(6,12)	Arthritis, cardiac conditions.	73.1% white	73	79
Liu & Lachman (2020) [18]	USA	60	2(1,2)	General	75% white	75	90.9
Bennell et al. (2020) [19]	Australia	110	1(6)	Arthritis	x	67	x
Mailey et al. (2020) [20]	USA	308	1(0.5)	General	95.5% white, 1,6% African American, 1.9% Asian, 1.6% other	83	93.8
Lo et al. (2020) [21]	Taiwan	43	1(3)	Multiborbility (cardiac conditions, diabetes)	x	51,2	37.2
Peacock et al. (2020) [22]	UK	204	2(3,12)	Cardiac conditions	88% white british	36	40
De Greef et al. (2010) [23]	Belgium	41	2(3,12)	Diabetes	x	31,7	70.7
Lorig et al. (2003) [24]	USA	551	1(4)	Cardiac	100% Latin	79	7.6 years
Plotnikoff et al. (2013) [25]	Canada	287	1(12)	Diabetes	Origin (76% Canada, 13.2% European, 4.2% Asia, 6.6% other)	46,3	42.7-47.9
Keyserling et al. (2002) [26]	USA	200	2(6,12)	African American women with diabetes	100% African American	100	10-11 years
Mouton et al. (2014) [27]	Belgium	149	1(12)	General	x	63,9	x
Di Loreto et al. (2003) [28]	Italy	340	1(24)	Diabetes	x	52,6	x
Hirschey et al. (2016) [29]	USA	148	2(1,12)	Cancer survivors	86.5% white	52,1	59-67

This is an Accepted Manuscript of an article published by Taylor & Francis in Health Psychology Review on 28th April 2025 available at: https://doi.org/10.1080/17437199.2025.2492042

Hageman et al. (2005) [30]	USA	30	1(3)	Women	93.6% white, 3.2% African American, 3.2% other	100	86.7
Pekmezi et al. (2017) [31]	USA	84	1(6)	African American women	100% African American	100	56
Young et al. (2019) [32]	USA	67	1(6)	Diabetes	40% white, 12% African American, 33% Latin, 15% other	61	92.5
Collins et al. (2019) [33]	USA	69	1(3)	Race, disease; Latin adults with cardiac conditions	100% Latin	86	84,1 completed highschool/
Moore et al. (2006) [34]	USA	250	1(12)	Cardiac conditions	81.2% white, 16.8 african american, 2% other	48,9	14 years
Irvine et al. (2013) [35]	USA	368	2(3,6)	General	59% white, 12,2%African American, 11,4% Asian American, 9,5% Latin, 7.9% other	69	82
Höchsmann et al. (2019) [36,37]	Switzerland	36	1(6)	Diabetes	x	47,2	x
Focht et al. (2014) [38]	USA	80	2(3,12)	Arthritis	68.8% white, 25% African American, 3.8% Latin, 2.5% Other	84	57.5
Kim & Kang (2006) [39]	Korea	73	1(3)	Diabetes	x	57	x
Callahan et al. (2014) [40]	USA	339	1(1.25)	Arthritis	78.17% white	83,4	53.3-62.5
Bennett et al. (2007) [41]	USA	56	2(3,6)	Cancer survivors	98.2% white, 1.8% Latin	89,3	73
Bennett et al. (2008) [42]	USA	72	1(6)	Adults living in rural areas	95.8% white, 4.2% other	89,9	40.5-42.9
Allen et al. (2008) [43]	USA	52	1(2)	Diabetes	90.4% white, 9.6% African american	51,9	63.5
Keyserling et al. (2008) [44]	USA	236	2(6,12)	Women with low economic status	58% white, 41% African American	100	78-80 completed grade 12
Butler et al. (2009) [45]	Australia	110	2(1.5,6)	Cardiac conditions	x	24,5	61,8-63,6 secondary level of education or less
Costanzo et al. (2006) [46,47]	USA	46	1(3)	Women	98% white, 2% other	100	52
De Greef et al. (2011) [48]	Belgium	92	2(6,12)	Diabetes	x	31	x
Van Keulen et al. (2011) [49,50]	Netherlands	1629	3(6.25,11.75,18.25)	Cardiac conditions	x	45	23
Toobert et al. (2011) [51]	USA	280	2(6,12)	Latin women with diabetes	100% Latin	100	28.6-30.1
Sawchuk et al. (2011) [52]	USA	36	1(6)	American Indian/ Alaska native adults	100% American Indian/Alaska Native	69	56
Buman et al. (2011) [53]	USA	81	2(4,18)	General	91.4% white, 3.7% African American, 1.2% Asian, 2.5% other	83	93.8
Strath et al. (2011) [54]	USA	81	1(3)	General	x	83	73
White et al. (2012) [55]	Australia	183	2(1.25,2.5)	Diabetes	99% white	68	x
Migneault et al. (2012) [56]	USA	337	3(4,8,12)	African Americans with cardiac conditions	100% African American	70,5	12.2 years
Ruiz et al. (2012) [57]	USA	21	1(1)	General	45.8% white, 54.2% African American	4	x
Bélanger-Gravel et al. (2013) [58]	Canada	101	2(2,8)	General	x	59	55.4

Hawkes et al. (2013) [59]	Australia	410	2(6,12)	Cancer survivors	x	46	89.8-91.7 high school
Sher et al. (2014) [60]	USA	80	3(61218	Cardiac conditions	65% white, 18.8% African American, 8.8% Latin, 6.2% Asian, 1.2% unknown	13	x
Griffin et al. (2014) [61]	UK	478	1(12)	Diabetes	97.5% white	37,7	61.4-61.9 Education after 16
Knittle et al. (2015) [62]	Netherlands	78	2(1.5,6)	Arthritis	x	66,7	16.7
Lilienthal et al. (2014) [63]	USA	86	2(1,6)	General	100% white	66	79.2-83.8
Trinh et al. (2014) [64]	Canada	32	1(3)	Cancer survivors	x	50	68.7
Bossen et al. (2013) [65]	Netherlands	199	2(3,12)	Arthritis	x	65	45.7
James et al. (2015) [66,67]	Australia	133	2(2,5)	Cancer survivors	x	77,4	71.9-73
Lin et al. (2016) [68]	Taiwan	115	1(3)	Women with cardiac conditions	x	100	x
Da Silva et al. (2015) [69]	Brazil	30	1(2)	Diabetes	x	66,7	x
Almeida et al. (2015) [70]	USA	452	1(1)	Cardiac	76.8% White, 11,8% African American, 11,4% Latin	59	40,1 highschool or less
Nahm et al. (2017) [71]	USA	866	1(2)	General	89.6 % White, 7.4% African American, 3% Other	37	87.2
Müller et al. (2016) [72]	Malasya	43	2(3,6)	General	x	74	68
Silfee et al. (2016) [73]	USA	23	1(1.25)	Diabetes	73.9%white, 21.7% African American, 4.4% other	69,6	95.7

Note. N: sample size at baseline; #fu(month): number of follow-ups and the months from baseline when each follow-up occurred. %higher studies: The percentage of the sample that had studies higher than high school. When this information was unavailable, other information was provided. An "x" is displayed when information is not reported. References for the included papers can be found in Supplementary Material 2.

Table 2. Results of the multilevel models adding one covariate at a time. Only significant covariates are shown.

SMD			SMC			SMCD		
Covariate	*β*	p	Covariate	*β*	p-value	Covariate	*β*	p-value
Prompts/cues	0.23	.003	Action Planning	0.27	0.015	MI	0.37	0.002
Web	-0.12	.016						
MI	0.17	.017						
Emails	-0.17	.027						

Note. MI: Motivational Interviewing

Figures

Figure 1. PRISMA flow chart

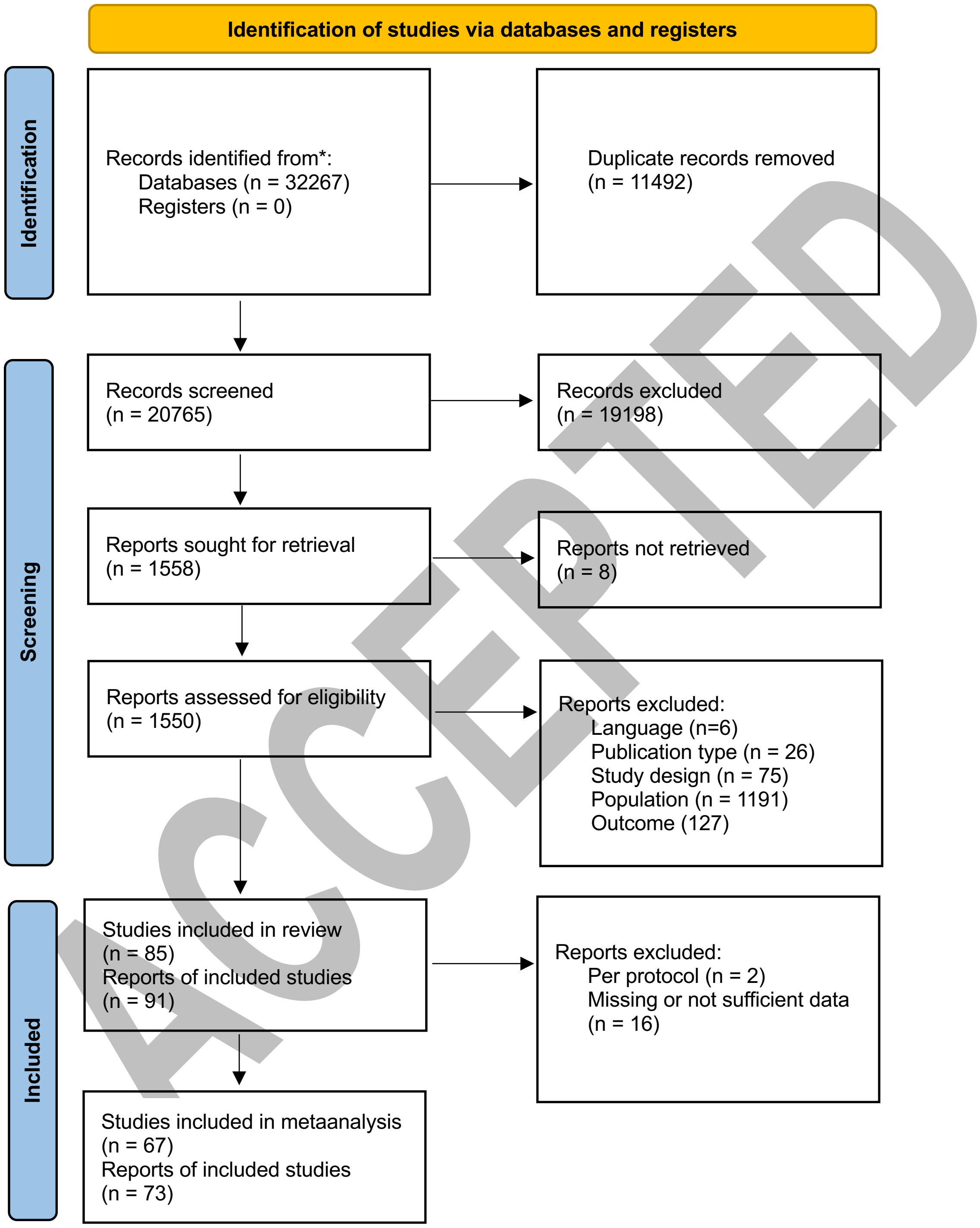

Figure 2. Forest plot for the aggregated data, standardized mean difference (SMD). Positive differences mean intervention group did better than control group.